



THÈSE

pour obtenir le grade de
DOCTEUR EN INFORMATIQUE

présentée et soutenue par

Guillaume CERUTTI

le 21 novembre 2013

**Segmentation et interprétation d'images
naturelles pour l'identification de feuilles
d'arbres sur smartphone**

JURY

Jean-Philippe DOMENGER	Rapporteur	Professeur des Universités	LaBRI, Bordeaux
Anne VERROUST-BLONDET	Rapporteur	Chargée de Recherche	INRIA, Rocquencourt
Jean-Yves RAMEL	Examineur	Professeur des Universités	LI, Tours
Erhan ABDULLAH	Examineur	Assistant Professor	Okan, Istanbul, Turquie
Pierre BONNET	Examineur	Cadre Scientifique	CIRAD, Montpellier
Laure TOUGNE	Directeur	Professeure des Universités	LIRIS, Lyon
Didier COQUIN	Directeur	Professeur des Universités	LISTIC, Annecy
Antoine VACAVANT	Directeur	Maître de Conférences	ISIT, Le Puy-en-Velay

RÉSUMÉ



LES espèces végétales, et en particulier les espèces d'arbres, forment un cadre de choix pour un processus de reconnaissance automatique basé sur l'analyse d'images. Les critères permettant de les identifier sont en effet le plus souvent des éléments morphologiques visuels, bien décrits et référencés par la botanique, qui laissent à penser qu'une reconnaissance par la forme est envisageable. Les feuilles constituent dans ce contexte les organes végétaux discriminants les plus faciles à appréhender, et sont de ce fait les plus communément employés pour ce problème qui connaît actuellement un véritable engouement. L'identification automatique pose toutefois un certain nombre de problèmes complexes, que ce soit dans le traitement des images ou dans la difficulté même de la classification en espèces, qui en font une application de pointe en reconnaissance de formes.


Cette thèse place le problème de l'identification des espèces d'arbres à partir d'images de leurs feuilles dans le contexte d'une application pour smartphones destinée au grand public. Les images sur lesquelles nous travaillons sont donc potentiellement complexes et leur acquisition peu supervisée. Nous proposons alors des méthodes d'analyse d'images dédiées, permettant la segmentation et l'interprétation des feuilles d'arbres, en se basant sur une modélisation originale de leurs formes, et sur des approches basées modèles déformables. L'introduction de connaissances *a priori* sur la forme des objets améliore ainsi de façon significative la qualité et la robustesse de l'information extraite de l'image. Le traitement se déroulant sur l'appareil, nous avons développé ces algorithmes en prenant en compte les contraintes matérielles liées à leur utilisation.

Nous introduisons également une description spécifique des formes des feuilles, inspirée par les caractéristiques déterminantes recensées dans les ouvrages botaniques. Ces différents descripteurs fournissent des informations de haut niveau qui sont fusionnées en fin de processus pour identifier les espèces, tout en permettant une interprétation sémantique intéressante dans le cadre de l'interaction avec un utilisateur néophyte. Les performances obtenues en termes de classification, sur près de 100 espèces d'arbres, se situent par ailleurs au niveau de l'état de l'art dans le domaine, et démontrent une robustesse particulière sur les images prises en environnement naturel.

Enfin, nous avons intégré l'implémentation de notre système de reconnaissance dans l'application **Folia** pour iPhone, qui constitue une validation de nos approches et méthodes dans un cadre réel.

MOTS-CLÉS : Reconnaissance de plantes, segmentation d'images, modèles déformables, contours actifs, *a priori* de forme, curvature-scale space, classification, application smartphone.

ABSTRACT

 PLANT species, and especially tree species, constitute a well adapted target for an automatic recognition process based on image analysis. The criteria that make their identification possible are indeed often morphological visual elements, which are well described and referenced by botany. This leads to think that a recognition through shape is worth considering. Leaves stand out in this context as the most accessible discriminative plant organs, and are subsequently the most often used for this problem recently receiving a particular attention. Automatic identification however gives rise to a fair amount of complex problems, linked with the processing of images, or with the difficult nature of the species classification itself, which make it an advanced application for pattern recognition.

This thesis considers the problem of tree species identification from leaf images within the framework of a smartphone application intended for a non-specialist audience. The images on which we expect to work are then potentially very complex scenes and their acquisition rather unsupervised. We consequently propose dedicated methods for image analysis, in order to segment and interpret tree leaves, using an original shape modelling and deformable templates. The introduction on prior knowledge on the shape of objects enhances significantly the quality and the robustness of the information we extract from the image. All processing being carried out on the mobile device, we developed those algorithms with concern towards the material constraints of their exploitation.

We also introduce a very specific description of leaf shapes, inspired by the determining characteristics listed in botanical references. These different descriptors constitute independent sources of high-level information that are fused at the end of the process to identify species, while providing the user with a possible semantic interpretation. The classification performance demonstrated over approximately 100 tree species are competitive with state-of-the-art methods of the domain, and show a particular robustness to difficult natural background images.

Finally, we integrated the implementation of our recognition system into the **Folia** application for iPhone, which constitutes a validation of our approaches and methods in a real-world use.

KEYWORDS : Plant recognition, image segmentation, deformable templates, active contours, shape priors, curvature-scale space, classification, smartphone application.

REMERCIEMENTS

Je voulais tout d'abord dire un grand merci à Laure, Didier et Antoine pour tout ce que nous avons partagé pendant ces trois années de thèse, et pour tout ce qu'ils ont pu m'apporter. Le fait d'être trop éloignés de certains (et peut-être trop près d'autre(s) !) n'a pas toujours facilité les choses, mais je crois que nous avons quand même réussi à construire quelque chose d'intéressant ensemble, et je suis en tout cas vraiment content d'avoir effectué ma thèse sous leur supervision à tous les trois.

Merci également aux membres de mon jury, M. Jean-Yves Ramel, Mme Anne Verroust-Blondet, M. Jean-Philippe Domenger, M. Erchan Aptoula et M. Pierre Bonnet, pour avoir accepté (sans doute sans savoir ce qui les attendait !) de s'atteler à ce copieux manuscrit, pour l'intérêt qu'ils ont montré pour son contenu, et pour les discussions enrichissantes que nous avons pu partager.

J'ai aussi une pensée particulière pour tous ceux qui ont pu m'aider dans mes travaux au sein du projet ReVeS, et avec qui j'ai vraiment apprécié d'avancer, de réfléchir ensemble, et surtout d'élargir mes horizons. Un grand merci donc à Violaine, à Julien, à Céline et Pierre-Olivier, à Houghui et à Rui, pour leur aide précieuse, pour les séances d'éclaircissements mutuels, et pour m'avoir permis de partager un peu de leur science.

Des remerciements en pagaille, évidemment, pour toutes les personnes du LIRIS Lyon 2 dont j'ai partagé pendant trois ans, dans ces locaux de prestige, le quotidien, la science du fléchiverbisme, la joie des chantiers, la cuisine thaïlandaise, les séances de répétitions subies, les matins chocolatés, et surtout la bonne humeur généralisée ! Et un gros merci particulier à Catherine pour sa disponibilité, son énergie et sa gentillesse.

Enfin, il me faut quand même remercier tous ceux, famille et amis, qui m'ont permis de me rappeler que tout ceci n'était qu'une thèse et de garder les priorités là où elles doivent être. Merci à tous, et un peu plus spécialement à mes parents et à ma sœur pour me faire me sentir toujours "à la maison" où qu'on soit, aux cousins pour la folie et les fous rires, à François pour la dolce vita de tous les jours, à Claire-Line pour l'énergie et la joie de vivre, à Julie pour le soutien à distance, à Julia pour la présence et pour l'espoir.

Table des matières

Introduction	1
1 L'identification en botanique	9
1.1 Taxinomie et clés de détermination	12
1.1.1 Nomenclature des espèces végétales	12
1.1.2 Classification phylogénétique et conséquences	15
1.1.3 Les clés de détermination en botanique	19
1.2 Petit lexique botanique appliqué aux feuilles	23
1.2.1 Qu'est-ce qu'une feuille ?	23
1.2.2 Descriptions des formes foliaires	26
1.2.3 Formes et espèces	33
1.3 La botanique comme guide pour la reconnaissance automatique	34
1.3.1 L'analyse par la forme	34
1.3.2 Des difficultés du problème d'identification	36
2 Etat de l'art	41
2.1 Traitement d'images pour la reconnaissance de plantes	43
2.2 Extraction d'objet et segmentation d'image	44
2.2.1 Extraction manuelle d'objet	44
2.2.2 Traitement global de l'image	46
2.2.3 Algorithmes de segmentation plus évolués	49
2.2.4 Détection de points d'intérêt	51
2.3 Caractérisation des objets extraits	52
2.3.1 Mesures géométriques simples	52
2.3.2 Descripteurs statistiques de forme	54
2.3.3 Caractérisation géométrique de contour	59
2.3.4 Modélisation explicite d'organes végétaux	60
2.3.5 Représentations de couleurs et textures	61
2.3.6 Descriptions de contextes locaux	62

2.3.7	Histogrammes de vocabulaire appris	64
2.4	Apprentissage automatique et recherche d'images	65
2.4.1	Comparaison et recherche d'images par le contenu	65
2.4.2	Méthodes de classification automatique	66
2.5	Cadre de travail et parti pris scientifique	67
2.5.1	Segmentation robuste d'images naturelles	68
2.5.2	Description géométrique inspirée par la botanique	69
2.5.3	Apprentissage automatique de nombreuses classes	69
3	Représentations paramétriques de feuilles	75
3.1	Modèle polygonal de feuille simple	77
3.2	Modèle polygonal de feuille palmatilobée	83
3.3	Modèle synthétique de feuille composée	87
3.4	Modèles polygonaux joints de folioles	90
3.5	Modèles de formes basales et apicales	93
3.6	Utilités des modélisations de formes	97
3.6.1	Segmentation avec <i>a priori</i> de forme	97
3.6.2	Pertinence de la description	98
3.6.3	Interprétation sémantique	98
3.6.4	Reconstruction de formes	98
3.6.5	Cohérence de la connaissance	99
4	Segmentation des images de feuilles	101
4.1	Segmentation d'objets en environnement complexe	103
4.1.1	Estimation de gradient et détection de contours	103
4.1.2	Méthodes de segmentation basées région	104
4.1.3	Modèles de contours actifs	105
4.1.4	Modèles déformables avec <i>a priori</i> de formes	107
4.1.5	Adaptation aux images de feuilles en environnement naturel	107
4.2	Modélisation de la couleur et cartes de dissimilarité	109
4.2.1	Distribution des couleurs et critère d'homogénéité	109
4.2.2	Variations locales de la couleur	116
4.2.3	Combinaisons de cartes de dissimilarité	122
4.3	Modèles paramétriques déformables de feuilles	127
4.3.1	Minimisation d'énergie	128
4.3.2	Un polygone actif pour les feuilles simples et palmatilobées	130
4.3.3	Modèles déformables de feuilles composées	139
4.4	Contours actifs guidés par les modèles de forme	145

4.4.1	Energie de guidage par la forme	145
4.4.2	Contours actifs guidés	148
4.4.3	Contours guidés multiples de folioles	153
4.5	Résultats et performances de segmentation	155
4.5.1	Mesures de qualité	156
4.5.2	Méthodes implémentées pour comparaison	159
4.5.3	Résultats et comparaisons	161
5	Description structurelle du contour	173
5.1	Interprétation du contour des feuilles	175
5.1.1	Partition sémantique du contour	175
5.1.2	Analyse du <i>Curvature-Scale Space</i>	178
5.1.3	Détection et description de points caractéristiques	182
5.2	Mots de bord	188
5.2.1	Une représentation structurelle pour le bord des feuilles	188
5.2.2	Outils de manipulation des chaînes	189
5.2.3	Application à l'identification des espèces	198
5.3	Description des formes basales et apicales	203
5.3.1	Modèles paramétriques déformables	203
5.3.2	Réponse au problème du pétiole	207
5.4	Pertinence de la description	210
6	Classification et fusion de données	215
6.1	Modèles d'espèces par agrégation	217
6.1.1	Modèles gaussiens d'espèces par descripteur	218
6.1.2	Une distance adaptée à la variabilité des classes	219
6.1.3	Combinaison de descripteurs par fusion de confiances	224
6.1.4	Contexte d'évaluation	227
6.1.5	Résultats et comparaisons	229
6.2	Etiquetage par forme et sous-classification	242
6.2.1	Apprentissage de classifieurs de formes foliaires	243
6.2.2	Modèles de fusion de sous-classifieurs	249
6.2.3	Performances et intérêts	256
6.3	Classification par forêts aléatoires	259
6.3.1	Modèles de classification	259
6.3.2	Sélection d'attributs et optimisation de paramètres	260
6.3.3	Résultats de classification	263

7	Intégration de données biogéographiques	269
7.1	Cartographie des facteurs biogéographiques	271
7.1.1	Paramètres écologiques déterminants	272
7.1.2	Indices intégrateurs	275
7.2	Mesures de plausibilité des espèces	277
7.2.1	Répartition phytogéographique théorique	277
7.2.2	Observations localisées et distribution empirique	279
7.3	Fusion de données et résultats	283
	Conclusions	289
A	Base de données botaniques	301
B	L'application Folia pour iPhone	333
B.1	Modèle de développement de l'application	335
B.1.1	Architecture générale	335
B.1.2	Déroulement du processus	336
B.2	Présentation de Folia	339
B.2.1	Processus d'identification	339
B.2.2	Fonctionnalités complémentaires	341
B.2.3	Perspectives d'extension	342

Introduction

Enjeux et vocations de l'identification automatique des végétaux



Acer campestre L.

“ ... car au fond, personne ne sait le nom des fleurs qu’il y a dans le jardin le plus simple.

”

Boris VIAN



La place des plantes dans notre monde moderne a indéniablement évolué, et de manière conséquente, au fil des décennies passées, qui les ont vu s'éloigner du quotidien de tout un chacun. Pour autant, la somme des connaissances accumulées sur le règne végétal par des générations de botanistes, remontant à l'antiquité grecque, est colossale, et ses applications dans la vie des hommes pour ainsi dire infinies.

Les espèces, briques de base du vivant, sont intégralement définies, référencées, décrites, organisées, en un mot parfaitement appréhendées par la botanique, mais elles demeurent une notion peu perceptible par une majorité de gens pour qui leur signification et leurs implications se sont estompées. L'expertise permettant l'identification d'une espèce n'en demeure pas moins un savoir concret, accessible à qui se donne les moyens et le temps de se plonger dans les arcanes de la botanique.

Cette expertise laisse à penser que les végétaux peuvent constituer un objet parfaitement adapté à un système de reconnaissance automatique, à une machine capable d'intégrer cette connaissance pour parvenir à prendre des décisions quant à l'appartenance d'un individu qui lui est présenté à telle ou telle espèce. Les défis à relever pour y parvenir sont immenses, au regard tant de l'acquisition des organes végétaux nécessaires à l'identification que des mécanismes à mettre en œuvre pour s'orienter sur le chemin menant à l'espèce correcte.

La richesse et la diversité des espèces végétales est elle aussi faramineuse, on ne compte ainsi pas moins de 300000 espèces distinctes et connues de par le monde, et la flore française, qui est pourtant bien loin d'être la plus fournie, affiche au bas mot 5000 espèces différentes. De tels ordres de grandeur surpassent de beaucoup les capacités de discernement de la machine, et il est évidemment nécessaire d'établir des restrictions, un système unique capable de couvrir l'ensemble du spectre si large en nature et en aspect d'objets paraissant bien illusoire.

Pour autant, même en se restreignant à un ensemble limité d'espèces la reconnaissance visuelle automatique des plantes pose un certain nombre de questions scientifiques et techniques qui méritent d'être traitées. C'est en tout cas bien dans ce cadre que la technologie peut jouer son rôle, et recréer, au travers d'un outil automatique, un lien vers une connaissance désormais hors de portée de tout un chacun.

Motivations

Le monde végétal n'est plus aujourd'hui l'environnement immédiat, source de bienfaits en tous genres, qu'il put être pour les générations passées. La société moderne, avec son urbanisation irrésistible et ses biens de consommation synthétiques, nous a peu à peu coupés de cette part de la nature, qui était autrefois une composante tangible de la vie de tous les jours. Dans cette séparation, nous avons pour la plupart oublié les noms et les usages de tant d'arbres, fleurs et plantes. Un savoir ancestral éparpillé dans la course effrénée du progrès. De nos jours pour autant, avec la prise de conscience grandissante que les ressources naturelles dans toute leur diversité sont un trésor à préserver, le désir de retrouver un certain contact avec la nature semble ne jamais avoir été aussi présent.

L'identification des espèces de plantes est la première et indispensable clé pour comprendre l'environnement végétal. A ce titre, doter le néophyte d'un outil technologique lui permettant de connaître les plantes qui l'entourent serait un progrès considérable. C'est d'autant plus vrai que l'alternative qui s'offre au botaniste en herbe passe par le défrichage des ouvrages de flore particulièrement touffus, et de leurs clés de détermination impénétrables, pour qui ne dispose pas du bagage théorique nécessaire.


Les technologies mobiles aujourd'hui omniprésentes rendent l'objectif de transmettre cette compétence d'identification envisageable, en permettant, à qui en ressent le besoin, d'accéder à des connaissances qui sont à présent un apanage de botaniste. Aider à identifier une espèce végétale, à apprendre son histoire et ses propriétés, est tout autant un moyen de transmettre un savoir égaré que d'offrir aux gens un aperçu de la richesse insondable de la nature.

L'émergence des *smartphones*, ces téléphones mobiles "intelligents", permet de propulser des applications technologiques dans la poche d'un nombre sans cesse croissant d'utilisateurs. Les fonctionnalités dont ils disposent (processeurs intégrés, connexion quasi-permanente, prise de photos et de vidéos, accéléromètres, géolocalisation) en font des supports idéaux pour l'informatique intelligente liée à l'image, offrant des possibilités accrues d'interaction et de traitement embarqué tout en posant des problèmes inhérents à la nature des appareils (faible capacité de calcul, qualité incertaine des images, mémoire limitée, utilisateurs peu consciencieux).

Il s'agit de plus d'un moyen formidable d'appliquer une technologie par ailleurs bien testée et validée dans un cadre contrôlé en laboratoire à un cas d'utilisation réel et nettement moins maîtrisé. La maturité des techniques d'analyse d'images et de reconnaissance automatique est aujourd'hui suffisante pour ce genre d'applications concrètes, et les confronter à un environnement plus complexe et moins rigoureux que celui dans lequel elles ont été développées constitue un bon moyen de tester leur validité et leur robustesse.

Le fait d'essayer une technologie à une si large échelle permet également d'en récupérer les résultats pour des applications en sciences participatives. Il est non seulement intéressant d'observer comment se comportent les méthodes scientifiques utilisées, comment les utilisateurs interagissent avec elles, mais surtout l'objet même de l'application peut fournir une source d'informations passionnante, via le contenu généré par les utilisateurs (*user generated content*, UGC).

Du point de vue de la botanique, l'utilisation d'une application mobile d'identification des espèces végétales constitue une mine d'observations localisées, qui s'apparentent à des relevés de terrain confirmant la présence d'espèces à un endroit donné. C'est un apport important de données brutes dont bénéficieraient les scientifiques du domaine, pour surveiller les populations végétales, mesurer les impacts des changements climatiques sur la répartition des espèces dans le cadre d'une science citoyenne déjà d'actualité (notamment en France via le réseau francophone TelaBotanica¹) mais rendue plus immédiate par la facilité d'accès et de communication qu'offre la technologie mobile.

1.  Tela Botanica : <http://www.tela-botanica.org>, réseau francophone de référence de la communauté botanique regroupant professionnels et amateurs, autour d'événements de publications et de projets participatifs.

La conception d'une application mobile d'identification des espèces végétales se présente donc comme un moyen idéal de recréer un lien entre les sciences botaniques et des citoyens qui s'en trouvent aujourd'hui bien distants. En convertissant une somme de ressources et connaissances botaniques, biogéographiques ou phytosociologiques en un outil technologique intelligent qui les rende accessibles au néophyte, la transmission d'un savoir délaissé peut être envisagée, tout en présentant un intérêt certain pour les acteurs du domaine. Une telle application laisse en tout cas entrevoir l'opportunité d'un rapprochement entre deux mondes que la marche du temps n'a jusque là tendu qu'à éloigner.

Contexte scientifique

Ce n'est pas un hasard si l'identification automatique des espèces de plantes est un problème de plus en plus fréquemment abordé dans le contexte de la recherche d'image par le contenu (*context-based image retrieval*, CBIR) et de la reconnaissance de formes. Les espèces végétales forment des classes bien définies au regard des méthodes de reconnaissance, et les éléments qui permettent de les distinguer sont généralement suffisamment visuels pour qu'on puisse envisager leur extraction et leur caractérisation dans des images numériques.

En témoignent d'ailleurs les différents challenges s'adressant à la communauté scientifique et les collections d'images mises à disposition des chercheurs ces dernières années. L'essentiel de ces initiatives ne concerne malheureusement pas véritablement des images correspondant à une situation réelle. On trouve ainsi de nombreuses images scannées, où vont pouvoir s'appliquer des méthodes qui n'auront aucune chance de passer à l'échelle d'une utilisation non supervisée en environnement naturel.

L'un des principaux verrous scientifiques que pose la reconnaissance d'organes végétaux dans des images naturelles concerne le **traitement de l'image** à bas niveau, et notamment tout ce qui vise à extraire des objets d'intérêt de l'image. Le peu de contraintes et de certitudes sur le contenu des photographies qui devront être traitées en font dans le cadre général un problème extrêmement difficile, en particulier si l'on s'attache à isoler explicitement les objets de leur environnement par une méthode de **segmentation**. Le défi va alors consister à trouver les moyens d'intégrer quelques certitudes de façon à traiter les images avec un présupposé minimal de départ quant à leur contenu et rendre le problème plus abordable.

Le problème de **classification** ou du moins de décision que représente l'identification des espèces est lui-même extrêmement complexe, comme souvent dans le cas d'objets naturels. Sans restriction, le nombre de classes serait évidemment démesuré, et une limitation vraisemblable à l'ordre de la centaine d'espèces constituerait encore une grandeur quelque peu excessive pour les méthodes de classification. Les classes elles-mêmes présentent des particularités qui viennent encore compliquer la tâche de reconnaissance, à commencer par une grande similarité inter-classe possible, parfois supérieure à l'écart qui peut se présenter entre deux individus d'une même espèce pour certains critères.

Il est toutefois possible de simplifier ce problème de reconnaissance par la prise en compte d'éléments extérieurs à l'image pouvant être fournis par les smartphones, à commencer par la position géographique. En effet, en se plaçant à l'interface avec la phytogéographie, il doit être possible d'améliorer la prise de décision en prenant en compte la répartition observée des espèces à l'échelle d'un territoire. Cela soulève malgré tout la question de la façon de traiter cette information d'un genre complètement différent, et renvoie donc à la **fusion de données hétérogènes**. Combiner diverses sources de connaissance de façon à réduire au mieux l'incertitude quant à la décision constitue une piste d'études des plus riches.


Les données biogéographiques (étude de la répartition spatiale du vivant et des facteurs l'influençant) et phytosociologiques (étude des populations d'espèces végétales et de leurs interactions) sont elles-mêmes assez difficiles à appréhender. Une carte précise de répartition de chaque espèce, même à l'échelle du territoire français, n'est pas concevable, et il s'agira donc de déterminer quels facteurs de l'habitat jouent sur la présence d'une espèce pour être capable de donner une décision relative à l'éventualité de rencontrer l'espèce en un point donné. Cette question n'est que peu abordée par les spécialistes du domaine qui se méfient des tentatives de généralisation, mais elle ne constitue pas moins une potentielle source d'informations déterminantes.

Enfin, la mise au point d'un tel système va passer par le développement d'une application mobile, une interface offrant la possibilité d'interagir avec un utilisateur. En plus des questions techniques liées à ce support, et de leur prise en compte dans la conception d'algorithmes adaptés, cette tâche soulève donc des questions d'ergonomie et de visualisation permettant un meilleur confort d'utilisation, ainsi qu'une réflexion sur la teneur pédagogique à lui donner. Les contextes d'utilisation, et les besoins comblés par une telle application sont également une dimension à prendre en compte.

Tous ces aspects entrent dans le champ des objectifs du projet **ReVeS**² (Reconnaissance de Végétaux pour des interfaces Smartphones) dans lequel est effectuée cette thèse, mené conjointement par trois partenaires académiques :

- ▶ Le LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) à Lyon (UMR 5205), dont l'expertise se situe au niveau du traitement d'image et de l'extraction de descripteurs
- ▶ Le LISTIC (Laboratoire d'Informatique, Systèmes, Traitement d'Information et de la Connaissance) à Annecy (EA 3703) spécialisée en traitement et en fusion d'information
- ▶ EVS (Environnement, Ville, Société) à Saint-Etienne (UMR 5600) laboratoire de géographie apportant notamment une compétence en géomatique et systèmes d'information géographiques

Les travaux menés dans le cadre de cette thèse vont donc s'attacher à répondre aux problématiques évoquées précédemment, avec pour but final de définir un système cohérent, adapté aux **supports**

2.  Projet financé par l'Agence Nationale de la Recherche **ANR** dans le cadre de l'appel à projets CONTINT 2010 (ANR-10-CORD-005), d'une durée de 3 ans, ayant débuté le 18/10/2010

mobiles, permettant par la prise en compte du contenu d'**images naturelles** et d'**informations géographiques** extérieures de fournir une aide à l'**identification des espèces** végétales.

Cadre de travail

Le scénario d'utilisation typiquement envisagé pour l'application est celui d'une personne se promenant dans la nature qui, tombant sur une plante inconnue, en prend une photographie à l'aide de son smartphone et, après quelques interactions sommaires avec l'application, obtient des informations sur les espèces les plus probables auxquelles peut appartenir la plante. Ce scénario impose quelques impératifs, notamment le fait que la reconnaissance doive s'effectuer à partir d'une seule photographie, et que, pour garantir le fonctionnement loin du milieu urbain connecté, les traitements puissent tous s'effectuer sur le téléphone lui-même.

Le public visé n'est ainsi pas un public de spécialistes, possédant déjà une certaine expertise botanique, mais bien un utilisateur néophyte, présentant un intérêt certain pour la nature, quoique dépourvu du bagage théorique souvent nécessaire pour identifier les espèces. L'un des objectifs est de transmettre une compétence d'identification, la visée étant non seulement de reconnaître l'espèce présente dans une image, mais en plus d'expliquer le cheminement qui permet d'arriver à cette décision. A l'opposé d'une boîte noire prenant une image en entrée et livrant simplement une réponse en termes d'espèces, notre système se devra d'être une analyse transparente des critères intuitifs permettant l'identification botanique.

Les botanistes se basent traditionnellement sur les formes, structures et compositions des fleurs, des fruits et des feuilles pour reconnaître les espèces. Dans le cadre d'une application pour non-spécialistes, l'utilisation prédominante des feuilles est la voie la plus prometteuse. Ce sont en effet les organes dont la définition correspond le mieux à la représentation que s'en fait le sens commun. Étant des objets généralement plans, elles ont de plus l'avantage d'être faciles à photographier, avec l'assurance qu'une seule image suffise à capturer l'ensemble des informations nécessaires à l'identification. Pour ne rien gâcher, on peut généralement en trouver quelle que soit la saison, et les paramètres permettant de les reconnaître font souvent intervenir une géométrie commode pour l'analyse.

Le domaine des espèces de la flore qui va pouvoir être considéré est donc celui pour lequel les feuilles sont les plus abondantes, à savoir les arbres, et plus précisément les essences d'arbres que l'on peut trouver à l'état spontané en France métropolitaine. Si celles-ci n'ont que peu d'intérêt pour les spécialistes, elles n'en demeurent pas moins suffisamment disparates et méconnues du grand public pour constituer un objet parfaitement pertinent. La délimitation du concept d'arbre n'existe pas véritablement en botanique, et on se bornera donc aux quelques 150 espèces natives d'arbres et d'arbustes pouvant atteindre plus d'1m20 que compte la flore française, et dont les principales ont été retenues et figurent dans l'Annexe A.

Les contraintes posées par le traitement sur les supports mobiles sont assez fortes. Une première limite pourrait venir de la qualité des images prises par ces appareils, bien que leur finesse et leur

robustesse soient en constante augmentation. Plus contraignant est l'écart qui existe entre les processeurs de téléphones et ceux d'un ordinateur (de l'ordre du facteur 10 en temps de traitement) ainsi que dans la capacité en mémoire. Les méthodes développées devront donc porter une attention particulière à ces aspects, sans pour autant que cela soit préjudiciable à la performance sur des images potentiellement complexes.

Pour faciliter le traitement de ces images, et pour disposer de suffisamment de certitudes avant d'appliquer des méthodes sur des hypothèses hasardeuses, l'interaction possible avec l'utilisateur permet d'envisager de fixer un cadre strict pour l'acquisition d'images. La prise de photo va ainsi pouvoir être guidée de façon notamment à capturer la feuille dans son intégralité, à forcer son alignement dans l'image, ou à s'assurer qu'elle soit disposée dans un sens donné, sa base vers le bas par exemple. Ce sont autant d'informations que nous supposons acquises dans la suite du traitement, et qui permettront le développement de méthodes moins générales mais plus spécifiquement dédiées à l'objet de notre étude, la feuille.

Enfin, la volonté de transmettre une connaissance relative à l'identification et à la description des espèces passe par la mise en place d'un processus d'identification explicatif. En particulier, les informations extraites de l'image devraient pouvoir être restituées de manière visuelle ou interprétable à l'utilisateur de telle sorte que le mécanisme de décision soit explicite. La donnée des critères menant à l'identification de la bonne espèce est un bon moyen de s'assurer que quelque chose peut être transmis par le système. Au contraire, en cas d'erreur, le fait d'avoir sous la main des éléments interprétables permet de localiser les défauts, et éventuellement dans le cadre d'une interaction complète de les corriger. Dans tous les cas, la portée pédagogique d'un système, sa capacité à délivrer la connaissance sur laquelle il est bâti, ne peut que bénéficier d'une transparence des méthodes d'identification.

L'objectif sera donc de développer un système dédié à la description et à l'interprétation des feuilles d'arbres, dans des images supervisées *a minima* prises en environnement naturel par un appareil photo de smartphone. Le traitement de ces images devrait permettre d'obtenir suffisamment d'informations pour déterminer un ensemble restreint d'espèces auxquelles semble appartenir la feuille, avec l'aide d'informations biogéographiques pour affiner le choix effectué parmi la centaine d'espèces natives. Le but avoué d'un tel processus est de fournir à l'utilisateur une représentation intelligible du processus de reconnaissance, de façon à lui ouvrir les portes d'un domaine par ailleurs peu accessible, et de lui livrer les moyens de se plonger plus avant dans l'univers fascinant de la botanique.

Chapitre 1

L'identification en botanique

De la place des feuilles dans la science des plantes



Betula pendula Roth

Résumé du Chapitre

Ce chapitre présente la manière dont les botanistes abordent le problème de l'identification des espèces végétales, en commençant par donner quelques notions sur ce qu'est une espèce, et ce qui permet de la caractériser. Il présente la façon dont sont classées les espèces dans le règne végétal, et les outils qui sont utilisés communément pour déterminer l'espèce d'un individu observé. En se focalisant uniquement sur les feuilles, il est alors possible de discriminer un nombre conséquent d'espèces d'arbres en étudiant les caractéristiques utilisées pour décrire les formes foliaires. La fin du chapitre expose l'idée centrale selon laquelle la botanique peut être utilisée comme guide dans un processus de reconnaissance automatique, dans l'intérêt à la fois de la machine et de l'utilisateur.

1. L'identification en botanique



IDENTIFICATION des espèces d'arbres à partir de leurs feuilles est une méthode bien répandue et sans aucun doute la plus accessible lorsque l'on commence à se pencher un peu sur les arcanes de la botanique. En témoignent les différents guides nature que chacun a pu croiser (Figure 1.1) et qui guident le botaniste en herbe entre les diverses essences, par le biais d'illustrations colorées et représentatives des feuilles. Des clés de détermination plus ou moins sommaires permettent de s'y retrouver entre les quelques grands groupes choisis pour structurer le guide. Les autres organes, s'ils sont évidemment décrits, passent le plus souvent au second plan, pour ce qui est de l'identification tout du moins.

On le comprend aisément, la feuille de l'arbre est un objet plus facile, plus intuitif et compréhensible que la fleur ou le fruit, qui requièrent un minimum de connaissance de leurs composants et de leur structure s'ils doivent être analysés. Elle a l'avantage indéniable de figurer de façon assez précise dans l'inconscient collectif, chacun ayant généralement une bonne idée de ce qu'est une feuille. Le fait que l'essentiel de ce qui permet leur identification soit des caractéristiques de forme, visuellement immédiates pour un observateur, renforce aussi cet aspect intuitif. Enfin, elles sont évidemment présentes sur une période de l'année plus longue que les autres organes (à l'exception de l'écorce) du printemps à l'automne, et peuvent encore être trouvées au sol en hiver (si l'enneigement ne le compromet pas) quand les fruits et surtout les fleurs présentent une durée d'observation bien plus limitée dans le temps.

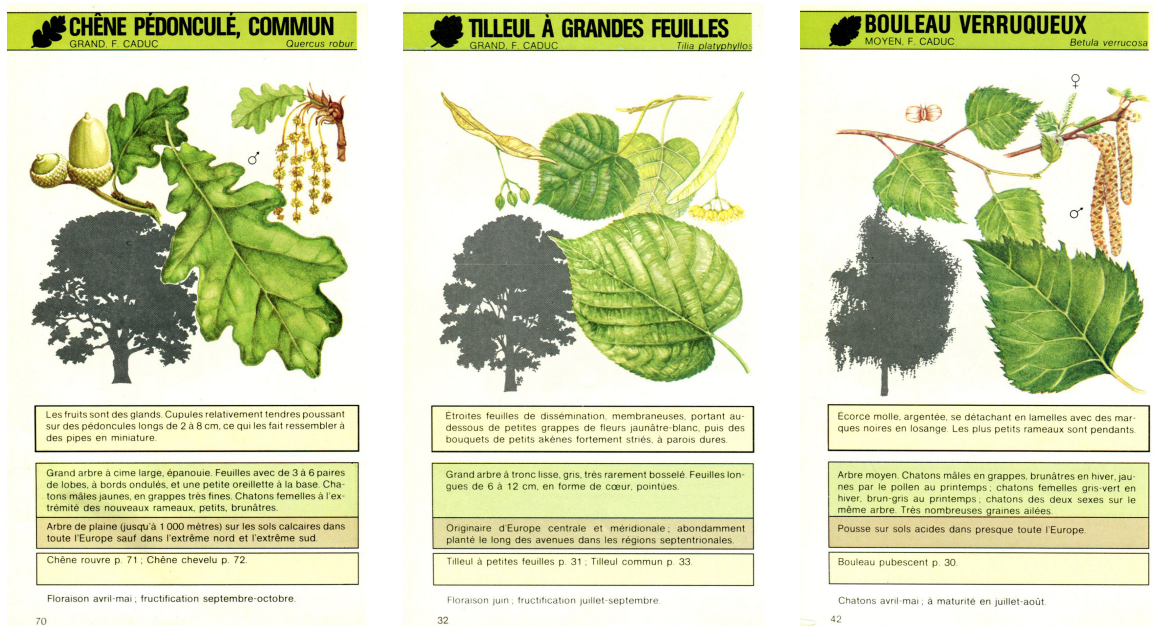


FIGURE 1.1 – Exemples d'illustrations tirées d'un guide nature [FF88]

Malgré cette évidence, les feuilles ne sont pas le critère primordial de la classification des espèces d'arbres, loin s'en faut. Pour être exact, le terme d'arbre lui-même ne correspond pas à une classe d'espèces végétales bien délimitée, qui serait indépendante des autres. La classification botanique

est basée sur les liens des espèces entre elles, et à ce titre, ce sont des propriétés plus fondamentales de reproduction et de développement de la plante qui sont à prendre en compte, bien plus que la forme des feuilles qui n'est au final qu'un caractère phénotypique de l'espèce, moins relié à sa nature intrinsèque. Il est donc important de comprendre comment sont définies les espèces avant de se proposer de les reconnaître.

Toutefois, la feuille constitue un objet d'étude de la botanique depuis ses origines, et il existe donc toute un ensemble d'outils permettant la compréhension, la description et la caractérisation des feuilles. Malgré l'absence de lien structurel entre la feuille et l'agencement des espèces entre elles, il n'en demeure pas moins que les feuilles d'une même espèce possèdent des propriétés communes, qui peuvent être décrites, et utilisées pour la reconnaître. Cette description utilise un vocabulaire assez complet, parfois un peu abscons, qui mérite d'être décortiqué pour comprendre les outils à disposition pour analyser une feuille, et puiser dans ce savoir acquis par des générations de botanistes, en vue d'un processus de reconnaissance automatique.

1.1 Taxinomie et clés de détermination

La paternité de la classification des espèces végétales est sans conteste à mettre au crédit de Carl von LINNÉ dont les travaux d'unification et de nomenclature [Lin53] ont jeté les bases de la botanique moderne. Si les notions d'espèce, de genre et de famille avaient émergé durant le XVII^{ème} siècle, c'est à LINNÉ que revient le mérite d'avoir synthétisé ces approches, unifié les dénominations, et établi une classification basée non plus sur les usages et l'apparence, mais utilisant les caractéristiques intrinsèques des plantes pour classer l'ensemble du règne végétal.

1.1.1 Nomenclature des espèces végétales

La nomenclature consiste tout simplement à donner un nom unique à chaque chose, de façon à ce qu'il soit connu et reconnu de tous. Cela suppose toutefois de se mettre d'accord sur ces choses que l'on souhaite nommer, à savoir les espèces végétales.

“ Un espèce est un ensemble d'individus semblables par leur forme, vivant au contact les uns des autres, s'accouplant exclusivement les uns aux autres et demeurant indéfiniment féconds entre eux. ”

Encyclopédie LAROUSSE

La notion d'espèce

Dans le monde du vivant, l'espèce est la brique de base de la taxinomie, mais sa définition n'est pas nécessairement chose évidente. On considère ainsi le plus souvent le critère d'interfécondité comme définition, à savoir qu'une espèce est essentiellement un groupe d'individus pouvant se reproduire

entre eux. En rajoutant le critère complémentaire, selon lequel deux individus d'espèces distinctes ne peuvent avoir de descendance féconde, on parvient à isoler des groupes indépendants.

Le cas du règne végétal est particulièrement complexe, puisque les hybridations y sont bien plus fréquentes que chez les animaux, rendant ce critère d'isolement reproductif un peu plus flou. Il n'en demeure pas moins l'idée fondamentale que la proximité des individus d'un point de vue taxinomique est directement liée à leur reproduction, et que c'est bien les organes reproducteurs qu'il faut étudier pour établir une classification.

Nomenclature binomiale

On doit à LINNÉ entre autres choses la formalisation de la dénomination des espèces, par le biais d'un binôme de noms dits latins, l'un représentant le **genre** (le nom générique) et l'autre l'**espèce** (l'épithète spécifique). Cette nomenclature sous forme de nom scientifique, ou **nom binomial**, permet de désigner de façon univoque toute espèce du monde vivant, sans avoir recours aux noms communs (dits vernaculaires) dans différentes langues, beaucoup moins précis.

Ces binômes genre-espèce sont souvent complétés par le nom (ou simplement l'initiale) du naturaliste ayant nommé l'espèce, voire de l'année de sa validation. C'est ainsi que l'on trouve apposée à de nombreux noms latins d'espèces végétales la lettre L. représentant LINNÉ.

Nom vernaculaire	Nom scientifique ou nom binomial
Erable champêtre	<i>Acer campestre</i> L., 1753
Erable de Montpellier	<i>Acer monspessulanum</i> L., 1753
Erable à feuilles d'obier	<i>Acer opalus</i> Mill., 1768
Erable plane	<i>Acer platanoides</i> L., 1753
Erable sycomore	<i>Acer pseudoplatanus</i> L., 1753

TABLE 1.1 – Quelques espèces du genre *Acer* et leurs dénominations scientifiques

La notion de genre vient d'un point de vue taxinomique juste au dessus de l'espèce, et constitue donc le premier regroupement d'espèces proches. C'est évidemment un concept qui existait avant sa formalisation au XVIII^{ème} siècle, une reconnaissance de la proximité d'espèces par ailleurs discernables entre elles. Par exemple, l'érable champêtre, l'érable sycomore et l'érable plane sont des espèces différentes, mais font tous partie du genre des érables. La Table 1.1 montre la transposition de cet état sous forme de noms binomiaux.

La définition du genre est tout de même moins facile à donner, si ce n'est qu'il s'agit d'espèces voisines que l'on peut regrouper. Etant donnée la définition première de l'espèce, c'est évidemment en observant les organes reproducteurs que sont les fruits et les fleurs que ce rapprochement fondamental va être établi. C'est d'ailleurs au sein d'un même genre que vont généralement se produire les hybridations naturelles entre espèces. On peut par exemple citer le genre *Quercus*, L. qui rassemble les chênes, espèces qui partagent toutes la propriété de porter des glands, un type de fruit qu'elles sont les seules à produire (Figure 1.2).

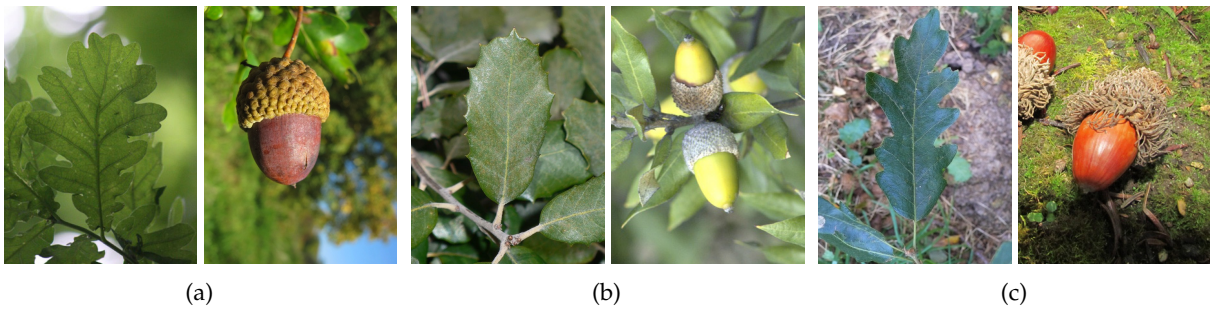


FIGURE 1.2 – Les chênes, diversité des feuilles, identité des fruits : (a) *Quercus robur* L. (b) *Quercus ilex* L. (c) *Quercus cerris* L.

La classification Linnéenne et ses héritières

Le modèle de classification initiée par LINNÉ et que l'on dénomme encore aujourd'hui classification classique (ou systématique classique) repose sur une hiérarchie de classes imbriquées, que l'on nomme toutes **taxons**. Il existe des taxons de différents rangs, dont les espèces et les genres constituent les premiers échelons, et se trouvent englobés dans des entités bien plus larges. La Figure 1.3 illustre l'agencement de ces rangs taxinomiques.

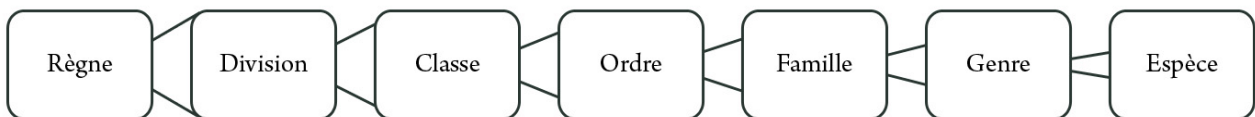


FIGURE 1.3 – Organisation des différents rangs taxinomiques dans la classification Linnéenne

Au dessus du genre se trouve la **famille**, désignée par des termes latins terminant en *-aceae* et regroupant les genres les plus proches. C'est à cet étage que commencent à se faire des rapprochements moins évidents. Si les espèces d'un même genre sont généralement suffisamment semblables du point de vue de l'apparence pour que leur rapprochement soit entériné par le langage courant, la parenté des genres d'une même famille est parfois moins visible. Par exemple, les *Oleaceae* famille de l'olivier (*Olea europaea* L.) comprennent également, entre autres, les frênes (*Fraxinus* L.) les lilas (*Syringa* L.) ou encore les forsythias (*Forsythia* Vahl), sans pour autant que leur proximité ne saute aux yeux.

Pire encore, la (très imposante) famille des *Rosaceae* qui ne contient pas moins d'une centaine de genres, pour plus de 3000 espèces, regroupe de façon assez surprenante nombre d'arbres fruitiers (*Prunus* L.) pommiers (*Malus* Mill.) avec des plantes moins arbustives telles les fraisiers (*Fragaria* L.), les ronces (*Rubus* L.), les potentilles (*Potentilla* L.) et bien évidemment les roses (*Rosa* L.).

Une fois encore, c'est par l'observation des organes reproducteurs des plantes que ces familles ont pu être constituées. Les fruits et les fleurs des espèces appartenant à une même famille ont en commun des propriétés morphologiques qui permettent leur rapprochement. Par exemple les *Betulaceae* (bouleaux, aulnes, noisetiers, charmes) présentent tous des inflorescences caractéristiques,

des chatons unisexués comme l'illustre la Figure 1.4.

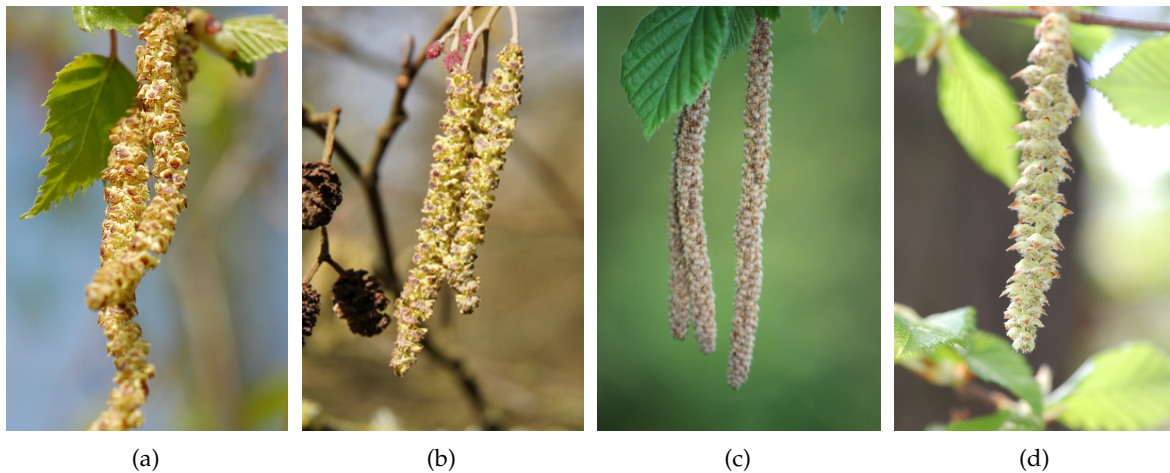


FIGURE 1.4 – Chatons d'espèces appartenant à différents genres de la famille des *Betulaceae* : (a) *Betula* (b) *Alnus* (c) *Corylus* (d) *Carpinus*

L'échelon supérieur est celui de l'**ordre**, marqué par la terminaison latine en *-ales* qui devient un regroupement très vaste d'espèces, puisque rassemblant plusieurs familles. Historiquement, cette notion apparaît plus tard, et les botanistes de la fin du XVIII^{ème} et du XIX^{ème} siècle vont s'évertuer à proposer des définitions de ces ordres, et des familles qu'ils contiennent, en suivant les bases taxinomiques édictées par LINNÉ. On peut citer parmi ces fondateurs les français Antoine-Laurent de JUSSIEU [Jus89] et Augustin de CANDOLLE [Can24] dont les travaux ont inspiré les classifications botaniques, jusqu'à la classification de CRONQUIST au XX^{ème} siècle.

Les rangs taxinomiques supérieurs correspondent à de grandes divisions du vivant, et sont de ce fait plus faciles à appréhender. Les plantes font toutes parties du règne végétal *Plantae* qui se divise en plusieurs grands groupes, notamment la division des **Angiospermes** (ou plantes à fleurs) qui sont celles qui nous intéressent particulièrement dans notre cadre de travail, par opposition aux **Gymnospermes**, division dans laquelle se trouvent les conifères. Parmi les Angiospermes, on distingue classiquement les deux grandes classes des **Monocotylédones** et des **Dicotylédones** (les cotylédons étant les feuilles primordiales qui naissent de la graine, leur nombre -un ou deux- permettant de faire la distinction) qui vont être elles-mêmes subdivisées en ordres.

1.1.2 Classification phylogénétique et conséquences

L'émergence des théories Darwiniennes de l'évolution à la fin du XIX^{ème} siècle vient quelque peu bouleverser cette belle organisation fortement hiérarchisée. L'idée que les espèces peuvent évoluer pousse à reconsidérer la façon de les agencer entre elles, et fait naître la notion de classification évolutive. Et ce n'est plus tant la similitude des critères naturels, que la filiation des espèces qui va alors présider à la construction de toute leur classification.

Proximité temporelle et clades

La classification phylogénétique, initialement conçue par le biologiste allemand Willi HENNIG dans la première moitié du XX^{ème} siècle, considère comme relation de base le degré de parenté entre les espèces. C'est la notion de proximité temporelle entre plusieurs espèces, au sens de l'évolution, qui permet d'établir une relation entre trois espèces E_1 , E_2 et E_3 , pouvant se traduire par E_1 est *plus proche* de E_2 que de E_3 . Cette relation a une traduction immédiate sous forme graphique, un simple arbre phylogénétique à trois feuilles comme celui de la Figure 1.5.

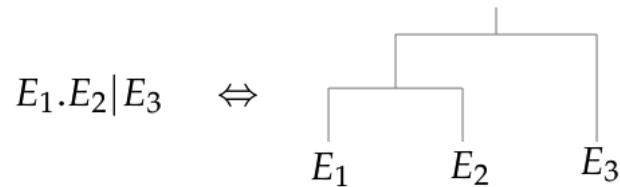


FIGURE 1.5 – La relation de base de la phylogénétique

De plus, l'interprétation en termes d'évolution de cette relation est limpide, elle signifie qu'il existe un ancêtre commun des espèces E_1 et E_2 qui ne fait pas partie des espèces ancestrales de E_3 . Autrement dit, E_1 et E_2 sont plus proches parents dans l'histoire de l'évolution des espèces que ne le sont E_1 et E_3 , ou E_3 et E_2 .

La connaissance de ces relations entre les taxons (obtenue plus facilement depuis le développement de la génétique) permet ainsi de construire un arbre phylogénétique des espèces végétales, qui représente en fait l'évolution des espèces à partir des leurs ancêtres communs, et les embranchements à partir desquels les espèces ont divergé. La structure d'un tel arbre est extrêmement complexe, mais elle traduit exactement ce qui va servir de base à la classification phylogénétique, à savoir la parenté des espèces. La construction effective de cette classification correspond à la mise en évidence de groupes monophylétiques, que l'on nomme aussi **clades**. Un clade est un groupe qui comprend une espèce ancestrale et l'ensemble de ses descendants. Cela correspond à une branche complète d'un arbre phylogénétique, dont toutes les feuilles peuvent remonter à un même nœud, faisant lui-même partie du clade.

Fort heureusement, les espèces, les genres, les familles et les ordres, au sens de la classification classique, constituent généralement des clades. Les rapprochements qui avaient été faits par l'observation de critères extérieurs traduisaient ainsi une proximité que l'on retrouve dans l'histoire de l'évolution des espèces. La génétique a certes permis de corriger certaines affectations, remodelant les contours de plusieurs ordres ou familles, mais les dénominations sont conservées. En revanche, les rangs supérieurs voient leur réalité remise en cause, et surtout la rigidité de la taxinomie vole en éclats. Le nombre de rangs taxinomiques ne peut plus être fixe, des clades pouvant constituer des branches indépendantes, ou bien être imbriqués sous d'autres clades. La Figure 1.6 donne une idée de la structure de l'arbre phylogénétique et de l'agencement des clades.

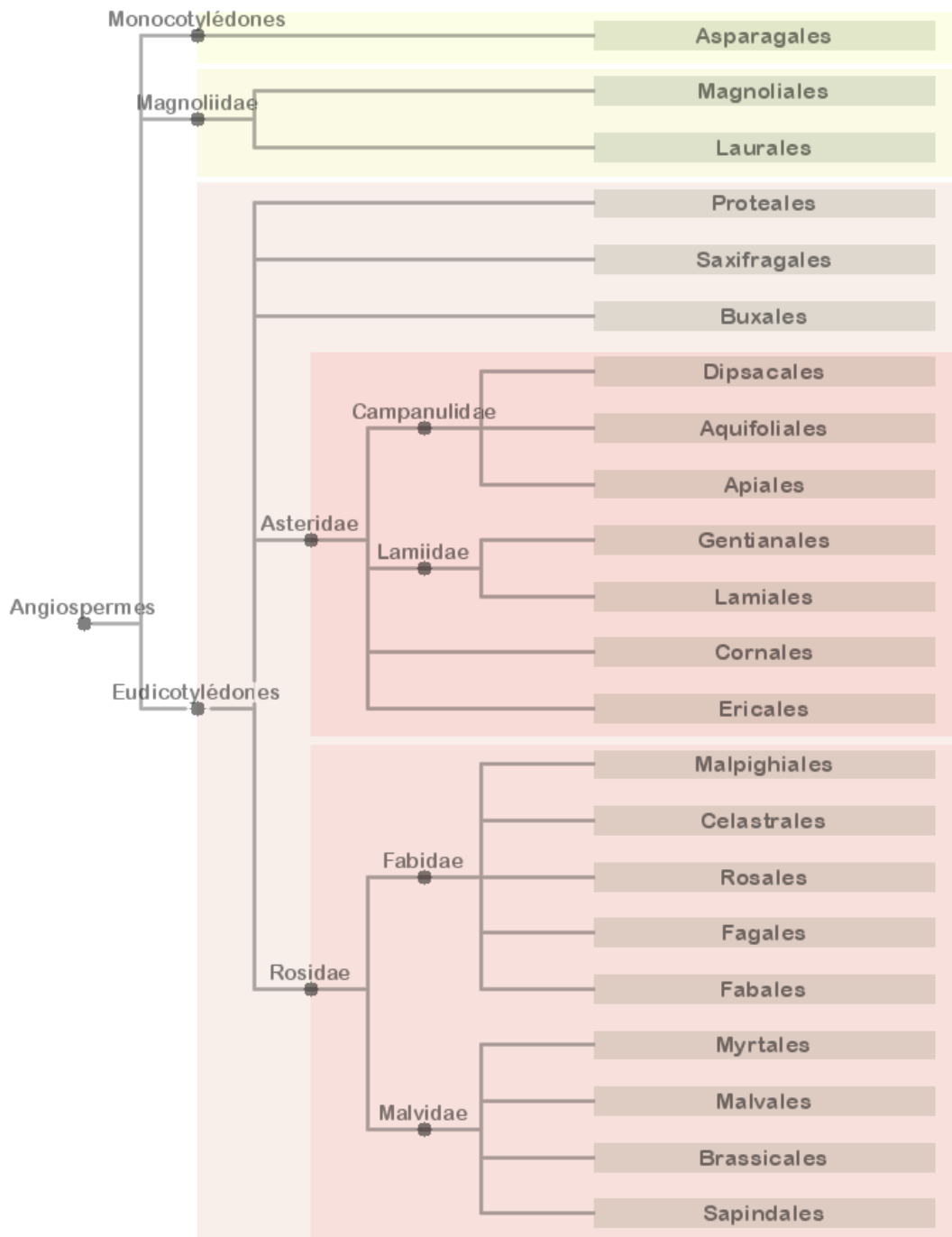


FIGURE 1.6 – Arbre phylogénétique des ordres des espèces considérées selon la classification APG III [Gro09]

Les espèces d'arbres dans la classification

La notion d'**arbre** ne se retrouve pas à proprement parler dans les critères qui président à la classification des espèces. Il est même difficile d'en donner une définition univoque, situant précisément la limite entre les espèces. Les éléments généralement retenus pour caractériser les espèces d'arbres sont avant tout la pérennité de la plante, sa taille (qui doit dépasser les 5, voire 7 mètres au dessus du sol) et le fait qu'elle soit ligneuse (tige composée de bois véritable) et possède un unique tronc autoportant ainsi que des branches à partir d'une certaine hauteur.

Ces caractéristiques ne sont hélas pas l'apanage d'un seul embranchement de la phylogénie des espèces, mais une solution partagée par de nombreuses espèces distinctes, une stratégie commune pour accéder plus facilement à la lumière. Cette convergence évolutive a généré des formes qu'il convient de rassembler sous le même terme, mais c'est un rapprochement qui ne se retrouve pas dans la structure de la classification phylogénétique des espèces.

La classification actuellement reconnue par la majorité des botanistes est la Classification APG III, établie en 2009 par l'Angiosperms Phylogeny Group [Gro09]. Il s'agit d'une classification phylogénétique regroupant l'ensemble des espèces d'Angiospermes connues à ce jour, mais elle ne concerne pas les Gymnospermes, et donc dans notre cas les espèces d'arbres conifères. Les espèces y sont donc regroupées en clades, les clades de niveaux inférieurs correspondant aux genres, aux familles et aux ordres. La figure 1.6 illustre la place dans la classification APG-III des principaux ordres comprenant des espèces d'arbres et arbustes natives d'Europe.

Ces ordres contiennent bien évidemment des espèces autres que celles qui constituent notre objet d'étude, mais la structure de l'arbre permet de mettre en évidence un fait capital : la plus petite subdivision des Angiospermes dans la classification botanique qui contienne toutes les espèces d'arbres n'est autre que la totalité des Angiospermes. Le concept d'arbre ne peut donc même pas être rattaché à un sous-ensemble de l'arbre phylogénétique des espèces, encore moins à un clade.

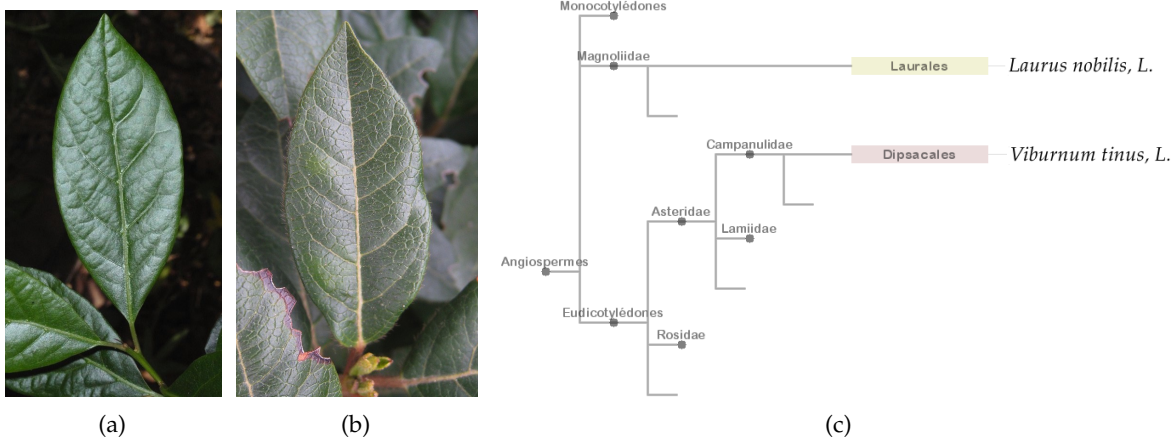


FIGURE 1.7 – Feuilles de laurier sauce (a), de laurier-tin (b) et leurs places respectives dans la classification APG III (c)

Cet état de fait rend assez compliquée la délimitation du problème d'identification des espèces d'arbres, puisqu'il est pour ainsi dire impossible de se borner à une partie seulement de la classification des plantes à fleurs. Il faut donc la considérer dans son ensemble, et définir au cas par cas les espèces, ou à la rigueur les genres, qui vont constituer l'ensemble à étudier.

L'autre inconvénient majeur est l'impossibilité de bâtir à partir de cette classification une quelconque hiérarchie qui permettrait l'identification à partir des critères de forme des feuilles. D'une part, parce que ce sont des critères bien peu représentatifs de la proximité temporelle des espèces, et qu'ils n'ont au final que peu de lien avec la façon dont est construite la classification. D'autre part, certaines similitudes de formes n'ont aucun sens lorsqu'on les transpose dans l'arbre phylogénétique. Par exemple, le laurier-tin (*Viburnum tinus* L.) et le laurier sauce (*Laurus nobilis* L.) ont des feuilles aux formes semblables, mais comme le montre la Figure 1.7, il faut remonter à la racine de l'arbre phylogénétique pour les relier. Autrement dit, leur dernier ancêtre commun serait aussi l'ancêtre commun de tous les Angiospermes.

Par conséquent, l'idée d'un processus de reconnaissance qui suivrait un chemin dans l'arbre phylogénétique des espèces, à partir de l'ensemble des Angiospermes jusqu'à une espèce déterminée, en se basant uniquement sur la feuille est proprement illusoire. La classification des espèces végétales est un passage obligé pour comprendre l'objet que l'on est amené à étudier, mais il n'offre quasiment aucune prise pour l'identification des espèces à partir de leurs feuilles, qu'elle soit humaine ou automatique.

1.1.3 Les clés de détermination en botanique

Les botanistes utilisent cependant des outils qui se rapprochent de ce cheminement dans un arbre, en respectant souvent la hiérarchie de la classification. Ce sont des **clés de détermination**, indispensables outils du naturaliste de terrain, qui offrent une méthode analytique pour assigner un individu à une espèce, par une suite d'observations de critères donnés.

Premiers critères de décision

Le problème de l'identification diffère du problème de classification puisqu'il n'a pas pour but de regrouper les espèces, mais d'associer un individu à un groupe issu d'une classification déjà existante. Nombre de clés de détermination faisant actuellement référence se basent sur des flores datant des années 1900, comme, pour la France, la flore de Bonnier [BL09], celle de Coste [Cos06] ou celle de Rouy [RBCF13]. Ces clés sont donc bâties sur des systématiques différentes de la cladistique aujourd'hui en vigueur, mais elles n'en demeurent pas moins fonctionnelles. En effet le but étant d'arriver jusqu'à l'espèce, les méthodes de détermination restent valides, puisque la définition des espèces (et, dans une large mesure, des genres et des familles) n'a pas varié au cours du dernier centenaire !

Une clé de détermination consiste en un ensemble de choix à réaliser par l'observation de critères qui, par renvois successifs, vont permettre de naviguer jusqu'à l'espèce de l'individu considéré.

TABLEAU GÉNÉRAL.

<p>⊕ Plante ayant des fleurs : en y trouvant des étamines, un pistil ou les deux à la fois.</p> <p>* Étamines et pistils sur la même plante, quelquefois dans des fleurs différentes.</p> <p>⊙ Fleurs non réunies en capitule, entourées d'une corolle ou d'une calice entouré d'une corolle de bractées.</p> <p>△ Fleurs à une seule enveloppe ou à deux enveloppes de consistance semblables, ou sans enveloppe florale.</p> <p>△ Fleurs à deux enveloppes (calice et corolle) de consistance différente.</p> <p>• Corolle non tubulaire.</p> <p>• Corolle papilionacée (fig. P, A, B, C) : c'est-à-dire irrégulière avec un lobe supérieur et deux lobes inférieurs tombés et libres).</p>		<p>⊕ Pétales libres entre eux, jusqu'à leur base.....</p> <p>⊕ Pétales soudés entre eux, au moins à la base.....</p> <p>— Feuilles à nervures non ramifiées (regarder par transparence) et parties semblables de la fleur disposées par 6 ou 3, ou moins de 3.....</p> <p>— Plante n'ayant pas à la fois ces caractères ; en général, feuilles à nervures plus ou moins ramifiées.....</p> <p>— Feuilles réunies en capitule, c'est-à-dire serrées les unes contre les autres, sans pédoncules et placées sur l'extrémité d'un rameau ou d'une tige, entourées d'une corolle de bractées (mûchero). [Exemples : dans le genre du Bleuet, de la Marguerite, de Chardon sont, en réalité, des capitules de fleurs.].....</p> <p>— Toutes les fleurs sans pistil, ou toutes les fleurs sans étamines.....</p> <p>— Plantes sans fleurs, n'ayant jamais ni étamines, ni pistil.....</p>	<p>Section A : Plantes à pétales séparés → p. X.</p> <p>Section B : Plantes à pétales soudés entre eux → p. XI.</p> <p>Section C : Plantes à corolle papilionacée → p. VIII.</p> <p>Section D : Plantes à une seule enveloppe florale → p. XVII.</p> <p>Section E : Plantes monocotylédones → p. XXII.</p> <p>Section F : Plantes à fleurs en capitule → p. XXV.</p> <p>Section G : Plantes à fleurs en capitule → p. XXV.</p> <p>Section H : Plantes gymnospermes → p. XXVI.</p> <p>Section I : Plantes cryptogames → p. XXVIII.</p>
---	--	---	--

FIGURE 1.8 – Premiers critères de déterminations issus de la flore de Bonnier [BL09]

Lorsque l'on observe les premières questions qui sont posées à l'observateur, on remarque tout d'abord qu'elles sont loin d'être à la portée de tout néophyte, et que de plus, elles concernent essentiellement l'organe reproducteur de tout Angiosperme : la fleur.

C'est ainsi qu'en suivant la clé générale tirée de la flore de Bonnier (Figure 1.8) l'observateur va par exemple devoir, pour progresser dans les sections, déterminer si la fleur présente une corolle papillonacée ou non, et le cas échéant si ses pétales sont soudés entre eux. Les choix suivants sont du même acabit, il s'agira de compter les étamines, de scruter les sépales, et de s'enquérir du calice. Autant de notions qui demandent un minimum de bagage botanique, et qui se trouvent bien loin des représentations que peut avoir tout un chacun concernant les fleurs.

C'est ce manque d'accessibilité qui crée une véritable rupture entre le novice qui s'intéresse aux plantes sans maîtriser les concepts parfois abscons de la botanique, et l'expert capable de s'y retrouver dans ce dédale d'ordres et de familles, grâce à ces connaissances de base. La barrière à franchir pour qui débute en botanique est simplement trop haute pour y parvenir seul. La complétude des clés de détermination fait qu'elles ne peuvent se passer de ces étapes hors de la portée du non-expert, si elles veulent couvrir l'ensemble de la classification des Angiospermes. C'est leur vocation à l'exhaustivité qui exclut de fait les non-initiés.

La feuille en dernier recours

Dans la pratique, un botaniste sautera bien souvent ces premiers critères pour s'intéresser directement à la clé spécifique à l'ordre ou à la famille qu'il aura su reconnaître au premier coup d'œil. Mais une fois encore, c'est uniquement sur la foi de son expérience, un savoir acquis, qu'il pourra le faire. Les étapes suivantes sont en revanche un peu plus compréhensibles, même au néophyte. Pour illustration, les critères dans la flore de Coste (Figure 1.9) permettant d'arriver à la famille des Cupulifères, aujourd'hui *Fagaceae* (la famille des chênes, hêtres et châtaigniers) font mention de fleurs en chatons, d'arbres non épineux, de feuilles simples et de fruits ou glands secs. Des critères qui, sans pour autant être évidents, sont plus accessibles que la composition du pistil.

- ☆ Plantes ligneuses.
 - > Fleurs toutes ou les mâles seules disposées en chatons ou épis denses.
 - ☒ Fleurs mâles seules disposées en chatons ou épis.
 - Arbrisseau épineux, couvert de petites écailles argentées; 4 étamines; fruit drupacé, d'un jaune orangé. **ÉLÉAGNÉES 96**
 - Arbres ou arbrisseaux non épineux, ni écailleux; 5-20 étamines; fruit jamais d'un jaune orangé.
 - Feuilles imparipennées, aromatiques, sans stipules; fruit ou noix à péri-carpe charnu et lisse, sans involucre distinct. **JUGLANDÉES 105**
 - Feuilles simples, non aromatiques, à stipules caduques; fruit ou gland sec, plus ou moins enveloppé par un involucre coriace ou foliacé.
 CUPULIFÈRES 106

FIGURE 1.9 – Critères de déterminations permettant d'arriver aux familles, issus de la flore de Coste [Cos06]

A ce stade, les feuilles font encore une fois office de caractéristiques secondaires, quand elles sont seulement mentionnées. Certaines familles, plutôt réduites, peuvent présenter certaines caractéristiques foliaires suffisamment discriminantes ; la famille des *Ulmaceae* qui comprend les ormes, est ainsi

caractérisée entre autres par ses feuilles dentées. Ce n'est en fait qu'au sein d'un même genre que l'on commence véritablement à utiliser les feuilles pour discriminer les espèces. La Figure 1.10 illustre le genre d'informations données sur la forme des feuilles qui permettent la reconnaissance des espèces de peupliers, mais seulement une fois parvenu au niveau du genre *Populus*.

Feuilles lobées ou sinuées-dentées; jeunes pousses, bourgeons et écailles des chatons poilus; 8 étamines.	
Feuilles adultes suborbiculaires, irrégulièrement sinuées-dentées, vertes et glabres, rarement pubescentes-grisâtres en dessous; pétioles comprimés; chatons à écailles incisées-digitées; stigmates purpurins	P. tremula 3310
Feuilles lobées ou anguleuses, vertes en dessus, très blanches-tomenteuses en dessous, à pétiole arrondi; chatons à écailles crénelées-dentées; stigmates jaunes.	P. alba 3314
Feuilles finement et régulièrement dentées; jeunes pousses, bourgeons et écailles des chatons glabres; 12-20 étamines.	
Feuilles toujours très glabres, ovales-rhomboidales; jeunes pousses arrondies; capsule ovoïde, marquée de 4 sillons, s'ouvrant en 2 valves	P. nigra 3312
Feuilles finement pubescentes aux bords dans leur jeunesse, largement ovales en cœur; jeunes pousses sillonnées anguleuses; capsule globuleuse, marquée de 6 sillons, à 3-4 valves	P. monilifera 3313

FIGURE 1.10 – Comment identifier les espèces du genre *Populus* selon Coste[Cos06]

La forme des feuilles apparaît finalement comme le dernier critère de décision qui va parfois permettre d'identifier l'espèce d'un individu après avoir longuement examiné ses fleurs et ses fruits pour arriver à connaître son genre. Dans certains cas, les espèces sont elles aussi séparées en utilisant les organes reproducteurs, qui fournissent des informations au moins aussi discriminantes. Tout juste quelques spécificités de l'agencement, du type général des feuilles sont elles égrenées au fil de la progression dans l'arborescence.

La démarche analytique commune en botanique pour identifier les espèces semble donc bien mal appropriée à leur reconnaissance par les feuilles. Le fait est qu'en se basant sur une classification des végétaux, il est évident que le rôle accordé aux feuilles ne peut être que minime. Puisqu'elles ne sont qu'un attribut phénotypique indirectement relié au processus clé sur lequel est bâtie la classification, à savoir la filiation, les feuilles ne peuvent qu'intervenir de façon tout aussi marginale dans la méthode d'identification qui en découle.

Il s'agit cependant d'une méthode créée par les experts et pour les experts. Il existe à côté de ces outils de référence des clés de détermination plus limitées, moins rigoureuses et départies du souci d'exhaustivité et de correspondance à une classification, qui vont se borner à considérer un ensemble très réduit d'espèces. Il en existe par exemple pour les essences forestières de feuillus¹ mais le nombre limité d'espèces considérées (quand elles ne s'arrêtent pas au genre) et les imprécisions scientifiques (noms vernaculaires, parfois ambigus) en font une base moins solide. Pour s'adresser à un néophyte, avec l'ambition de lui donner les moyens de franchir la barrière initiale qui rend l'accès à la botanique si ardu, il est indispensable de repartir des définitions botaniques établies, malgré toute leur complexité.

1. Clés de déterminations proposées par l'Office National des Forêts : http://www.onf.fr/activites_nature/++oid++8ac/@@display_advise.html

1.2 Petit lexique botanique appliqué aux feuilles

En dépit de leur inadéquation avec les critères fondamentaux utilisés pour la classification et l'identification des espèces de plantes dans leur globalité, les feuilles sont des organes végétaux qui présentent une grande diversité et dont les caractéristiques bien étudiées sont dénommées par tout un vocabulaire spécifique qu'il convient de décortiquer. Si elles ne servent pas à délimiter des ordres ou des familles, il n'en demeure pas moins qu'au niveau de l'espèce, et notamment lorsque l'on parle d'espèces d'arbres, les propriétés des feuilles peuvent s'avérer tout à fait discriminantes. Il faut avant tout définir ce qui constitue l'objet feuille, qui n'est pas nécessairement ce que le sens commun met derrière ce mot.

1.2.1 Qu'est-ce qu'une feuille ?

“ Les feuilles sont des organes généralement verts qui sont répartis le long des tiges et dans lesquels se déroule l'essentiel de l'activité photosynthétique des végétaux. ”

Encyclopédie LAROUSSE

Chez la plante, les feuilles sont donc simplement les organes dédiés à la photosynthèse. En d'autres termes, ce sont les choses vertes reliées à la tige, ou au bout des branches pour un arbre, et c'est bien là la seule définition générale qu'on peut en donner !

Diversité des feuilles

En effet, la diversité des formes et des spécialisations qu'elles peuvent prendre rend quasiment impossible une définition générale qui pourrait s'appliquer à l'ensemble des feuilles. On les voit tout aussi bien vertes et planes comme on le conçoit communément, que composées de dizaines de parties indépendantes, prenant formes d'aiguilles, d'écailles, ou même spécialisées en épines ou en vrilles. La Figure 1.11 illustre la diversité couverte par la notion de feuille.

Le fait est pourtant que toutes les feuilles ont à la base la même fonction, puisque ce sont les organes de la plante qui sont dédiés à la photosynthèse. Le limbe foliaire, la partie verte, est donc une surface déployée de façon à recueillir la lumière nécessaire à ce processus. C'est également par ses feuilles que la plante respire et transpire. Dans le cas des conifères, la feuille prend généralement la forme d'aiguille, afin de limiter la transpiration et mieux résister à la sécheresse. La perte de surface exposée à la lumière est compensée par le nombre d'aiguilles et leur disposition qui leur permet de présenter malgré tout la surface maximale au soleil.

Une unité de nature

Dans le cas des Angiospermes qui nous intéresse plus particulièrement, puisqu'ils regroupent l'ensemble des arbres dits feuillus, les feuilles sont assez semblables dans leur nature, mais peuvent



FIGURE 1.11 – Différentes formes de feuilles d'espèces d'arbres variées

prendre une variété impressionnante de formes différentes. Les botanistes utilisent des termes relativement précis pour permettre de décrire ces formes, et ce sont ces termes qui sont utilisés dans les ouvrages de flores ou dans les clés de détermination. A la différence des fleurs où la plupart des termes font référence à des éléments précis de leur anatomie qu'il faut avoir appris, ceux utilisés pour les feuilles sont essentiellement descriptifs et donc plus accessibles, même s'il faut les avoir rencontrés au moins une fois pour en comprendre le sens.

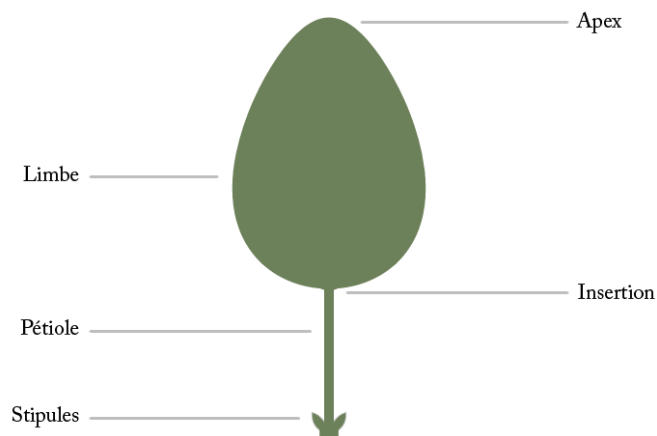


FIGURE 1.12 – Schéma détaillant la structure générale d'une feuille et les composants communs à l'ensemble des types de feuilles

La structure d'une feuille chez les Angiospermes représentée dans la Figure 1.12 est généralement toujours la même, bien qu'elle puisse donner lieu à des dégénérescences. La feuille est reliée à la tige de la plante par un **pétiole** qui est soit rattaché directement, soit par le biais d'une gaine. Le pétiole peut chez certains cas être absent, les feuilles sont alors dites sessiles. L'autre partie principale de la feuille est le **limbe** foliaire, tissu contenant généralement les cellules photosynthétiques qui s'étend à la suite du pétiole. Le limbe peut être d'un seul tenant ou divisé en plusieurs parties que l'on nomme

folioles, qui ont la même structure qu'un limbe simple, et qui sont reliées à un prolongement du pétiole : le **rachis**. Les derniers composants usuels de la feuille sont les **stipules**, pièces situées à la base de la feuille de part et d'autre du pétiole, bien qu'elles soient souvent très réduites voire absentes.

Plusieurs feuilles en une ?

Le fait que le limbe puisse être divisé est la source des plus grandes variabilités morphologiques. En effet les feuilles que l'on va alors appeler **composées** ne constituent bien qu'une seule et unique feuille malgré la présence de plusieurs petits organes que l'on pourrait eux aussi apparenter à des feuilles. Il s'agit bien là de folioles, donc de pièces d'une feuille, qui se trouve par conséquent assez éloignée de sa représentation canonique. La façon dont sont organisées ces folioles à partir du pétiole est elle aussi hautement variable, et donne naissance à une diversité conséquente de formes possibles visibles sur la Figure 1.13. Certaines espèces comme l'arbre à soie (*Albizia julibrissin* Durazz.) poussent le vice jusqu'à voir leurs folioles elles-mêmes composées.

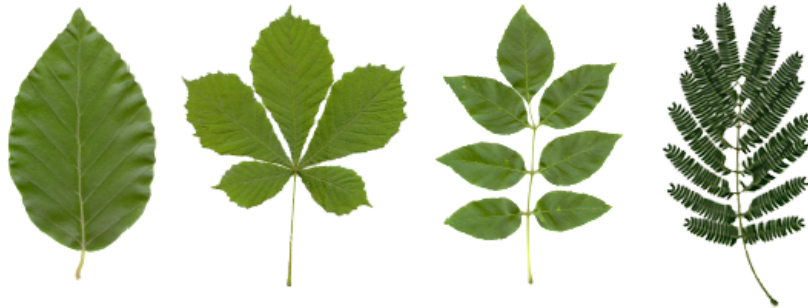








FIGURE 1.13 – Feuilles simple de hêtre (*Fagus sylvatica* L.) digitée de marronnier (*Aesculus hippocastanum* L.) imparipennée de frêne à fleurs (*Fraxinus ornus* L.) et bipennée d'arbre à soie (*Albizia julibrissin* Durazz.)

On distingue ainsi plusieurs organisations bien distinctes qui compliquent légèrement la définition de ce qu'est la feuille du point de vue de la forme. Ces distinctions sont pourtant capitales tant la composition des feuilles est parfaitement caractéristique de l'espèce. Les compositions possibles sont les suivantes :

	Simple	limbe d'un seul tenant, continu
	Trifoliée	limbe divisé en trois folioles uniquement qui se rejoignent à l'extrémité du pétiole
	Digitée	limbe divisé en plusieurs folioles qui se rejoignent toutes à l'extrémité du pétiole
	Paripennée	limbe divisé en un nombre pair de folioles échelonnées le long du rachis
	Imparipennée	limbe divisé en un nombre impair de folioles avec une foliole terminale
	Bipennée	limbe divisé en folioles disposées le long du rachis et elles-mêmes pennées

Cette diversité fait qu'il est particulièrement difficile de modéliser la forme des feuilles d'une façon unique, qui serait valable quelle que soit la feuille considérée. La seule caractéristique qui semble pouvoir s'appliquer à l'ensemble des feuilles est leur bilateralité, la présence d'un axe de symétrie, ou à plus proprement parler d'une direction principale de développement de la feuille. C'est généralement la **nervure principale** des feuilles simples qui matérialise cet axe, ou bien le rachis dans le cas des feuilles composées.

Le développement des feuilles est donc directement lié à leur forme, et ce par le processus de leur croissance. C'est un point qui a intéressé les botanistes, et on peut citer le français VAN TIEGHEM un des premiers à incorporer des considérations géométriques et des modélisations mathématiques dans l'étude de la structure des plantes. L'essentiel du développement de la feuille se produit en fait dans le bourgeon terminal [VT98] et c'est à ce stade que vont se produire les ramifications éventuelles, et non pas dans une hypothétique transformation à l'air libre. La croissance une fois la feuille épanouie est alors essentiellement intercalaire, même si les nervures continuent de se ramifier, mais elle est généralement localisée, selon les espèces, à la base ou au sommet (que l'on préfère nommer **apex**) de la feuille.

1.2.2 Descriptions des formes foliaires

Pour décrire l'ensemble du spectre des formes ainsi prises par les feuilles, les ouvrages botaniques utilisent un vocabulaire assez précis. Il s'agit d'un ensemble de termes, peu ou pas usités en dehors de ce contexte, qui vont s'évertuer à qualifier les différentes propriétés morphologiques de la feuille. Ce sont ces critères qui permettent de discriminer les espèces, et qui sont parfois mentionnés dans les clés de détermination. Il faut donc se familiariser avec cette terminologie qui, en plus de permettre l'interprétation des ouvrages botaniques, permet d'avoir une idée assez claire de ce qu'il est nécessaire d'observer pour identifier une feuille.

Forme générale et nervation

La première caractéristique, la plus facilement observable, est la forme générale du limbe de la feuille, ou de celui des folioles dans le cas d'une feuille composée. Il est à noter que les folioles d'une feuille composée présentent généralement les mêmes formes, se contentant de varier de façon progressive, à l'exception de l'éventuelle **foliole terminale**, celle située à l'extrémité supérieure du rachis, qui peut avoir sa forme propre. Dans cette impression globale de la forme de la feuille, on met avant tout en avant la présence de lobes, c'est-à-dire de grandes séparations du limbe. Ces lobes sont à mettre en relation avec la nervation de la feuille, puisqu'ils témoignent en fait de plusieurs directions principales de développement. On distingue ainsi plusieurs types de feuilles :
















	Simple	limbe continu, sans séparation apparente
	Palmatilobée	limbe séparé en plusieurs lobes aux nervures se rejoignant à l'extrémité du pétiole
	Pennatilobée	limbe séparé en plusieurs lobes aux nervures échelonnées sur la nervure principale



FIGURE 1.14 – Feuilles palmatilobée d'érable plane (*Acer platanoides* L.), lobée à nervation palmée de peuplier blanc (*Populus alba* L.), pennatilobée d'aubépine (*Crataegus monogyna* Jacq.) et lobée à nervation pennée de chêne pubescent (*Quercus pubescens* Willd.)

Malheureusement, comme le montre la Figure 1.14 la frontière est mince entre ce qui constitue indubitablement un lobe, nettement séparé, et les découpages et ondulations moins structurales mais plutôt caractéristiques du bord du limbe. L'origine en est pourtant la même, visible clairement à travers le réseau des nervures. Le problème vient donc surtout du fait que ces grandes catégories morphologiques sont un peu artificielles et viennent se superposer à des objets naturels qui peuvent prendre des formes suivant un spectre bien plus continu. Cependant ces distinctions s'avèrent bien utiles dans le cas où elles sont flagrantes, même si les cas limites peuvent poser problème.

Pour qualifier la forme générale des feuilles simples, ou des lobes des feuilles qui en comportent, la botanique emploie des mots vaguement géométriques qui cherchent en tout cas à associer une forme de référence à la feuille observée. Il existe ainsi un certain nombre de qualificatifs, dont les plus souvent rencontrés sont les suivants :

	Linéaire	très étroite et allongée, aux bords quasiment parallèles
	Lancéolée	étroite et allongée, pointue au sommet et généralement plus large à la base
	Elliptique	terme vague, étroite ou large, généralement arrondie et égale à la base et au sommet
	Oblongue	allongée, aux bords quasiment parallèles, arrondie à la base et au sommet
	Ovale	en forme d'œuf, de largeur moyenne, plus large et arrondie à la base qu'au sommet
	Obovale	opposé à ovale, de largeur moyenne, plus large et arrondie au sommet qu'à la base
	Rhomboidale	de la forme d'un losange, de largeur moyenne, égale à la base et au sommet
	Spatulée	assez étroite, plus large et arrondie au sommet, la base s'étirant vers le pétiole
	Orbiculaire	de la forme d'un cercle, très large et très arrondie à la base et au sommet
	Suborbiculaire	presque orbiculaire, un peu moins large que longue mais très arrondie
	Triangulaire	de la forme d'un triangle, plus ou moins large à la base, très pointue au sommet
	Cordée	de la forme d'un cœur, large et incurvée à la base, plus ou moins pointue au sommet

Une fois encore le problème que l'on rencontre avec cette terminologie est qu'il ne faut en aucun cas la prendre comme un ensemble de catégories bien délimitées. Les feuilles, en tant qu'objets naturels, voient leurs formes globales varier de manière assez continue, plus ou moins larges, plus ou moins allongées, plus ou moins arrondies, ainsi que l'illustre la Figure 1.15.



FIGURE 1.15 – Variations continues de la forme globale de feuilles de châtaignier (*Castanea sativa* Mill.), passant progressivement de lancéolées à oblongues

Pourtant, les mots qui vont les qualifier renvoient à l'idée d'une forme unique, canonique, qui ne saura jamais correspondre rigoureusement à la forme véritable de la feuille. Il faut donc plutôt voir ces termes comme des jalons, des points connus de l'espace des formes possibles, auxquels on peut se référer pour situer une forme donnée. Bien plus que l'appartenance à une classe, c'est la similitude à

un modèle qui est évoquée par ces termes, et une même feuille pourra tout à fait se retrouver à cheval entre plusieurs formes.

Ces termes de forme générale sont donc aussi ceux utilisés pour qualifier la forme spécifique d'un lobe ou d'une foliole. On pourra ainsi rencontrer des feuilles palmatilobées à lobes triangulaires (érable de Montpellier, *Acer monspessulanum* L.), des feuilles digitées à folioles obovales (marronnier, *Aesculus hippocastanum* L.), ou encore tout bonnement des feuilles imparipennées à folioles lancéolées (frêne, *Fraxinus excelsior* L.).

Comme on l'a vu, c'est la nervation qui tient un rôle central dans la forme générale de la feuille comme dans sa structure. Les types de nervations sont peut-être plus facilement catégorisables que les formes qu'elles produisent, il n'y a en effet pas la même continuité dans la façon dont elles s'agencent, une feuille étant soit d'un type soit d'un autre. On est donc cette fois-ci bien en présence de classes de nervation, qui sont elles aussi utilisées dans la description des formes foliaires :








	Uninervée	une seule nervure principale suivant l'axe principal de la feuille
	Parallèle	plusieurs nervures orientées selon l'axe principal de la feuille sans intersection
	Pennée	plusieurs paires de nervures secondaires échelonnées le long de la nervure principale
	Alterne	plusieurs paires de nervures secondaires mais pas en face les unes des autres
	Arquée	plusieurs paires de nervures secondaires se rejoignant vers l'apex de la feuille
	Palmée	un nombre impair de nervures principales divergeant depuis la base de la feuille
	Réticulée	nervures secondaires et tertiaires formant un réseau complexe
















FIGURE 1.16 – Les nervures au plus près du bord










Des formes plus locales

Selon que les nervures atteignent ou non le bord du limbe, celui-ci présentera des dents, plus ou moins marquées, ou au contraire de simples ondulations voire un profil parfaitement lisse (Figure 1.16). La caractérisation de la forme du bord est indispensable pour s'y retrouver dans l'ensemble des formes possibles. D'autant plus que la présence éventuelle de lobes est parfois assimilée à une caractéristique du bord. Ce n'est au final pas si surprenant puisque ce qui se produit sur le bord n'est jamais qu'une version locale de ce qui fait la structure de la feuille, au vu de son réseau de nervures. C'est donc le problème de l'échelle qui va ici compliquer les choses, tous les termes utilisés












ne renvoyant pas nécessairement à des éléments de la même taille. On peut donc se retrouver avec des feuilles cumulant plusieurs qualificatifs, ce qui rajoute un peu à l'ambiguïté et à la subjectivité de leur définition. Les principales formes de bord sont les suivantes :

	Entière	bord parfaitement lisse, sans aucune aspérité ni aucune échancrure
	Ciliée	bord présentant à sa marge de petits poils ou cils fins
	Ondulée	bord sans aspérités mais présentant des sinuosités larges et peu marquées
	Crénelée	bord présentant des dents de petite taille et très arrondies
	Denticulée	bord présentant des dents marquées, mais très fines
	Dentée	bord présentant des dents bien marquées, qu'elles soient de taille petite ou moyenne
	Dentée en scie	bord présentant des dents bien marquées, toutes orientées vers l'apex de la feuille
	Doublement dentée	bord présentant deux niveaux de dents, les petites entre les plus grandes
	Épineuse	bord présentant dents larges et très pointues, voire des épines
	Anguleuse	bord sans dents mais présentant angles larges et assez marqués
	Sinuée	bord sans dents mais présentant des sinuosités marquées, de taille moyenne
	Lobée	bord présentant des lobes plus ou moins marqués, et de larges sinus
	Pennatifide	bord présentant des sinus larges et profonds, presque jusqu'à la nervure

Enfin, plus localement, les formes que prend le bord du limbe à sa base et à son apex peuvent être décrites plus ou moins géométriquement, à la manière de ce qui est fait pour la forme générale. La base de la feuille correspond à l'insertion du pétiole, et c'est à partir de ce point que naissent la ou les nervures principales, qui constituent le prolongement du pétiole dans le limbe. En fonction des angles que font les nervures secondaires à cet emplacement, le tissu foliaire va être développé dans des directions plus ou moins larges, et le bord résultant du limbe va pouvoir suivre des formes diverses, dont les plus courantes sont :

	Atténuée	limbe très étroit à la base, qui se prolonge en pétiole
	En coin	limbe assez étroit à la base, qui forme un angle aigu assez marqué
	Arrondie	limbe plutôt large à la base, qui forme un angle obtus arrondi et peu marqué
	Oblique	limbe à la base dissymétrique, l'angle s'ouvrant plus largement d'un côté que de l'autre
	Décrochée	limbe à la base dissymétrique, un côté du limbe s'attachant plus bas que l'autre
	Tronquée	limbe aplati à la base, qui forme un angle plat sur une largeur plus ou moins importante
	En Cœur	limbe large à la base, qui forme un angle rentrant, qui s'arrondit de part et d'autre
	Echancrée	limbe arrondi à la base, qui forme une encoche plus ou moins profonde
	Auriculée	limbe assez large à la base, qui forme deux petites bosses de part et d'autre

L'apex, dénomination utilisée en botanique pour toutes les extrémités des organes, correspond donc au sommet de la feuille, où vient généralement se terminer la nervure principale. A son sommet, la forme d'une feuille est une fois encore à mettre en rapport avec la façon dont viennent se terminer toutes les nervures secondaires près du bord du limbe, les formes prises par celui-ci à cet endroit donnent lieu à différents termes, dont on peut citer les plus représentés :

	Acuminée	limbe très étroit au sommet, qui se termine par une pointe effilée plus ou moins longue
	Aiguë	limbe se rétrécissant au sommet, en formant une pointe avec un angle aigu
	Obtuse	limbe se rétrécissant au sommet, en formant une pointe avec un angle obtus
	Épineuse	limbe se rétrécissant au sommet, qui se termine par une pointe voire une épine
	Recourbée	limbe se rétrécissant au sommet, qui se termine par une pointe inclinée d'un côté
	Cuspidée	limbe se rétrécissant au sommet, qui se termine par une pointe courte marquée et rigide
	Mucronée	limbe se rétrécissant au sommet, qui se termine par une pointe courte dure et arrondie
	Arrondie	limbe assez large au sommet, sans pointe mais s'arrondissant
	Tronquée	limbe aplati au sommet, qui forme un angle plat plus ou moins large
	Rétuse	limbe assez large au sommet, qui forme un angle rentrant
	Echancrée	limbe assez large au sommet, qui forme une encoche marquée

Ici le problème est double. D'une part on ne sait pas exactement quelle portion de la feuille est qualifiée par ces termes ; est-ce qu'il s'agit simplement de la petite pointe qui se trouve tout à l'extrémité de la feuille ou bien la façon dont tout le tiers supérieur s'arrondit ou se rétrécit ? C'est en pratique un peu au cas par cas, certains termes correspondant à des spécificités (pointes rigides, échancrures) qui ne peuvent pas être ignorées, d'autres se contentant de mettre un mot sur l'angle formé par le sommet de la feuille, tous ne pouvant s'appliquer en même temps.

Et d'autre part, du moment que les qualificatifs renvoient à des formes géométriques de référence, on se trouve dans une problématique qui n'est pas véritablement de la classification, mais à nouveau un exercice d'association par similitude. Il n'est pas difficile d'imaginer comment les bases et les sommets des feuilles peuvent combler l'espace qui existe entre ces formes canoniques, comme l'illustre la Figure 1.17 et les frontières seront fatalement floues et la subjectivité de mise.



FIGURE 1.17 – Variations continues de la forme de l'apex de feuilles d'olivier (*Olea europaea* L.), passant progressivement de aiguës à obtuses

La forme du limbe est ce qui concentre le plus de variabilité et se trouve être par suite la plus richement fournie en termes botaniques. Il existe cependant d'autres critères importants qui vont jouer un rôle dans la descriptions des espèces et que l'on peut rajouter aux outils de caractérisation des feuilles :

- ▶ Taille : explicitement en cm, ou qualifiée de petite, assez grande, grande, longue
- ▶ Longueur du pétiole : en fonction du limbe, ou feuille longuement, courtement pétiolée, voire (sub)sessile (absence de pétiole)
- ▶ Couleur : souvent décrite bien que variable avec les saisons, surtout pour la distinction dessus-dessous
- ▶ Pilosité : présence de poils, qualifiée de glabre, glabrescente, pubescente, soyeuse, laineuse, tomenteuse
- ▶ Rigidité : définit la consistance de la feuille, qualifiée de molle, épaisse, ferme, rude, coriace
- ▶ Luisance : qualifiée de mate ou luisante, parfois pour la distinction dessus-dessous

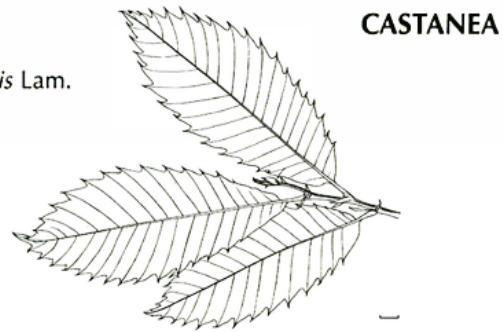
Ces caractéristiques, peut-être moins morphologiques, sont malgré tout des indices importants qui sont accessibles à qui tient la feuille entre ses mains, et renvoient donc à une identification de terrain. En observant une photographie en revanche, certains sont plus difficiles à observer, voire impossibles si on ne dispose que d'un seul angle de vue. Moins visuellement immédiats, ils restent tout de même souvent discriminants pour reconnaître une espèce à coup sûr.

1.2.3 Formes et espèces

Chacune des espèces au sein des Angiospermes peut ainsi être caractérisée au regard de la forme de ses feuilles. Même si cela fait rarement partie des critères primordiaux utilisés pour déterminer l'espèce en suivant sa famille et son genre, selon l'arbre arpenté par les clés, les critères morphologiques de la feuille n'en demeurent pas moins propres aux espèces. Chaque individu d'une même espèce possède donc des feuilles du même type, malgré une irréductible variabilité naturelle, influencée par l'environnement et l'âge de la plante. Cette variabilité intra-espèce diffère selon les critères considérés, et n'est naturellement pas la même pour toutes les espèces.

FAGACÉES

Castanea sativa Mill. = *C. vulgaris* Lam.
Châtaignier



CARACTÈRES DIAGNOSTIQUES

– feuilles **alternes**, grandes, **lancéolées**, **dentées** (longues de 10-20 cm et larges de 4-8 cm), à **nervation pennée**, **glabres** et **brillantes dessus**, légèrement pubescentes et à **nervures saillantes dessous**;

FIGURE 1.18 – Description des feuilles du châtaignier (*Castanea sativa* Mill.) tirée de la flore de RAMEAU [RMD⁺89]

On trouve ainsi dans tous les ouvrages de flore, à la page décrivant l'espèce (soit généralement le point d'arrivée de la clé de détermination préalable) un recensement complet des propriétés des feuilles, ainsi qu'illustré sur la Figure 1.18. Il est donc possible de répertorier pour chaque espèce une liste de propriétés morphologiques que vont partager les feuilles de ses individus. Ce ne sont pas nécessairement de critères discriminants en eux-mêmes, mais ils constituent déjà une fiche d'identité descriptive de la feuille associée à une espèce. La liste des critères de forme recensés pour les espèces d'arbres et d'arbustes retenues comme cadre de travail figure dans l'Annexe A. C'est aussi dans ces descriptions que vont se révéler les ambiguïtés liées à la définition des termes botaniques.

Il n'est en effet pas rare de trouver plusieurs termes de forme accolés à une même espèce, comme le micocoulier (*Celtis australis* L.) aux feuilles **ovales-lancéolées** selon COSTE [Cos06]. Il est alors difficile de savoir si l'espèce peut, dans sa diversité, présenter des feuilles dont la forme générale va couvrir continument le spectre de formes entre les formes canoniques ovale et lancéolée, ou au contraire si ses feuilles vont toutes se ressembler, d'une forme intermédiaire entre ces deux références. Il est de toute façon peu vraisemblable d'affirmer que les formes prises par une espèce soient unique, il va toujours s'agir d'apprécier la largeur de l'ensemble des formes possibles, ce qui laisse inévitablement place à l'interprétation.

Un autre cas de figure des plus courants est l'absence de caractérisation d'une certaine propriété.

La forme générale ou celle du bord sont généralement renseignées, mais il peut arriver qu'il ne soit pas fait mention de la forme au sommet, par exemple. Faut-il alors comprendre que le sommet peut décrire l'ensemble des formes existantes ? Ou simplement qu'il est trop variable pour être pris comme une caractéristique constitutive de l'espèce ? Mais même dans ce cas, l'espèce ne prendra des formes que dans une plage limitée, bien que trop large pour être explicitée.

Quoiqu'il en soit, il s'agit là d'un outil très performant dès lors que l'on restreint l'ensemble d'espèces considérées. En se limitant aux arbres, une fois connus pour toutes les espèces les ensembles de formes possibles de chaque critère, on dispose d'un crible assez complet qui permet à un observateur de reconnaître une espèce uniquement à partir de sa feuille. Ce sont bel et bien ces caractéristiques qui sont mises à contribution dans les clés de détermination des feuillus déjà existantes, malgré le fait qu'elles soient destinées à un public débutant et donc souvent délestées de la terminologie consacrée. Il existe certes un certain recouvrement entre certaines espèces aux formes très proches, mais le doute pourra bien vite être levé entre l'éventuelle poignée d'espèces correspondant à un jeu de formes par quelque considération sur l'allure générale de la plante, de ses fruits ou de ses fleurs.

Par conséquent, il est bien possible de s'appuyer sur la description morphologique des feuilles qui est employée par les botanistes pour bâtir un système d'identification des espèces d'arbres. Un système qui diffère totalement des clés de détermination puisqu'il ne s'appuierait pas sur l'arborescence de la classification des espèces mais utiliserait les définitions botaniques des formes foliaires propres aux espèces pour prendre des décisions.

1.3 La botanique comme guide pour la reconnaissance automatique

La botanique fournit par définition des méthodes analytiques menant à l'identification des espèces végétales. Ce n'est malgré tout pas ce qui va permettre à un expert de reconnaître au premier coup d'œil une espèce qu'il a déjà rencontrée, une identification qui se fait de manière bien moins consciente, certainement par une foule de détails visuels acquis, fruits de son expérience, qu'il aurait bien du mal à expliciter. En revanche, le débutant va devoir s'engager dans cette voie laborieuse d'observation et d'annotation de critères pour arriver au résultat escompté.

1.3.1 L'analyse par la forme

L'idée de transposer ces méthodes d'analyse dans un cadre automatique est séduisante mais elle pose un certain nombre de problèmes. Tout d'abord à cause de l'acquisition des données qui ne peut que différer de celle d'un être humain observant, tâtant, disséquant un objet sous toutes ses coutures. Le comportement de la machine diffère lui aussi de celui de l'observateur, qui est capable d'interprétation et de subjectivité, de jugements au cas par cas de ce qu'il est pertinent de prendre en compte. Enfin, est-ce bien le processus analytique de la détermination qui est souhaitable, ou ne vaudrait-il pas mieux s'inspirer de la reconnaissance immédiate et implicite de l'œil exercé ?

Mesures géométriques et termes botaniques

Le cas des feuilles est tout particulièrement intéressant en ceci que, étant généralement des objets plans, et leurs propriétés se résumant essentiellement à des appréciations géométriques, une image suffit à capturer l'essentiel de ce qu'il y a à analyser pour produire une caractérisation au sens botanique. Un observateur humain serait en effet capable de déterminer la structure des folioles, compter les lobes, qualifier la dentition ou mesurer l'angle formé par le sommet de la feuille. On peut donc espérer que la machine fasse de même.

Il est en effet parfaitement envisageable d'automatiser intégralement une procédure qui vise dans un premier temps à la compréhension de la structure et des composants de la feuille, puis à l'évaluation numérique de ce qui permet déjà à l'observateur d'apposer des mots sur les formes. Les définitions des termes botaniques évoquent la largeur, la longueur, l'axe principal, l'angle, la courbure, la circularité, autant de choses qui sont par essence mesurables. Le seul véritable obstacle pour la machine est alors de déterminer où effectuer ces mesures, ce qui suppose une inévitable étape d'interprétation des données qui lui sont soumises, en y plaquant des concepts humains. Cette étape que toute personne est capable de mener à bien quasi-instinctivement, constitue un défi pour la machine, d'autant plus exigeant qu'elle est cruciale pour que les mesures effectuées par la suite aient un quelconque sens.

Le fait de voir la question de l'analyse de la forme des feuilles comme un problème purement géométrique n'est pas nouveau, et depuis les premiers botanistes-mathématiciens du XIX^{ème}, des modèles mathématiques plus rigoureux ont été proposés [Gro99]. Ce n'est cependant pas la façon dont sont caractérisées les espèces qui en constitue le sujet d'étude, mais bien l'objet feuille lui-même, et les ouvrages botaniques à proprement parler continuent d'utiliser le vocabulaire spécifique pour décrire les formes foliaires.

Un vocabulaire à transmettre

Cela a donc un sens de chercher non seulement à extraire les grandeurs numériques recouvertes par les termes botaniques, mais également de s'en servir pour mettre un mot sur les formes ainsi analysées. En effet, être capable de reconnaître automatiquement les différentes formes existantes, et surtout de les présenter à un utilisateur peut permettre de lui rendre intelligible un vocabulaire qui, sans cela, lui serait sans doute resté largement hermétique. La maîtrise de cette terminologie est cruciale puisqu'elle constitue, du moins en ce qui concerne les feuilles, un point d'entrée obligatoire vers les flores et autres ouvrages botaniques de référence, qu'il est difficile d'aborder par soi-même.

L'apposition automatique de ces termes n'est pas à proprement parler un problème de classification, puisque, ainsi qu'on l'a vu, ils ne définissent pas pour la plupart des catégories rigoureusement fermées. On est plutôt en présence d'un problème d'annotation ou d'étiquetage qui n'en demeure pas moins éminemment compliqué. En effet, ce sont souvent des concepts sans définition formelle bien établie, et qui laissent la part belle à l'interprétation de l'observateur. Une feuille va être facilement qualifiée d'obovale du moment qu'elle est nettement plus large vers son sommet, et ce indépendamment de sa

largeur, de sa circularité ou d'autres facteurs qui rentrent en ligne de compte dans la définition des autres concepts. Il y a tout de même de quoi faire avec l'idée que chacun renvoie malgré tout à une forme canonique à laquelle il est possible de se comparer.

L'introduction de la démarche botanique et du vocabulaire qui y est associé a donc un intérêt double. Il permettent à la fois d'aider à circonscrire les mesures à effectuer dans le cadre d'une reconnaissance analytique, qui ne sont autres que celles préconisées par les botanistes pour parvenir à nommer les formes. Cela fournit un guide de la manière dont il faut regarder une feuille si on souhaite la caractériser et l'identifier. Enfin, ils constituent un savoir qu'il est particulièrement intéressant de transmettre à un néophyte en botanique, puisqu'il lui permettra peut-être de se plonger plus aisément dans les arcanes de cette science.

1.3.2 Des difficultés du problème d'identification

Caractériser les feuilles en suivant l'approche analytique dérivée de la botanique ne permet toutefois pas de résoudre à coup sûr le problème de la détermination de l'espèce. Et de fait, ce dernier recèle de nombreux écueils qui en font un problème de classification des plus complexes. Idéalement, les classes que représentent ici les espèces, devraient pouvoir être bien séparées selon des critères qui pris dans leur totalité permettraient de les discriminer. Une telle chose ne se produit que rarement dans la nature, et les possibilités sont telles que lorsqu'on se penche sur les formes, on se trouve en face de classes bien mal agencées.

Variabilité intra-espèce et similitude inter-espèce

L'un des principaux problèmes rencontrés lorsque l'on considère la caractérisation des feuilles d'une espèce donnée à l'aide des termes botaniques provient du fait qu'un simple archétype ne parvient pas à recouvrir la variété conséquente de formes que peuvent produire la croissance, les facteurs environnementaux, et les aléas naturels. Même si toutes les feuilles d'une même espèce sont issues du même processus de croissance, et partagent tout de même en cela un certain nombre de caractéristiques communes, certains résultats peuvent différer du tout au tout, comme par exemple la forme globale voire même le nombre de lobes des feuilles présentées dans la Figure 1.19.



FIGURE 1.19 – Variabilité des feuilles de peuplier blanc (*Populus alba* L.) en termes de forme générale

Le fait est que certains critères vont pouvoir se montrer extrêmement capricieux pour une espèce donnée alors que d'autres resteront stables à travers les individus, et que ce sera l'inverse chez une



FIGURE 1.20 – Variabilité des feuilles de houx (*Ilex aquifolium* L.) en termes de bord

seconde espèce comme le démontre la Figure 1.20 où la forme globale est constante quand le bord varie du tout au tout. Cette grande variabilité est un problème dans l'optique d'une classification, puisque les classes ainsi considérées peuvent en fait être très larges, sans que l'on sache *a priori* où se situe cette largeur.

Cela ne serait rien sans la difficulté conjointe qui est que la similitude inter-espèce peut parfois excéder cette variabilité intra-espèce. En d'autres termes, il peut se trouver qu'en observant la forme on trouve une feuille plus proche d'une feuille d'une espèce différente que d'une autre feuille de la même espèce ainsi que l'illustre la Figure 1.21. On va certes toujours pouvoir trouver un facteur qui va permettre de distinguer deux feuilles en particulier, mais l'important est de mettre en rapport la proximité qui peut exister entre deux individus de classes distinctes et la distance qu'on peut trouver entre deux individus d'une même classe.



FIGURE 1.21 – Similitude des feuilles, de gauche à droite, du houx (*Ilex aquifolium* L.), du cornouiller mâle (*Cornus mas* L.), du laurier-tin (*Viburnum tinus* L.) et du paliure (*Paliurus spina-christi* Mill.)

Ces réalités viennent ainsi compliquer un problème de classification automatique dans lequel l'hypothèse selon laquelle deux individus proches ont de grandes chances de faire partie de la même classe ne tient plus tout à fait. C'est en tout cas un élément capital à prendre en compte dans la définition de ce qui caractérise une espèce pour la machine.

La place de l'incertain

Construire un modèle unique par espèce, une forme archétypale à laquelle toutes les feuilles de l'espèce pourraient se rapporter, semble donc assez peu réaliste. La signature définissant une espèce se doit de comporter une forme d'incertitude et de souplesse. Certaines des caractéristiques de la feuille pourront, et devront, être fixées avec exactitude lorsque ce sont des traits qui se retrouvent

chez tous les spécimens, mais d'autres devront être définies de manière incertaine, non comme des grandeurs pouvant varier plus ou moins autour d'un centre canonique, mais comme des paramètres pouvant prendre indifféremment des valeurs diverses.

Cette idée est aussi à mettre en lien avec le fait que certaines caractéristiques ne soient pas renseignées dans les descriptions botaniques des feuilles des espèces. Cela renvoie à une notion d'incertitude, d'indétermination, du fait qu'une caractéristique en particulier peut n'apporter que peu d'information pour l'identification d'une espèce donnée. En revanche la donnée d'un certain critère peut s'avérer crucial pour le choix de retenir ou d'éliminer une autre espèce. La façon pour la machine de représenter les espèces devrait donc clairement stipuler quels caractères sont des traits certains et donc déterminants, et où, à l'opposé, se situe la latitude et l'incertitude sur d'autres paramètres.

La botanique est dans tous les cas particulièrement riche d'enseignements pour le problème de l'identification des espèces à partir de leur feuille. La définition même de l'espèce permet de se rendre compte à quel point la feuille est marginale dans ce qui permet leur classification et leur identification. Pourtant, en se penchant sur le cas plus restreint des feuilles d'arbres, et en leur appliquant la terminologie descriptive bien usitée des botanistes, on remarque une diversité suffisante pour être convaincu de la pertinence de leur utilisation comme organe discriminant. La détermination n'est alors pas sans obstacle, loin s'en faut, tant la nature s'évertue à rendre ses formes indiscernables, mais l'utilisation des outils botaniques d'analyse de la forme des feuilles constitue déjà une piste incontournable pour l'étude du problème. Il s'agit là de plus, dans le contexte d'un outil automatique interactif, d'un moyen de familiariser un naturaliste en herbe avec un domaine scientifique passionnant, qui lui serait probablement resté étranger.

Bibliographie

- [BL09] Gaston Bonnier and G. De Layens. *La végétation de la France, Suisse et Belgique 1ère partie - Flore complète portative de la France et de la Suisse*. Librairie Générale de l'enseignement, Paris, 1909.
- [Can24] Augustin Pyrame de Candolle. *Prodromus systematis naturalis regni vegetabilis*. Parisii : Sumptibus Sociorum Treuttel et Wurtz, 1824.
- [Cos06] H. Coste. *Flore descriptive et illustrée de la France de la Corse et des contrées limitrophes*. Librairie de l'Académie impériale de Médecine, Paris, 1906.
- [FF88] P. Forey and C. Fitzsimons. *Arbres, Nature-Poche*. Gründ, 1988.
- [Gro99] Leaf Architecture Working Group. *Manual of Leaf Architecture - morphological description and categorization of dicotyledonous and net-veined monocotyledonous angiosperms*. Department of Paleobiology Smithsonian Institution, Cornell University Press, 1999.
- [Gro09] The Angiosperm Phylogeny Group. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants : Apg III. *Botanical Journal of the Linnean Society*, 161(2) :105–121, 2009.
- [Jus89] Antoine-Laurent de Jussieu. *Genera plantarum, secundum ordines naturales disposita juxta methodum*. Horto Regio Parisiensi exaratum, 1789.
- [Lin53] Carl von Linné. *Species plantarum : exhibentes plantas rite cognitatas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas*. Holmiae : Impensis Laurentii Salvii, 1753.
- [RBCF13] Georges Rouy, Jean Nicolas Boulay, E.-G. Camus, and Julien Foucaud. *Flore de France ; ou description des plantes qui croissent spontanément en France, en Corse et en Alsace-Lorraine*. Asnières (Seine), 1893-1913.
- [RMD⁺89] J.C. Rameau, D. Mansion, G. Dumé, J. Timbal, A. Lecointe, P. Dupont, and R. Keller. *Flore forestière française : Guide écologique illustré*. Institut pour le Développement Forestier, 1989.
- [VT98] Philippe Edouard Léon Van Tieghem. *Eléments de botanique*. Masson et cie., 1898.

Chapitre 2

Etat de l'art

Travaux parallèles et problèmes adjacents



Castanea sativa Mill.

Résumé du Chapitre

Ce chapitre propose une étude comparative des différents travaux menés sur le même sujet, ainsi que des problèmes connexes que l'on rencontre quand on s'intéresse à l'identification automatique des espèces végétales à partir d'images de leurs feuilles. Il détaille les méthodes utilisées dans des systèmes existants de reconnaissance de végétaux, en offrant une grille de lecture des différentes approches. Les différentes étapes du processus de reconnaissance sont étudiées en détail, en étudiant les solutions proposées dans la littérature pour répondre aux principaux problèmes soulevés par la reconnaissance des végétaux. Enfin, à partir de ces références, nous dégageons des pistes de travail pour répondre à nos problématiques particulières : l'analyse d'images naturelles complexes, leur interprétation en termes de descripteurs de forme discriminants, et la prise en compte de données hétérogènes pour mener à l'identification de l'espèce.

SES feuilles d'arbres, par leur facilité d'acquisition et leur intelligibilité en termes de formes géométriques, constituent un objet de choix pour le traitement automatique des images. D'ailleurs nombre de chercheurs de la communauté se sont confrontés au problème de leur identification. Dans les quinze dernières années, l'engouement s'est fait de plus en plus marqué, traduit par la soumission de benchmarks et de bases de données de référence (Swedish Leaves [Söd01], ICL [ICL10], Smithsonian Institute [BCF⁺08]) et par l'apparition de challenges dédiés à une communauté de plus en plus étendue (ImageCLEF depuis 2011 [GBJ⁺11, GBJ⁺12, GBJ⁺13]).

On peut relier ce mouvement à une tendance plus globale à replacer la nature un peu plus au centre des activités humaines, ou à la maturité suffisante des méthodes de vision artificielle permettant de s'attaquer à des objets naturels complexes. Également, avec l'émergence des technologies mobiles dites intelligentes, les usages se font plus évidents, et ce n'est pas un hasard si un certain nombre de projets partagent l'objectif de créer une application pour smartphone (LeafSnap, PlantNet, et Folia).

Toutefois, s'il est indispensable de considérer les travaux qui s'intéressent très exactement au même objet d'étude, le problème de la reconnaissance automatique de feuilles est liée à beaucoup de problèmes connexes qui tiennent tant de l'analyse d'image, notamment du problème de la segmentation d'images prises dans un environnement naturel, que de l'analyse et la description de formes d'un point de vue général, ou encore du problème de classification, avec un nombre de classes potentiellement élevé et des spécificités qui le rendent particulièrement épineux.

2.1 Traitement d'images pour la reconnaissance de plantes

L'identification des espèces végétales est un problème idéal pour le traitement d'images et la reconnaissance de formes, que ce soit dans un contexte de classification automatique d'images ou de recherche d'images par le contenu (CBIR - *content-based image retrieval*). Il s'agit en effet d'un problème bien posé, où les classes sont clairement délimitées (les espèces végétales formant, sauf hybridation, des ensembles distincts) et pour lequel on peut espérer qu'une simple image contienne suffisamment d'informations discriminantes pour représenter l'objet qu'elle contient et parvenir à l'identifier.

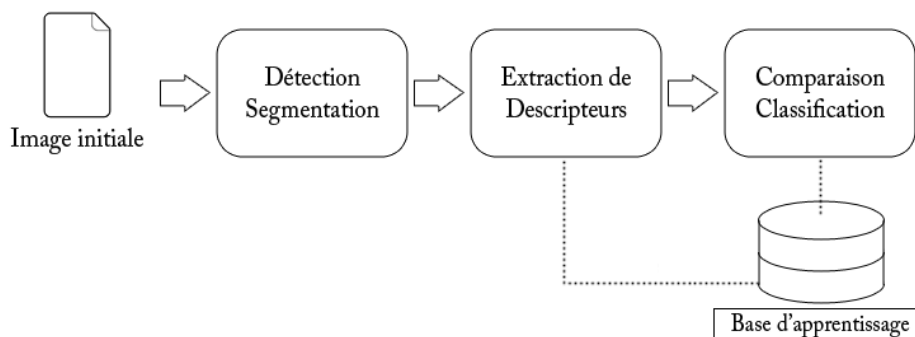







FIGURE 2.1 – Vue d'ensemble du fonctionnement des systèmes de reconnaissance d'objets

Les méthodes employées pour arriver à ce but que l'on peut trouver dans la littérature suivent toutes le même schéma général, somme toute assez classique dans le domaine de la reconnaissance d'objets. Ainsi que l'illustre la Figure 2.1, on peut distinguer trois grandes phases qui reviennent invariablement :

- ▶ **Extraction / Segmentation** : la plupart des méthodes commencent par considérer l'image dans sa globalité et la transformer en un objet qu'il sera ensuite possible de décrire, que ce soit une partition de l'image en régions distinctes (segmentation) ou un ensemble de points d'intérêt extraits sur l'image, pour citer les plus fréquents.
- ▶ **Description / Extraction de Caractéristiques** : Une fois les objets d'intérêt localisés et extraits de l'image brute, il s'agit de décrire leurs propriétés de façon complète et efficace ; les descripteurs vont alors capturer les informations de contour, de forme, de couleur, de texture s'il s'agit d'une région, ou de contexte local si l'on ne considère que des points.
- ▶ **Classification / Matching** : Les caractéristiques extraites sur l'image constituent une signature de l'objet qu'elle contient, qui peut être utilisée pour représenter la classe (l'espèce) qui lui est associée lorsqu'elle est connue. Si l'on dispose de suffisamment d'images ainsi étiquetées, elles forment une base d'apprentissage à laquelle il va être possible de comparer la signature d'une nouvelle image pour en déterminer la classe. Cela peut passer par une phase d'entraînement d'un classifieur ou par une comparaison individuelle à l'ensemble des images, qui se rapproche alors d'une recherche d'images.

C'est un schéma que l'on retrouve, avec évidemment des variations notables à chacune des étapes, dans les différents travaux consacrés à l'identification automatique d'espèces végétales qui ont vu le jour ces dernières années. La Table 2.1 récapitule ces différences de choix dans la conception et le déroulement des systèmes de reconnaissance automatique de végétaux étudiés.

2.2 Extraction d'objet et segmentation d'image

La façon dont est appréhendée l'image brute qui constitue le point de départ de tout algorithme de reconnaissance de plantes dépend énormément de l'objet d'étude qui est considéré. Selon l'organe qui est étudié (feuille , fleur , fruit , écorce , aiguille ) les méthodes à employer et les résultats escomptés ne peuvent être que différents. La modalité d'acquisition des images est aussi un point de divergence, et vient même définir des problèmes radicalement différents, selon que l'on travaille sur des images sur fond uni (et généralement blanc) ou des photographies prises en environnement naturel, avec plus ou moins de certitudes sur le contenu et l'agencement de la scène.

2.2.1 Extraction manuelle d'objet

Dans certains cas, l'extraction de l'objet d'intérêt se fait de façon complètement manuelle, requérant l'intervention assez extensive d'un utilisateur. Pour certains organes le choix d'une extraction manuelle




















Méthode	Objet	Extraction	Description	Reconnaissance
[AMK97] [MA04]		Seuillage (Uni)	Géométrique, Statistique	Recherche
[INK98]		Seuillage (Uni)	Modélisation	Classification
[WCFW00] [WCF03]		Seuillage (Uni)	Statistique	Recherche
[SK00] [SK04]		Segmentation (Naturel)	Modélisation, Géométrique, Couleur	Classification
[CNMJS06]		Manuelle	Statistique	Classification
[HZD ⁺ 06]		Manuelle	Texture	Classification
[NZ06] [NZ07]		Segmentation (Naturel)	Statistique, Vocabulaire, Couleur	Classification
[DWZ07]		Seuillage (Uni)	Géométrique, Statistique	Classification
[WBX ⁺ 07]		Seuillage (Uni)	Modélisation, Géométrique	Classification
[BCF ⁺ 08]		Seuillage (Uni)	Statistique	Recherche
[WHD ⁺ 08]		Segmentation (Naturel)	Statistique	Classification
[TKC09]		Segmentation (Naturel)	Statistique	Recherche(Forme)
[CA10]		Seuillage (Uni)	Géométrique, Statistique	Recherche
[FS11]		Points d'intérêt	Vocabulaire	Classification
[YAT11] ⚠ [YAT12]		Seuillage, Segmentation	Statistique, Texture, Couleur, Géométrique	Classification
[CBFMB11] ⚠ [CBFNGMB12]		Seuillage, Manuelle	Statistique, Texture	Classification
[BYM ⁺ 12] ⚠ [MYVB12]		Points d'intérêt	Locale, Statistique, Géométrique	Recherche
[AGB ⁺ 12] ⚠		Seuillage, Segmentation	Statistique, Modélisation, Géométrique	Classification
[KBB ⁺ 12]		Seuillage (Uni)	Géométrique	Recherche

TABLE 2.1 – Résumé des types d'approches de différentes méthodes de reconnaissance de plantes, en particuliers des travaux participant au challenge ⚠ ImageCLEF d'identification de végétaux

rapide et approximative peut faire sens, dans le cas par exemple où la forme de la région d'intérêt va apporter nettement moins d'information que les couleurs, textures, ou structures particulières contenues à l'intérieur.

C'est par exemple le cas pour les écorces d'arbres, où l'on ne va pas chercher à analyser la forme d'un tronc ou d'un morceau d'écorce mais bien les stries, les craquelures et les fissures présentes à la surface. La définition d'une région d'intérêt (RoI, *region of interest*) rectangulaire à la main est donc envisageable [HZD⁺06], d'autant que ce n'est pas une action très coûteuse en temps. Une initialisation rectangulaire peut de toute manière être vue comme un cadrage implicite, et donc dans le cas où l'utilisateur du système est responsable de la prise de vue, des consignes claires suffiraient à couvrir cette phase d'extraction (l'objet extrait étant alors l'image dans son ensemble).

Certains travaux considèrent ce cadrage de l'objet comme acquis, et ne s'embarrassent pas de traitement spécifique, considérant d'office l'image entière comme RoI, la plupart du temps pour ne pas se lancer dans un traitement potentiellement peu fiable sur des images complexes [YAT11].

En revanche, il arrive que l'interaction demandée à l'utilisateur soit de nature plus complète, et qu'il doive réaliser une réelle segmentation pour séparer l'objet d'intérêt de l'image [CBFMB11]. Si l'objectif est d'obtenir le contour exact de l'objet, il existe des outils pour l'aider comme la populaire baguette magique du logiciel Adobe[®] Photoshop[®] [CNMJS06] d'utilisation assez intuitive. On est alors proche de méthodes de segmentation interactive, à la différence près que toutes les décisions (tolérance de couleur, placement des points de départ, annulations éventuelles, nombre d'itérations) sont prises par l'utilisateur ; il serait ainsi plus exact de parler de segmentation passive.

Toutefois, dans la plupart des cas, ces méthodes d'extraction n'ont pas pour vocation d'être utilisées dans le contexte d'un réel système d'identification. Elles sont essentiellement là pour fournir un point de départ stable à la phase suivante de description, avec un moindre effort pour une tâche qui peut s'avérer particulièrement épineuse dans le cas d'images en environnement naturel.

2.2.2 Traitement global de l'image

Dans un contexte d'images simples, où l'objet d'intérêt se détache de façon évidente de son environnement que ce soit par sa couleur ou même simplement par sa luminosité, il est possible d'appréhender l'image dans sa globalité, sans prendre en compte de dimension spatiale. La segmentation résultante se fait alors en prenant une décision binaire pour chaque pixel (appartenance ou non à l'objet) au regard uniquement de l'information qu'il contient.

Ce genre de traitement est tout particulièrement adapté aux images de feuilles sur fond uni. Les feuilles, et en particulier les feuilles d'arbres, sont suffisamment planes et généralement assez petites pour passer dans un scanner, et il n'est pas surprenant que certaines bases de données d'images de feuilles (Swedish Leaves [Söd01], ICL [ICL10], PlantLeaves [GBJ⁺12] partiellement) soient composées d'images ainsi acquises. Dans ces images le fond est uniformément blanc et l'absence d'ombres portées et de reflets fait que la feuille ressort idéalement par rapport au fond, comme sur les exemples de la Figure 2.2.

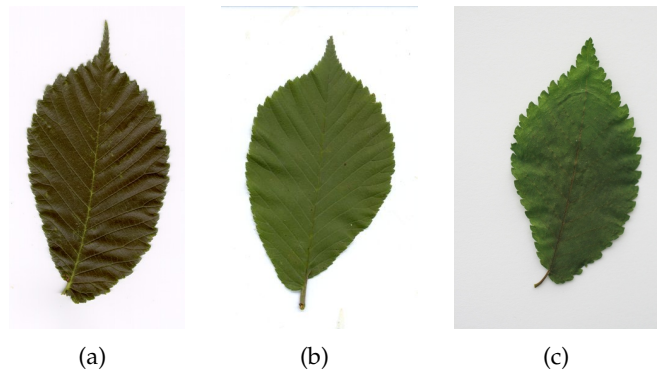


FIGURE 2.2 – Feuilles d'ormes scannées : *Ulmus glabra* Huds. de Swedish Leaves (a), *Ulmus chenmoui* Cheng de ICL (b) et *Ulmus minor* Mill. de Pl@ntLeaves (c)

Dans ce type d'image, on peut considérer que la feuille a déjà été isolée de son environnement et qu'il ne reste plus qu'à affecter les pixels au vu de leur valeur à la feuille ou au fond. Ce fond étant généralement blanc, ou d'une couleur très claire, traiter les images en niveaux de gris peut suffire, et la définition d'une valeur seuil de l'intensité de l'image permet alors de séparer à coup sûr les deux classes, étant donné qu'elles couvrent des plages bien distinctes de luminosité.

Pour cette approche de seuillage, la valeur peut être fixée empiriquement [Söd01, WBX⁺07] mais le plus efficace est de définir un seuil automatiquement pour chaque image et de manière optimale. C'est la méthode d'Otsu [Ots79] qui fait figure de référence pour déterminer ce seuil, et elle est utilisée par la plupart des systèmes traitant des images scannées de feuilles [WCF03, MA04, DWZ07, CA10, YAT11, CBFMB11, MYVB12, AGB⁺12] ou comme point de départ dans le cas d'images naturelles [WHD⁺08].

C'est une méthode de sélection qui se base sur l'histogramme de niveaux de gris de l'image, en mesurant pour chacune des 256 valeurs possibles la variance intra-classe des deux composantes séparées par le seuil. La valeur seuil finale est calculée comme la moyenne des deux meilleurs seuils trouvés lors de ce processus. La Figure 2.3 illustre le fonctionnement de cette méthode.

Dans le contexte d'images scannées, où l'acquisition des images est à ce point maîtrisée, il ne fait aucun doute qu'une approche même aussi simple garantit un taux de réussite très intéressant. D'autres méthodes de séparation des pixels sont malgré tout possibles, et peut-être plus extensibles à des cas de figure plus ardues. Elles utilisent notamment des algorithmes de *clustering* toujours dans la même optique de considérer l'image dans sa globalité et d'associer les pixels individuellement au fond ou à l'objet.

En considérant la distribution des pixels de l'image dans un espace colorimétrique (rarement RGB, on préfère souvent des espaces distinguant luminance et chrominance, comme HSV ou L*a*b*) une méthode de *clustering* va en effet permettre de séparer deux classes. La plus commune est celle des *k-means* qui propose une approche itérative pour séparer un ensemble de vecteurs en *k clusters* en alternant successivement :

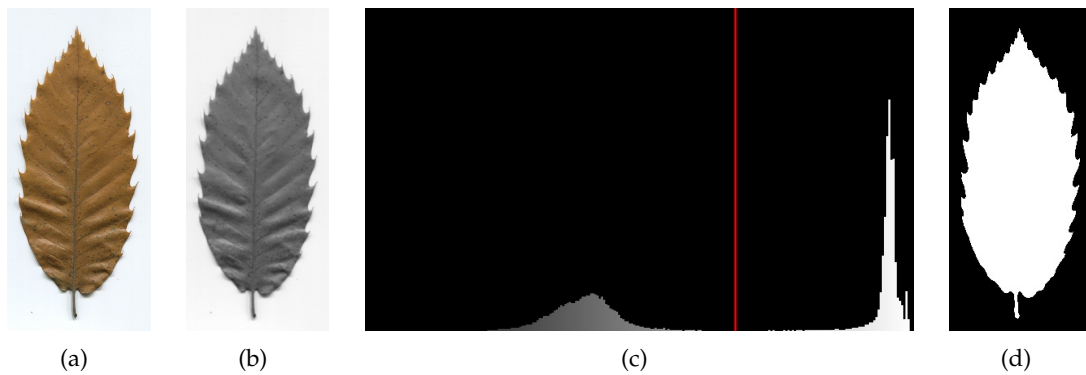


FIGURE 2.3 – Segmentation globale par seuillage en utilisant la méthode d'Otsu : image initiale (a) image en niveaux de gris (b) histogramme et position optimale du seuil (c) et image seuillée en utilisant cette valeur (d)

- l'affectation de chaque vecteur au *cluster* dont le centroïde lui est le plus proche
- la mise à jour des centroïdes comme la moyenne des vecteurs affectés à chaque *cluster*

jusqu'à ce que les affectations soient stables d'une itération sur l'autre. Le résultat est une partition des données dont la variance intra-classe est un minimum local.

En considérant deux *clusters* seulement, on peut donc produire une binarisation de l'image initiale en regroupant les pixels de couleurs proches (indépendamment de leur localisation). C'est une méthode parfois utilisée pour obtenir la segmentation [SK00] ou bien comme point de départ d'une autre méthode [BCF⁺08]. Elle demande simplement une heuristique pour décider lequel des deux *clusters* correspond à l'objet, par exemple choisir celui dont la couleur moyenne est la plus proche de la couleur moyenne du centre de l'image [SK00] en supposant que l'objet soit centré.

Une autre approche de *clustering* est l'algorithme EM (*Expectation-Maximisation*) qui suit lui aussi une approche itérative mais en utilisant un modèle probabiliste. Les distributions des différentes classes sont par exemple modélisées par des distributions gaussiennes, et l'on cherche à trouver les paramètres de ces modèles qui maximisent la vraisemblance des échantillons, en alternant, jusqu'à convergence, le calcul de ces vraisemblances et la mise à jour des modèles probabilistes. En optimisant de la sorte deux modèles gaussiens, initialisés par exemple par le résultat d'un *k-means* [BCF⁺08, KBB⁺12], on obtient une modélisation de la distribution des couleurs de la feuille et de celle du fond, qui va permettre d'affecter chaque pixel à la classe pour laquelle il est le plus vraisemblable.

Ces méthodes ont l'avantage de ne pas faire de suppositions sur la luminosité relative de la feuille et du fond, et permettent surtout d'isoler l'objet dans un contexte où la couleur est plus importante que la luminosité. Comme le montre la Figure 2.4, elles peuvent donc s'appliquer à des images prises dans un contexte moins strict que les feuilles scannées, mais où le choix de distinguer deux classes de couleurs homogènes a un sens. Il sera néanmoins bien plus compliqué d'obtenir une segmentation correcte dès lors que le fond se fera plus complexe.

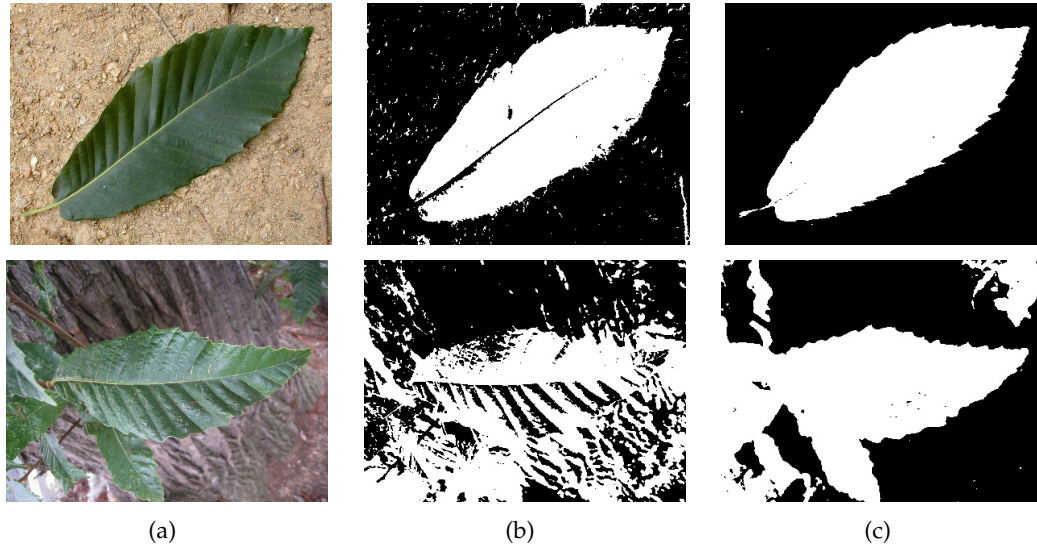


FIGURE 2.4 – Segmentations obtenues par *clustering* : image initiale (a) résultats de l'algorithme *k-means* (b) et résultats de l'algorithme EM (c)

2.2.3 Algorithmes de segmentation plus évolués

Des méthodes de segmentation d'images plus évoluées, ne considérant pas l'image dans sa globalité mais prenant en compte des informations plus locales, vont permettre de traiter les images naturelles de façon plus robuste. En travaillant sur une région ou un contour, ou en s'appuyant sur des données comme le gradient de l'image, il est possible de s'approcher de façon plus exacte de l'objet à extraire de l'image.

Dans la plupart des cas, les systèmes de reconnaissance de végétaux sont destinés à être manipulés par un utilisateur, et une part d'interactivité dans la phase de segmentation est envisageable, d'autant que les performances sur des images naturelles vont s'en trouver inévitablement améliorées. La méthode des *Intelligent Scissors* [MB95] qui permet à partir d'un contour grossier tracé par l'utilisateur d'identifier des contours de coût minimal (un coût basé sur le gradient et le laplacien) peut ainsi être utilisée pour la segmentation de fleurs [SK04]. Il est même possible d'automatiser la phase de définition manuelle du contour pour des images où l'arrière-plan est suffisamment flou pour que les minima de la fonction de coût soient des points du contour de la fleur.

Un autre outil de segmentation interactive et itérative est l'algorithme *GrabCut* [RKB04] qui se base sur les coupes de graphes pour minimiser une fonction d'énergie dépendant à la fois d'une modélisation de la couleur et d'un terme de lissage. Elle requiert simplement le tracé d'un rectangle englobant l'objet pour produire une première segmentation, qui peut être affinée rapidement en marquant manuellement des régions comme appartenant à l'objet ou non. Sa simplicité et sa robustesse en font une méthode qui se prête bien à l'extraction de feuilles dans des photographies complexes [AGB⁺12], comme l'illustre la Figure 2.5, même si elle nécessite parfois de nombreuses interactions répétées pour parvenir à une segmentation parfaite.

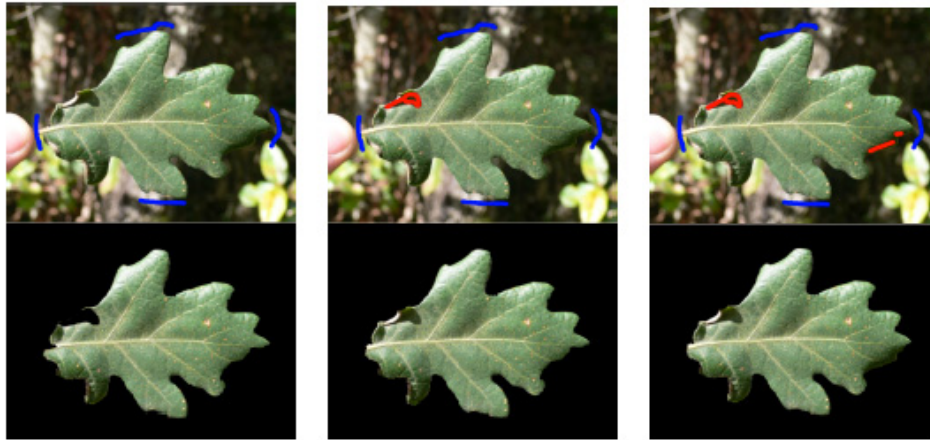


FIGURE 2.5 – Plusieurs itérations d'une segmentation interactive de feuille utilisant l'algorithme *GrabCut* ; images tirées de [AGB⁺12]

Parmi les méthodes de segmentation plus automatiques, l'approche *watershed* [BM92] qui voit le gradient d'un point de vue topographique, et fait le parallèle entre les composantes indépendantes de l'image et des bassins versants, est assez utilisée [WHD⁺08, YAT12]. Le fait que l'information de gradient soit particulièrement bien adaptée pour délimiter des feuilles dont les couleurs peuvent par ailleurs varier grandement (reflets, taches, couleurs d'automne, textures) plaide en tout cas en sa faveur. Cette méthode nécessite toutefois des marqueurs pour initialiser les bassins et éviter une sur-segmentation, régions initiales qui peuvent être obtenues manuellement [YAT12] ou après une première phase de traitement global [WHD⁺08].

La complexité du problème de segmentation des feuilles dans des images naturelles est telle qu'elle peut justifier l'introduction de la 3D pour en faciliter l'extraction [TKC09]. Cela nécessite toutefois une phase de calibration de l'appareil photo pour pouvoir espérer reconstituer un nuage de points à partir de plusieurs images de points de vue proches. L'information de profondeur rend alors plus aisée la tâche de segmentation, mais c'est au détriment d'un coût en termes de calcul relativement élevé.

Enfin, l'utilisation d'un modèle de forme pendant la phase de segmentation est également une solution, introduisant des connaissances externes pour mieux isoler l'objet. Dans le cas des fleurs [NZ07], à partir d'une première segmentation obtenue par coupe de graphe pour minimiser un coût basé sur une modélisation *a priori* de la couleur, l'application d'un modèle de forme de fleur permet de faire une sélection de pixels cohérents et de construire un modèle de couleur plus adéquat, afin d'obtenir une meilleure segmentation par la suite (Figure 2.6). C'est ici la connaissance antérieure de la forme à obtenir qui joue le même rôle que l'utilisateur dans l'algorithme *GrabCut*, permettant d'affiner itérativement la segmentation, bien aidée, il est vrai, par le fait que la couleur est particulièrement discriminante pour les fleurs. L'une des difficultés (en particulier pour le cas des fleurs) est de produire un modèle qui s'adapte correctement à tous les objets considérés, sous peine de se trouver incapable de traiter certains exemples.

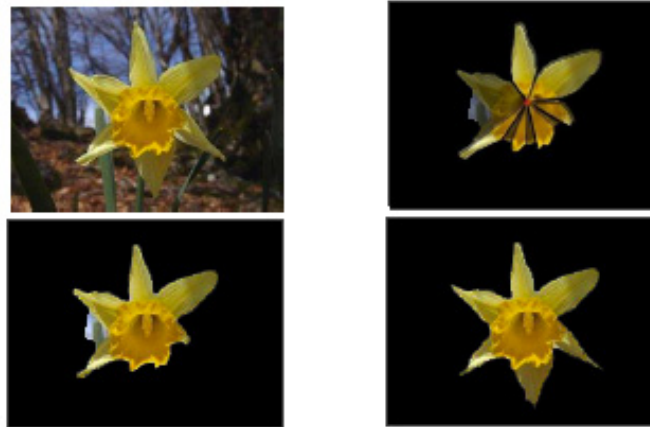


FIGURE 2.6 – Segmentation itérative de fleurs en utilisant un modèle de forme ; images tirées de [NZ07]

Ces différentes méthodes montrent bien la nécessité d'introduire une source d'informations extérieures, que ce soit par le biais d'un utilisateur, de connaissances *a priori*, ou de la prise en compte de la profondeur, pour résoudre un problème des plus complexes. La segmentation purement automatique des feuilles en environnement naturel est en effet un défi considérable, d'autant plus que c'est une phase critique pour toute la suite du traitement.

2.2.4 Détection de points d'intérêt

Dans certains cas, on peut préférer, pour ne pas risquer de compromettre toute la reconnaissance par une étape de d'extraction de l'objet peu robuste, ne considérer que des points comme objet d'intérêt, plutôt qu'une région ou un contour extraits. Il va alors s'agir de repérer dans l'image les points les plus saillants, potentiellement porteurs de plus d'informations, de façon à obtenir un ensemble de points qui soit représentatif du contenu de l'image.

Parmi les détecteurs les plus communément utilisés, l'un des précurseurs est le détecteur de coins de Harris [HS88] qui cherche à localiser les points correspondant à un changement brutal de direction de contour. Cette détection s'effectue sur l'image en niveaux de gris, en examinant localement ses dérivées, mais elle peut aussi être généralisée à des images en couleur. L'extraction de points de Harris à différentes résolutions dans une image de feuille [BYM⁺12] permet d'obtenir un objet à partir duquel on peut espérer capturer les spécificités de la feuille.

Tout aussi répandus, les *Scale Invariant Feature Transform* (SIFT) [Low04] reposent sur un mécanisme de détection similaire basé cette fois sur des différences de Gaussiennes (DoG) multi-échelles. Le résultat est le même, un ensemble de points d'intérêt (*keypoints*) qui, une fois décrit leur contexte local, pourra former une signature de l'image. Ces points peuvent ainsi être détectés sur tout type d'image, et sont une solution pour représenter des images d'organes végétaux très différents [FS11] aux formes

difficilement interprétables comme les aiguilles ou les écorces.

La détection de points d'intérêt fournit ainsi une alternative à l'extraction explicite de l'objet, dans un contexte où cette entreprise peut s'avérer risquée. Les points extraits constituent un objet moins immédiatement compréhensible, mais qui s'avère être une source fiable d'information. Les mécanismes à l'œuvre pour l'extraction des *keypoints* ne sont pas sans rappeler la façon dont l'œil humain se fixe sur des détails saillants, et la robustesse des points entre deux image proches tend à appuyer la pertinence de cette approche en comparaison de l'approche intuitive de segmentation.

En revanche, les objets ainsi produits sont très peu explicites, et il est difficile pour un utilisateur de se faire une idée de ce sur quoi travaille le système. En particulier, dans le cadre de la reconnaissance, la décision qui sera prise en bout de chaîne est beaucoup moins explicable lorsqu'elle est basée sur un ensemble disparate de points que sur une région ou un contour, visuellement beaucoup plus explicite. Le lien qui existe entre l'information extraite de l'image et la suite du processus est donc nettement moins apparent avec ce genre d'approche par points d'intérêt.

2.3 Caractérisation des objets extraits

Une fois identifié dans l'image un objet que l'on considère représentatif de l'espèce végétale, il faut parvenir à convertir cette donnée brute en une représentation (généralement sous forme de valeurs numériques, mais pas nécessairement) qui, tout en capturant le maximum d'informations pertinentes quant à la classe de l'objet, va permettre une comparaison avec des données tirées d'autres images ou de modèles appris. Cette étape de description dépend énormément de la façon dont l'objet a été extrait lors de la phase précédente. Elle offre en tout cas la possibilité d'inclure une connaissance *a priori* de ce qui permet de discriminer les espèces, ou au contraire d'évaluer des méthodes génériques à un cas assez complexe d'objets naturels irréguliers.

2.3.1 Mesures géométriques simples

A la suite d'une segmentation, l'objet obtenu est une région dont on peut chercher à caractériser la forme. Plus précisément, soit I_{seg} l'image binaire résultant de la segmentation, la **région** d'intérêt Ω correspondant à l'objet peut être définie par :

$$\Omega = \{(x; y) \mid I_{seg}(x, y) = 1\}$$

Il est également possible de ne considérer que le **contour** de la région, séquence de pixels se trouvant sur son bord. Ce contour $C(\Omega)$ peut facilement être obtenu à partir de l'image binaire par un algorithme de suivi de contour, par exemple si l'on considère un contour 8-connexe (on note alors $V_4(x; y)$ les 4 pixels formant le voisinage du pixel $(x; y)$) :

$$C(\Omega) = \{(x; y) \mid I_{seg}(x, y) = 1, \exists(x', y') \in V_4(x; y), I_{seg}(x', y') = 0\}$$

D'autres objets utiles vont être l'**enveloppe convexe** de la région $Conv(\Omega)$, soit la plus petite région convexe contenant Ω , ainsi que sa **boîte englobante** ou *bounding box* $BB(\Omega)$ c'est-à-dire le plus petit rectangle qui contienne tout Ω .

L'idée la plus intuitive, et par ailleurs largement employée, est alors de mesurer des grandeurs géométriques sur la forme ainsi définie, de façon à la définir dans sa globalité et à la comparer à d'autres formes. Certaines mesures sont ainsi très communément utilisées pour décrire les régions représentant des feuilles :

- **Aire** de la région [WBX⁺07, CA10] : souvent simplement approximée par le nombre de pixels contenus dans la région, à normaliser éventuellement si la taille des images n'est pas fixe.

$$A(\Omega) = \sum_{(x;y) \in \Omega} 1$$

- **Périmètre** de la région [WBX⁺07] : souvent approximé par le nombre de pixels appartenant au contour de la région, soit la longueur de son bord.

$$Per(\Omega) = \sum_{(x;y) \in C(\Omega)} 1$$

- **Diamètre** de la région [WBX⁺07, CA10] : la longueur maximale entre deux points du contour de la région :

$$Diam(\Omega) = \max_{(x;y), (x';y') \in C(\Omega)} \|(x;y) - (x';y')\|_2$$

- **Format** ou **Aspect Ratio** de la région [SK00, DWZ07, WBX⁺07] : rapport des dimensions de la boîte englobante $BB(\Omega)$, en notant L la plus grande et l la plus petite.

$$AR(\Omega) = \frac{L(BB(\Omega))}{l(BB(\Omega))}$$

- **Rectangularité** de la région [DWZ07, WBX⁺07, GJY⁺11] : mesure à quel point la région est proche de son rectangle englobant, en comparant leurs aires :

$$Rect(\Omega) = \frac{A(\Omega)}{L(BB(\Omega)) \times l(BB(\Omega))}$$

- **Circularité** ou **Compacité** [SK00, MA04, WBX⁺07, GJY⁺11] : compare la région à la forme la plus compacte de même périmètre, à savoir le cercle :

$$Circ(\Omega) = \frac{4\pi \times A(\Omega)}{Per(\Omega)^2}$$

- **Convex Area Ratio** [DWZ07, YAT11, BYM⁺12] : rapport d'aires entre la région et son enveloppe

convexe :

$$CA(\Omega) = \frac{A(\Omega)}{A(\text{Conv}(\Omega))}$$

- **Convex Perimeter Ratio** [DWZ07, YAT12, GJY⁺11] : rapport de périmètres entre la région et son enveloppe convexe :

$$CP(\Omega) = \frac{\text{Per}(\Omega)}{\text{Per}(\text{Conv}(\Omega))}$$

- **Excentricité** [DWZ07, CA10, GJY⁺11] : rapport des longueurs de l'axe d'inertie principal de la région $E_0(\Omega)$ et de l'axe secondaire (perpendiculaire) $E_1(\Omega)$:

$$\text{Exc}(\Omega) = \frac{E_0(\Omega)}{E_1(\Omega)}$$

- **Sphéricité** [DWZ07, GJY⁺11] : rapport de longueurs entre le rayon du cercle englobant de la région r_C et le rayon de son cercle inscrit r_c :

$$\text{Spher}(\Omega) = \frac{r_c}{r_C}$$

Ces grandeurs sont des moyens basiques mais très efficaces pour décrire les formes, et représentent fidèlement certaines des propriétés qui vont permettre de reconnaître les espèces. *L'aspect ratio* va ainsi très bien isoler les feuilles allongées, le *convex perimeter ratio* les feuilles lisses, la sphéricité les feuilles lobées. Ce ne sont toutefois pas des mesures spécifiquement dédiées aux formes foliaires, et elles peuvent aussi bien s'appliquer à tout autre type d'objet.

2.3.2 Descripteurs statistiques de forme

De la même façon, il existe nombre de descripteurs génériques qui ont pu être appliqués aux feuilles, avec une approche plus statistique. En effet l'objet d'intérêt peut être vu non plus comme une forme sur laquelle mesurer des grandeurs, mais plutôt comme un ensemble de pixels dont on va chercher à caractériser la distribution dans l'image et la répartition les uns par rapport aux autres.

Dans cette catégorie de descripteurs se retrouvent particulièrement les **moments**, descripteurs très appréciés pour leurs propriétés d'invariance [Söd01, DWZ07, WHD⁺08, KPP10]. La notion mathématique de moment relève de la statistique et des probabilités, où l'on parle de moment d'ordre p d'une variable aléatoire X pour désigner l'espérance $E[X^p]$. Cette notion peut tout à fait être transposée pour les images numériques [Hu62], les coordonnées des pixels formant deux variables aléatoires. Il faut donc deux valeurs p et q pour définir l'ordre des moments et on parle alors de moments d'ordre $p + q$. La formulation générale des moments ordinaires d'une image I sera donc :

$$m_{pq}(I) = \sum_x \sum_y x^p y^q I(x, y)$$

Après segmentation et donc dans le cas d'une image binaire à valeurs 0 ou 1, seuls les pixels de la région seront alors pris en compte dans la sommation, et on peut par conséquent écrire le moment d'ordre $p + q$ d'une région Ω :

$$m_{pq}(\Omega) = \sum_{(x;y) \in \Omega} x^p y^q$$

Cette définition statistique est très générale et permet en fait de couvrir un certain nombre de notions plus intuitives, comme l'aire $Area(\Omega) = m_{00}(\Omega)$, ou le barycentre de la région $(\bar{x}; \bar{y}) = (\frac{m_{10}(\Omega)}{m_{00}(\Omega)}, \frac{m_{01}(\Omega)}{m_{00}(\Omega)})$. On peut ainsi définir les **moments centrés** qui permettent de s'affranchir de la position exacte des pixels dans l'image et ne considérer que leur position relative en x et en y :

$$\mu_{pq}(\Omega) = \sum_{(x;y) \in \Omega} (x - \bar{x})^p (y - \bar{y})^q$$

Ces moments ne sont pertinents qu'à partir de l'ordre 2, puisqu'évidemment nuls à l'ordre 1, et égaux à l'aire à l'ordre 0. Les moments centrés d'ordre 2 correspondent aux éléments de la matrice de covariance Σ de l'ensemble des pixels de la région, qui peut alors s'écrire :

$$\Sigma(\Omega) = \begin{bmatrix} \mu_{20}(\Omega) & \mu_{11}(\Omega) \\ \mu_{11}(\Omega) & \mu_{02}(\Omega) \end{bmatrix}$$

Cette mesure de variance estime la distribution des pixels dans la région et ses axes principaux, et permet notamment de recalculer l'excentricité comme le rapport des valeurs propres de la matrice Σ [WCFW00, MA04]. Les moments d'ordre 3 mesurent quand à eux les dissymétries dans la distribution. De façon à les rendre de plus invariants à l'échelle, il est possible de les normaliser :

$$\eta_{pq}(\Omega) = \frac{\mu_{pq}(\Omega)}{\mu_{00}(\Omega)^{\frac{p+q}{2}+1}}$$

L'ensemble de ces moments centrés normalisés $(\eta_{2,0}, \eta_{1,1}, \eta_{0,2}, \eta_{3,0}, \eta_{2,1}, \eta_{1,2}, \eta_{0,3})$ est généralement utilisé pour calculer des valeurs invariantes à la rotation par combinaison non-linéaire [Hu62]. Ce

sont les **moments invariants de Hu**, ou moments géométriques, définis de la façon suivante

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) ((\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2) \\
&\quad + (\eta_{03} - 3\eta_{21})(\eta_{03} + \eta_{21}) ((\eta_{03} + \eta_{21})^2 - 3(\eta_{30} + \eta_{12})^2) \\
\phi_6 &= (\eta_{20} - \eta_{02}) ((\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2) + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}) \\
\phi_7 &= (\eta_{30} - 3\eta_{12})(\eta_{03} + \eta_{21}) ((\eta_{03} + \eta_{21})^2 - 3(\eta_{30} + \eta_{12})^2) \\
&\quad - (\eta_{03} - 3\eta_{21})(\eta_{30} + \eta_{12}) ((\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2)
\end{aligned}$$

L'avantage de ces descripteurs est de résumer en un nombre réduit de valeurs des propriétés géométriques de la forme (élongation, asymétrie) uniquement via des considérations statistiques, et d'assurer une invariance de ces mesures à la fois à la translation, au changement d'échelle et à la rotation. Leur robustesse en fait des outils très intéressants pour la recherche de formes similaires.

Dans la catégorie des moments viennent aussi se rajouter les moments améliorés de Chen [Che93] qui ne considèrent plus que le contour comme région d'intérêt, et où les mêmes calculs peuvent être faits à partir d'une formulation différente des moments ordinaires [DWZ07, WHD⁺08] :

$$\tilde{m}_{pq}(\Omega) = \sum_{(x;y) \in C(\Omega)} x^p y^q$$

On peut aussi citer les **moments orthogonaux de Zernike** [Tea80], basés sur les polynômes de Zernike [Zer34] qui forment un ensemble de polynômes orthogonaux sur le cercle unité. Une fois ramenée l'image au cercle unité la sommation des valeurs des 16 premiers polynômes sur la région permet de définir 16 moments dotés des mêmes propriétés d'invariance que les moments de Hu, et donc tout aussi robustes pour décrire les formes [WHD⁺08].

Dans un registre un peu différent, certaines méthodes vont chercher à retranscrire le profil de la forme sans véritablement chercher à l'interpréter, généralement en se basant sur son contour. Le contour va alors souvent être décrit par son abscisse curviligne. A tout point $(x; y)$ du contour $C(\Omega)$ correspond ainsi une valeur d'abscisse u de telle sorte que la courbe puisse être paramétrée par la variable u :

$$C(\Omega) = \{(x(u); y(u)), u \in [1; N]\}$$

L'exemple le plus simple est la courbe de distance centre-contour (*Centroid-Contour Distance* ou **CCD**) [WCFW00, WCF03, TKC09, KPP10] qui va simplement tracer la distance de points consécutifs

du contour au barycentre $(\bar{x}; \bar{y})$ de la région, en fonction de leur abscisse curviligne u . Dans un souci de comparabilité, les contours vont alors souvent être ramenés à la même longueur N par sous-échantillonnage :

$$\forall u \in [1; N], CCD(u) = \| (x(u); y(u)) - (\bar{x}; \bar{y}) \|_2$$

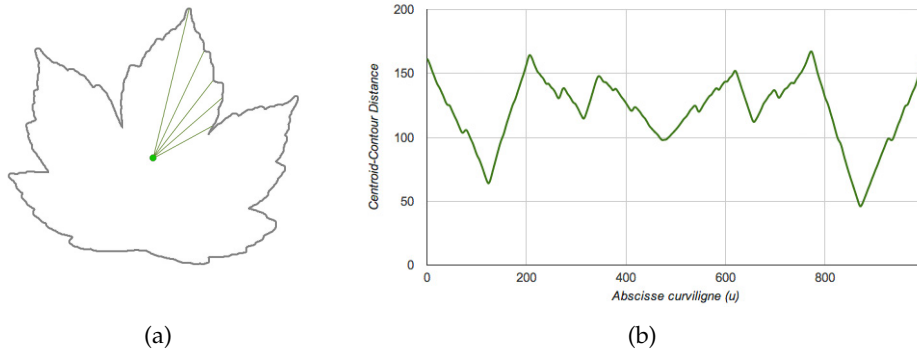


FIGURE 2.7 – Calcul de la courbe distance centre-contour (CCD) : contour original et centre de gravité (a) et courbe résultante sur une échelle fixe (b)

Cette description très brute, illustrée dans la Figure 2.7, est ainsi invariante à la translation, et au changement d'échelle à condition d'opérer une normalisation sur la distance. En revanche, il est plus compliqué de parvenir à une invariance à la rotation. La seule solution possible est en effet de pouvoir considérer tous les points du contour comme point d'abscisse 1, avec à chaque fois un décalage de la courbe. Cela implique donc que la comparaison entre deux de ces courbes requiert en fait N comparaisons pour trouver le décalage qui minimise l'écart entre les deux courbes, ce qui ne se fait qu'au prix d'un nombre d'opérations forcément très élevé.

Des statistiques basiques peuvent être dérivées de cette courbe (distance moyenne, minimale, maximale, variance) [YAT11] de façon à obtenir une description plus robuste, mais nettement moins complète. Comme tout processus statistique agrégatif, l'effet pervers est de perdre les relations spatiales entre les points au profit d'une mesure unique.

D'autres descripteurs génériques de forme sont les *Shape Contexts* [BMP01] qui cherchent à représenter les relations spatiales de différents points du contour. L'idée est de produire en chaque point d'un contour sous-échantillonné un histogramme décrivant la distribution des autres points par rapport au point considéré. L'histogramme accumule les points en fonction à la fois de leur distance et de leur direction, dans des *bins* uniformes dans l'espace log-polaire, comme ceux présentés sur la Figure 2.8.

Pour adapter cette description aux formes plus complexes, pouvant notamment présenter des concavités assez marquées ou des parties presque indépendantes comme les feuilles, il est possible de considérer la distance interne entre les points, à savoir la longueur du plus court chemin entre deux points du contour passant uniquement par des points de Ω . Ces *Inner-Distance Shape Contexts* (IDSC)

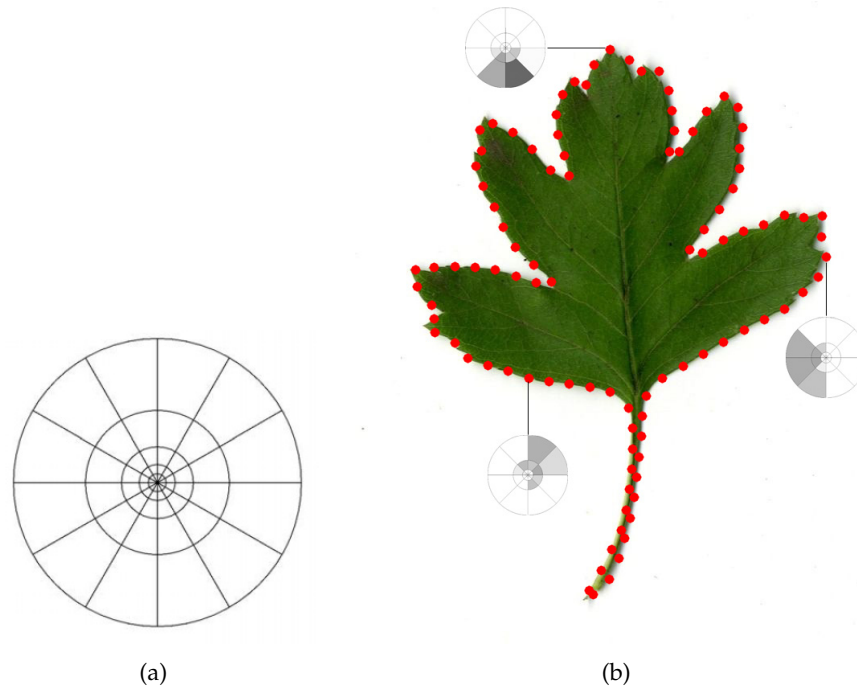


FIGURE 2.8 – Shape Contexts : Histogramme en coordonnées log-polaires utilisé pour accumuler les points (a) et descripteurs obtenus en différents points d'une forme (b)

[LJ07, BCF⁺08] permettent ainsi de représenter plus fidèlement des formes potentiellement flexibles et d'apporter de l'information sur la morphologie des objets.

Une autre extension [MYVB12, BYM⁺12] consiste à ne considérer dans le calcul non plus uniquement des points du contour mais également des points d'intérêt préalablement extraits, de façon à prendre en compte des informations locales en plus de la forme.

Une approche connexe est de considérer les points du contour comme les sommets d'un graphe, dont les arêtes seront pondérées par les distances euclidiennes entre les points, un objet qui conserve également les propriétés spatiales du contour. Après élimination des arêtes au dessous d'un certain nombre de seuils de distance, on obtient plusieurs réseaux complexes (*Complex Network*) [BCMB09, CBFMB11, AGB⁺12] qu'il est possible de caractériser en termes de degré moyen ou degré maximal.

Enfin, il est possible d'analyser le contour de façon statistique en l'assimilant à un signal. Par exemple, si l'on considère la séquence de codes de Freeman (qui définissent la direction du segment formé par deux points consécutifs du contour parmi les 8 possibles) on obtient un signal qui se prête parfaitement à une décomposition en **coefficients de Fourier** [CNMJS06, YAT11]. Cette description fréquentielle peut, de plus, être rendue invariante à la rotation si l'on ne prend pas en compte l'information de phase (la rotation correspondant en effet à une translation du signal).

2.3.3 Caractérisation géométrique de contour

Le contour des objets peut aussi être vu de façon moins analytique en considérant un peu plus sa géométrie. C'est notamment le cas dès lors qu'on s'intéresse à la courbure des contours, une mesure qui permet de capturer les changements de direction de la courbe et donc les variations locales de la forme. Le calcul de la courbure peut s'effectuer de différentes manières, mais sa formulation mathématique pour une courbe paramétrée par l'abscisse curviligne u est la suivante :

$$\kappa(u) = \frac{x'(u) \cdot y''(u) - x''(u) \cdot y'(u)}{(x'(u)^2 + y'(u)^2)^{\frac{3}{2}}}$$

Un outil particulièrement utile est la transformation dans l'espace courbure-échelle (*Curvature-Scale Space* ou **CSS**) [MM92] qui consiste à empiler des mesures de courbures $\kappa(u)$ extraites sur des contours progressivement lissés par des filtres gaussiens d'échelle s croissante, ramenés à un nombre de points constant. On notera $\kappa(u, s)$ la valeur de courbure correspondant au contour lissé à l'échelle s .

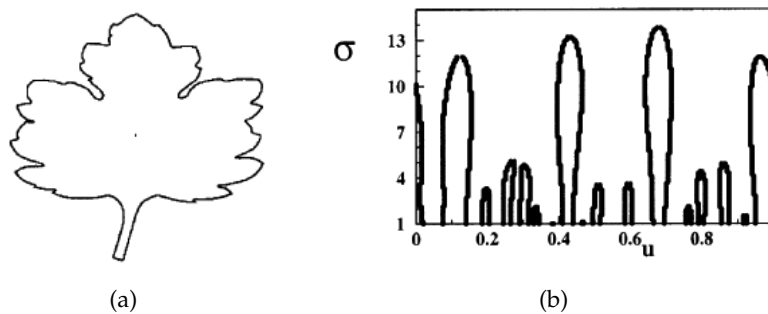


FIGURE 2.9 – Curvature Scale Space : contour initial (a) et contours (*zero-crossings*) extraits de la transformation dans l'espace courbure-échelle (b) ; images tirées de [MA04]

Le résultat de cette transformation, illustré en Figure 2.9, est une matrice particulièrement visuelle où les convexités et les concavités présentes sur un contour se traduisent en stries de courbure positive ou négative. Pour décrire les formes, les **CSS-contours**, à savoir les points de changement de signe de la courbure aux différentes échelles, sont particulièrement utilisés [AMK97, MA04], en considérant le vecteur formé par les points terminaux de ces contours comme représentatifs de la forme. En considérant les extrema de courbure au lieu des passages à zéro, on obtient également un ensemble de points caractéristiques de la forme [CA10].

Il est aussi possible d'appliquer une vision plus globale à cet objet, en se servant du CSS pour générer un histogramme des courbures en fonction de l'échelle (**HoCS**) [KBB⁺12]. En sommant pour un certain nombre de niveaux d'échelles les occurrences de valeurs de courbures, on obtient une description assez compacte du contour qui va à la fois résumer sa géométrie locale et les propriétés plus globales de sa forme, mais en perdant une fois encore la dimension spatiale.

Pour caractériser le contour, les changements locaux d'orientation constituent une source considérable d'information, permettant par exemple de différencier les bords lisses pour lesquelles l'orien-

tation ne varie que très peu, des bords dentés aux changements plus marqués. En échantillonnant le contour, on obtient une séquence de triplets de points qui définissent chacun un angle. Une fois convertie la mesure de ces angles en code, on peut créer un histogramme (*Angle Code Histogram* ou **ACH**) [WCF03, YAT12] en accumulant les codes sur la longueur du contour.

De façon plus explicite encore, on peut chercher à décomposer un contour en segments de direction relativement constante. Cette décomposition en fragments directionnels permet alors de composer un descripteur type histogramme, en comptant les segments en fonction de leur orientation et de leur longueur [GJY⁺11, BYM⁺12].

Ces descripteurs viennent considérer le contour des objets d'un point de vue plus local, et parviennent ainsi à capturer des différences subtiles inaccessibles à des mesures faites sur l'ensemble de la forme. Ils restent malgré tout des descripteurs génériques pouvant être adoptés pour bon nombre d'objets pour lesquels les particularités géométriques du bord jouent un rôle important dans la reconnaissance.

2.3.4 Modélisation explicite d'organes végétaux

Dans un contexte comme la reconnaissance de végétaux, où les critères discriminants pour identifier les espèces sont par avance connus, cela a du sens de chercher à évaluer explicitement certaines de ces propriétés. Le fait de savoir à l'avance ce que l'on recherche dans l'image permet de définir des descripteurs dédiés spécifiquement à l'objet d'étude, supposément plus efficaces pour la reconnaissance que des mesures génériques.

Cela suppose généralement d'être capable d'interpréter l'objet extrait de l'image, en détectant par exemple la base et l'apex d'une feuille, ce qui implique soit une intervention manuelle [WBX⁺07], une contrainte sur l'acquisition d'images [SK00] ou un processus plus complexe de localisation utilisant la courbure [INK98, MYB12] ou le squelette de la forme [WCF03].

Ainsi, il est possible de rajouter à des mesures géométriques plus conventionnelles des valeurs explicites représentant l'angle formé par la feuille au niveau de sa base ou de son apex [SK00], la longueur de la nervure principale [WBX⁺07], le nombre de lobes de la feuille ou de pétales de la fleur [INK98, SK00, SK04]. La nervation est aussi un objet intéressant à caractériser [WBX⁺07], mais la difficulté réside dans son extraction de manière précise qui, si elle est possible pour des images scannées [LC06], demeure un vrai problème dans le cas général.

C'est surtout la dentition de la feuille qui va pouvoir être modélisée de façon précise. Les dents, comme maxima de courbure, forment en effet des points particulièrement remarquables qui vont pouvoir être détectés et localisés comme sur la Figure 2.10. Pour caractériser la forme du bord, on peut alors simplement compter ses dents ainsi détectées, éventuellement avec différents seuils de détection [INK98, AGB⁺12] ou même chercher à caractériser la forme de chaque dent pour en faire des statistiques à l'échelle du contour [INK98].

De tels descripteurs exclusivement dédiés à la représentation des organes végétaux s'avèrent très



FIGURE 2.10 – Détection des dents sur le contour de la feuille avec différentes valeurs de seuil (méthode tirée de [AGB⁺12])

riches d'un point de vue sémantique, et permettent de capturer uniquement l'information pertinente pour l'identification, à la condition près que l'interprétation sur laquelle ils reposent soit exacte. Leur utilisation correspond à l'introduction d'une connaissance *a priori* qui peut se montrer intéressante mais qui exige une cohérence avec le reste du processus pour être efficace, là où des descripteurs génériques de forme se contenteront de retranscrire l'information brute de la forme extraite, délaissant l'interprétation aux étapes suivantes.

2.3.5 Représentations de couleurs et textures

En plus de la forme de l'objet considéré, une quantité non négligeable d'informations peut être tirée de l'intérieur de la région segmentée, essentiellement via la couleur (notamment si l'on étudie les fleurs) et la texture, à savoir tous les motifs que l'on rencontre à la surface de l'objet.

En ce qui concerne la couleur, chaque pixel de la région peut être décrit par un vecteur tridimensionnel, dont la valeur va dépendre de l'espace colorimétrique choisi. L'espace **RGB** est le plus commun et celui dans lequel est représentée l'image brute, chaque pixel étant ainsi localisé par sa valeur sur les canaux rouge, vert et bleu. On peut ainsi considérer la région Ω comme une distribution de points dans l'espace colorimétrique, par exemple en RGB :

$$Col(\Omega) = \{(R(x, y); G(x, y); B(x, y)) \mid (x, y) \in \Omega\}$$

La définition serait la même si l'on adoptait d'autres espaces comme l'espace Teinte-Saturation-Luminosité (**HSL**) ou l'espace perceptuel **L*a*b*** qui ont tous deux l'avantage de détacher l'information de luminance de l'information de chrominance et pouvoir ainsi s'affranchir quelque peu des variations dues à l'éclairage.

Conserver la distribution complète des couleurs n'offre pas une description efficace, et il faut passer par un processus statistique pour condenser cette information. Une solution est de passer par une quantification de l'espace des couleurs, ou d'un sous espace, pour construire un histogramme représentant la distribution. L'histogramme complet peut alors constituer un descripteur, ou alors simplement le ou les deux *bins* les plus représentés [SK00].

Considérer le nuage de points dans un espace 3-D permet aussi d'utiliser des méthodes statistiques pour caractériser sa forme. Ainsi en généralisant les moments au cas tri-dimensionnel, on obtient une description rendant compte de la position, de la déviation et de l'orientation de la distribution des couleurs [YAT11], qui peut de plus être rendue moins dépendante à l'éclairage, même avec l'espace RGB, si l'on considère des moments centrés.

La texture s'évalue également à partir de l'intérieur de la région, mais elle se base généralement sur l'image en niveaux de gris uniquement. Les motifs à la surface des feuilles (dûs à des variations de couleur ou à la nervation), ceux visibles sur des images de fleurs ou encore les stries et craquelures des écorces vont pouvoir être détectés en utilisant des outils classiques comme les ondelettes de **Gabor**, les *Local Binary Patterns (LBP)* ou encore les champs aléatoires de Markov (**MRF**).

La description de la texture passe le plus souvent par une phase de filtrage de l'image en niveaux de gris, obtenu suite à sa convolution par un jeu de filtres différents. Par exemple, les filtres de Gabor sont généralement utilisés avec plusieurs tailles et différentes orientations (Figure 2.11), et ce sont les différentes réponses qui sont employées comme sources d'information par exemple en considérant leurs valeurs moyennes et leurs variances [HZD⁺06, KYÜ11], ou des matrices de co-occurrences [CRBW10].

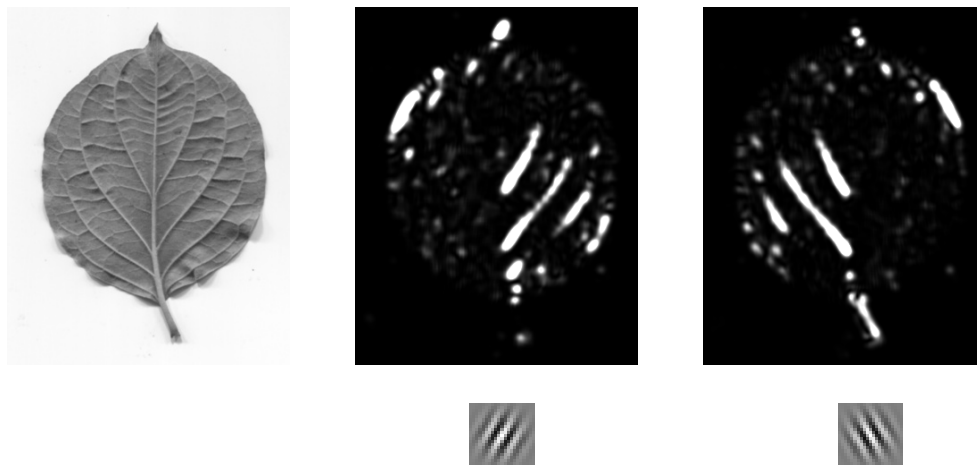


FIGURE 2.11 – Résultat du filtrage par des ondelettes de Gabor d'orientations différentes

D'autres approches passent par un filtrage par des filtres différents [NZ06], des fonctions dédiées à la détection des nervures [LC06], ou bien par le calcul d'orientation dominante de chaque pixel [YAT12], ou de covariance [YAT11] souvent dans le but de produire des histogrammes (*Circular Covariance Histogram, CCH*).

2.3.6 Descriptions de contextes locaux

Lorsque l'objet extrait de l'image n'est pas une région de pixels dont on peut décrire la forme et le contenu, mais un ensemble de points saillants, les méthodes de description ne peuvent fatalement

pas être les mêmes. La raison d'être initiale de ces points d'intérêt caractéristiques est de pouvoir être reconnus et suivis dans des vues différentes d'objets identiques ou du moins similaires. Les points détectés sont donc associés à une description de la région locale de l'image qui va permettre de retrouver des points similaires d'une image sur l'autre.

Ainsi, les **SIFT** reposent sur un histogramme des orientations du gradient (**HOG**) de l'image, calculées relativement à l'orientation du point d'intérêt pour assurer l'invariance à la rotation, établis sur une grille 4x4 de fenêtres de même taille [Low04]. Ces histogrammes une fois normalisés sont concaténés pour former la description du point SIFT sous forme d'un vecteur de dimension 128.

De manière analogue, les *Speeded-up Robust Features* (**SURF**) se basent sur les réponses au filtrage par des ondelettes de Haar sur le même type de grille [BTVG06], résultant en un descripteur plus compact de dimension 64, et également plus rapide à calculer.

Il existe de nombreuses méthodes concurrentes de description de contexte local, reprenant toutes la même idée d'une fenêtre autour du point d'intérêt dont la taille et l'orientation sont déterminées lors de la détection et qui sert de support à l'évaluation de mesures décrivant la forme et la texture locale (HOG, Fourier, transformée de Hough, ...). Elles donnent également toutes un vecteur de taille fixe qui va permettre d'effectuer un *matching*, c'est-à-dire de trouver des correspondances entre les points d'une image à ceux extraits d'une autre.

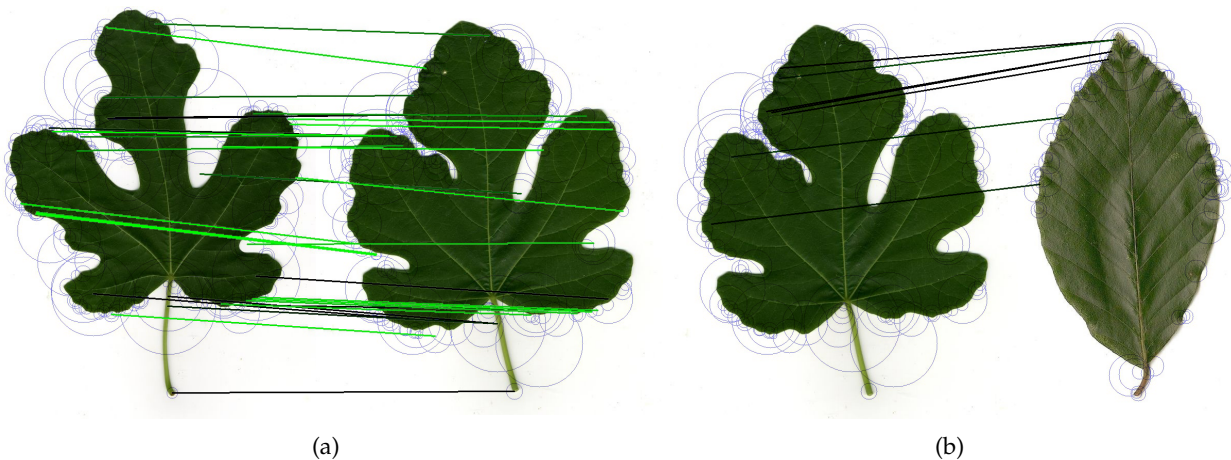


FIGURE 2.12 – Matching rigide de descripteurs de points d'intérêt (SURF) entre deux feuilles de la même espèce (a) et de deux espèces différentes (b)

Ce *matching* permet de retrouver des points similaires sans aucune contrainte de spatialité. Aussi peut-il être intéressant, dans le cadre de la description d'objets, d'ajouter une contrainte de rigidité (comme sur la Figure 2.12) à ces correspondances de façon à ne conserver que les points qui correspondent en plus à la transformation globale entre les deux images (elle-même évaluée à partir des correspondances de points par exemple via un algorithme type RANSAC) [GJY⁺11, BYM⁺12]. Le nombre de correspondances donne alors une mesure de la similarité entre deux images.

Les méthodes de description d'images basées sur les points d'intérêt permettent ainsi de capturer des similarités visuelles entre les images, en s'attachant sur des détails caractéristiques, un peu à la

manière de l'œil humain. A condition d'y inclure un peu de spatialité, les objets complets peuvent être représentés et leur similarité mesurée d'une image sur l'autre. Telles quelles, elles rendent en revanche difficile une généralisation au niveau de l'espèce et contraignent à effectuer des comparaisons image par image.

2.3.7 Histogrammes de vocabulaire appris

Une approche de plus en plus commune en reconnaissance visuelle consiste agréger un nombre important d'informations tirées de l'image dans un sac de mots (*Bag-of-Words* ou **BoW**). Ce modèle, introduit à la base pour le traitement du langage naturel, repose sur l'idée que toute information peut être décrite à l'aide d'un vocabulaire tiré d'un dictionnaire, et qu'en comptant les occurrences pour un objet, de chacun des mots du dictionnaire, on en obtient une description assez complète.

Ce modèle s'adapte tout particulièrement aux points d'intérêt extraits de l'image, qui constituent d'ailleurs leur première application dans le cadre du traitement d'images [SZ03, CDF⁺04]. La description du contexte local autour d'un point permet de représenter l'élément visuel contenu dans la région de l'image autour du point. L'idée est alors de décrire ces morceaux d'images par ce que l'on va nommer des **mots visuels**, en apprenant automatiquement à partir des descriptions extraites le dictionnaire des mots pertinents dans le contexte étudié.

Cette phase d'apprentissage en fait des descripteurs relativement à part, et surtout totalement déconnectés d'un quelconque *a priori* quant à ce qui pourrait être pertinent pour un objet particulier. Une fois extraites les données d'un nombre important d'images (par exemple les descriptions de quelques milliers de points SIFT) c'est une méthode de clustering, avec un nombre de classes fixé (la taille du dictionnaire), qui va permettre de faire émerger les mots du vocabulaire. Chacun des éléments des données initiales peut alors être associé à un cluster, c'est-à-dire à un mot du dictionnaire, et on représente une image par l'histogramme des occurrences de chacun des mots, dont la taille fixe est celle du dictionnaire.

Cette méthode permet de décrire de façon assez robuste des images végétales complexes, sans se préoccuper d'une segmentation risquée, et en tirant parti de descripteurs de contextes locaux comme les SIFT [FS11] aussi bien pour des images de feuilles que d'écorces ou d'aiguilles, ou encore les SURF [AGB⁺12] pour des images complexes de feuilles composées difficiles à interpréter. Ces descripteurs locaux peuvent aussi être calculés uniquement dans une région d'intérêt, sur une grille régulière après segmentation [NZ06], représentant ainsi des informations sans doute plus pertinentes.

Cependant, il est possible d'utiliser la même approche sur d'autres données que des descripteurs locaux d'images. Par exemple, en considérant la couleur des pixels des régions segmentées d'un grand nombre d'images, et en appliquant la même procédure d'apprentissage de vocabulaire et de conversion en histogramme on obtient une décomposition de l'espace des couleurs non plus arbitraire comme pourrait l'être une quantification mais directement liée aux données des images [NZ06]. La même stratégie peut être employée sur des descripteurs de texture en chaque point, avec

la constitution d'un dictionnaire de *textons* [NZ06].

En dépit de leur côté "boîte noire", ces descripteurs présentent l'avantage d'être à la fois parfaitement génériques et d'assurer une cohérence avec les données qu'ils doivent représenter via cette phase d'apprentissage. Ils permettent tout de même assez peu de généralisation, ou demandent en tout cas une base d'apprentissage vraiment conséquente pour que la description produite soit pertinente.

2.4 Apprentissage automatique et recherche d'images

La dernière phase du processus de reconnaissance consiste à prendre une décision quant à la classe de l'objet analysé, à savoir, dans le cas de l'identification de végétaux, son espèce. Pour cela, l'objectif est de mettre en rapport les attributs extraits comme description de l'objet avec un ensemble de référence, une base d'exemples dont la classe est connue, eux-mêmes représentés par le même ensemble de descripteurs.

2.4.1 Comparaison et recherche d'images par le contenu

Dans de nombreux cas, la base d'exemples est utilisée telle quelle pour la reconnaissance d'un nouvel individu. Les descripteurs extraits sur celui-ci doivent alors être comparés individuellement avec ceux de chacun des exemples de référence afin d'établir une mesure de proximité entre l'individu, dont la classe est inconnue, et l'ensemble de la base de données étiquetées. Cette méthode de plus proches voisins est communément utilisée [WCF03, BCF⁺08, CA10, MYVB12, KBB⁺12] et permet d'associer à l'individu l'espèce de l'exemple le plus similaire, ou de classer les espèces en fonction du nombre d'exemples de chacune, parmi les exemples les plus similaires.

Le calcul de la mesure de proximité au cœur de ce processus passe par la définition d'une métrique qui est généralement spécifique au descripteur utilisé. La façon dont celle-ci est définie va avoir une influence non-négligeable sur la complexité de la recherche. En effet certains descripteurs spatialisés qui nécessitent une forme d'alignement sont souvent associés à une mesure de distance qui se calcule en temps quadratique [BCF⁺08], et rend les comparaisons extrêmement coûteuses.

La recherche est généralement effectuée de manière extensive, et génère donc un volume de calcul relativement important. Toutefois, les temps de calcul peuvent être réduits en utilisant par exemple des méthodes de *hashing* [MYVB12] qui réduisent le nombre de comparaisons à effectuer. Une autre possibilité pour limiter le volume de calculs est d'évaluer hors-ligne la distance de tous les exemples de la base entre eux de façon à savoir rapidement quels exemples sont susceptibles d'être proche d'un nouvel individu [BCF⁺08].

Le principal obstacle à l'usage de ce genre de méthodes dans notre contexte est qu'elles ne produisent pas de généralisation des données de référence, et nécessitent donc l'ensemble des exemples dont la classe est connue pour fonctionner. Le volume de données que cela représente peut être très important, ce qui rend pratiquement impossible leur embarquement sur un dispositif léger comme un smartphone. Nous préférons donc nous pencher plutôt sur des approches capables, par un

apprentissage, de condenser l'information présente dans la base d'exemples, en même temps que d'en effectuer une généralisation à ce qui permet véritablement de définir les classes qui nous intéressent.

2.4.2 Méthodes de classification automatique

L'apprentissage automatique, qui consiste de façon générale à entraîner un système prenant en entrée un ensemble d'attributs et donnant en sortie une information de classe, est un domaine bien trop large pour que nous en fassions un inventaire détaillé. De nombreux modèles développés pour la classification de données, représentées le plus souvent sous la forme de vecteurs, existent avec chacun leurs avantages, et nous nous limiterons à citer les plus usuels qui n'ont pas manqué d'être utilisés dans le cadre de la reconnaissance de végétaux.

L'une des méthodes les plus répandues est celle des réseaux de neurones [SK00, WBX⁺07] qui est dans sa conception originelle supposée simuler le comportement du cerveau humain. La définition de ces réseaux peut également inclure une part probabiliste. L'entraînement d'un tel système permet d'ajuster les poids et fonctions de transfert qui régissent son comportement, sans qu'il soit véritablement possible d'appréhender de manière intuitive ce qui est modélisé.

Un autre outil populaire est l'utilisation des *support vector machines* ou **SVM** [NZ08, FS11, YAT11, YAT12], qui, bien que développés initialement pour des problèmes à deux classes, se généralisent bien aux problèmes plus complexes. C'est surtout la possibilité d'utiliser les noyaux, et donc de discriminer des ensembles non-linéairement séparables, qui rend ces méthodes attractives dans un cadre où la distribution des classes peut s'avérer particulièrement complexe.

Enfin, récemment, les forêts aléatoires ont été appliquées avec succès au cas de l'identification des végétaux [AGB⁺12]. Ce sont des objets assez simples, puisqu'ils se basent sur des arbres binaires de décision, générés en nombre important et de façon aléatoire. Chaque arbre n'atteint alors qu'une performance moyenne mais pris dans leur ensemble, et en incluant aussi la prise en compte de la distribution des données dans la phase d'apprentissage, ils permettent de bien s'adapter à des données complexes et potentiellement hétérogènes.

Le choix des attributs est dans tous les cas très important en apprentissage automatique et la performance dépend énormément de la pertinence des descripteurs sélectionnés pour représenter les objets d'intérêt. Des techniques de sélection d'attributs [YAT12] ou de réduction de la dimensionnalité comme l'analyse en composantes principales [WBX⁺07] sont souvent employés pour fournir au classifieur les données les plus pertinentes possibles en entrée. Il n'en demeure pas moins que la qualité des descripteurs extraits est primordiale pour la performance de la classification, et que le choix d'un modèle de classifieur va bien souvent dépendre essentiellement des données sur lesquelles il devra travailler.

2.5 Cadre de travail et parti pris scientifique

Le problème abordé dans le cadre de cette thèse peut se résumer de la façon suivante : l'**identification des espèces** d'arbres sur **smartphone** à partir de photographies **naturelles** peu supervisées de leurs **feuilles** en suivant un processus analytique **inspiré par la botanique**. La nature du problème d'identification des espèces, le traitement sur smartphone ou encore la qualité et la teneur des images posent un certain nombre de points incontournables qui vont permettre de délimiter un peu plus précisément la question traitée, et de faire certains choix quant aux méthodes à employer tout au long du traitement :

- ▶ Les images à traiter proviendront d'appareils photos de smartphones, et leur qualité ne sera pas toujours optimale, à cause notamment des problèmes de netteté dus à la mise au point et au mouvement.
- ▶ Pour des raisons pratiques évidentes, on ne peut demander à un utilisateur sur le terrain de photographier les feuilles qu'il rencontre sur un fond uni, le cadre le plus général est donc celui d'images sur fond naturel, éventuellement végétal et complexe.
- ▶ L'analyse botanique se base avant tout sur des critères de forme qui nécessitent évidemment l'extraction d'une région explicite à caractériser : il faudra donc passer par une phase de segmentation de l'image.
- ▶ L'algorithme de segmentation à utiliser devra être suffisamment léger pour être implémenté sur smartphone et pourra tirer parti d'une forme de connaissance *a priori* sur les formes pouvant être prises par les feuilles d'arbres.
- ▶ Les définitions des formes utilisées dans les ouvrages botaniques fournissent une grille d'analyse et une idée des éléments discriminants pour l'identification des espèces qu'il sera intéressant de considérer comme base de la description des formes.
- ▶ Dans une visée pédagogique et éducative, le but sera non seulement de reconnaître les espèces mais aussi de donner à l'utilisateur des éléments de compréhension de la décision prise au final, de façon à transmettre une compétence d'identification.
- ▶ Pour que la reconnaissance puisse s'effectuer entièrement sur le téléphone, la phase de comparaison doit être aussi peu coûteuse que possible tant en temps de calcul qu'en espace nécessaire : le matching sur une grande base de données est donc à exclure.
- ▶ Les problèmes liés à l'identification des espèces (variabilité intra-classe, similarité inter-classe, importance non-uniforme des différents critères) devront absolument être pris en compte par l'algorithme de classification.

Ces contraintes situent donc à l'avance la méthode envisagée parmi les approches existantes (Table 2.2), par l'utilisation d'un algorithme de segmentation suffisamment robuste pour traiter des images



Méthode	Objet	Extraction	Description	Reconnaissance
		Segmentation (Naturel)	Modélisation, Géomé- trique	Classification

TABLE 2.2 – Caractérisation de l'approche retenue pour la reconnaissance des plantes dans le cadre de la thèse

naturelles, par l'introduction d'une description modélisant les propriétés géométriques discriminantes des feuilles, et une méthode efficace de classification adaptée aux spécificités du problème.

2.5.1 Segmentation robuste d'images naturelles

Les images de feuilles sur fond naturelles, d'autant plus si elles sont acquises via un smartphone, constituent un véritable défi pour des algorithmes de segmentation. Le fait que l'objet d'intérêt soit situé au sein d'un environnement composite, dans lequel on risque de retrouver des objets de couleurs très proches, rend totalement inopérantes les méthodes globales comme les seuillages. Dans l'ensemble, les conditions d'acquisitions sont peu maîtrisées, ce qui produit des images difficiles comme le montre la Figure 2.13.

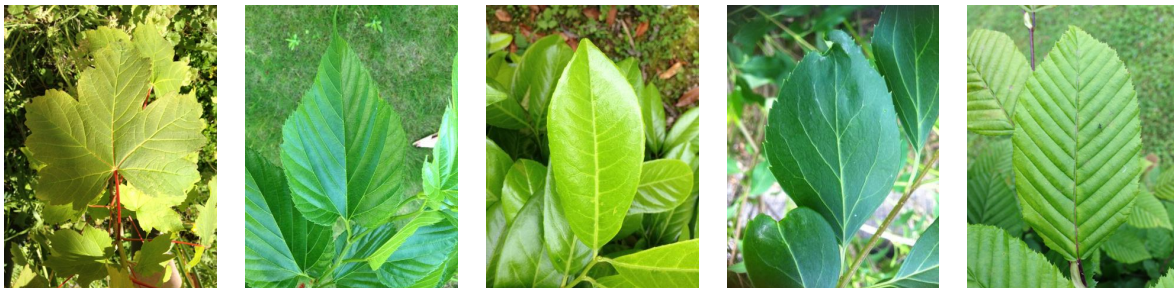


FIGURE 2.13 – Images de feuilles prises en environnement naturel par un appareil photo de smartphone

Les régions de l'image définies par les feuilles sont d'ailleurs assez peu homogènes, avec la présence potentielle de reflets, d'ombres, de taches, ou de variations importantes de couleur à l'automne. On va de plus très souvent trouver des contours marqués à l'intérieur des feuilles, du fait des nervures notamment, ce qui complique la détection du contour des objets. Les critères à prendre en compte pour isoler la feuille de son environnement ne sont donc pas si évidents que cela.

La méthode de segmentation devra donc faire en sorte d'éviter ces difficultés, et prendre en compte à la fois l'information d'homogénéité de couleur et les contours apparents pour espérer obtenir une estimation correcte de la région définie par la feuille. L'intervention d'un utilisateur pourra dans ce sens s'avérer utile pour limiter le risque d'erreur. Enfin, l'introduction de connaissances *a priori* sur la forme des objets à extraire présente un intérêt certain pour s'assurer que les régions obtenues aient une forme conforme à ce qu'il est possible de trouver parmi les feuilles.

2.5.2 Description géométrique inspirée par la botanique

L'étude attentive des critères utilisés dans les clés de détermination botaniques pour caractériser les espèces permet de donner une idée précise des formes qu'il va falloir évaluer afin de faciliter l'identification. Des années d'expertise botanique constituent un fort argument en faveur de descripteurs dédiés expressément à la caractérisation des formes foliaires, fournissant une connaissance extérieure des endroits où regarder et des structures à chercher.

Le principal intérêt de descripteurs spécifiques, destinés à capturer indépendamment les formes de la feuille (forme globale, bord, base, sommet), réside dans la possibilité de décorrélérer les différents aspects nécessaires à son identification, quand des descriptions de formes plus globales et statistiques ne permettraient pas cette séparation. C'est un point important, puisque dans une même espèce, certains critères botaniques peuvent être bien plus variables que d'autres. Il est donc important de les considérer de manière indépendante.

La modélisation des propriétés botaniques à suffisamment haut niveau offrirait en outre la possibilité d'une interprétation sémantique. Si les mesures sont prises avec le même degré d'abstraction impliqué par la caractérisation botanique, il serait possible de mettre directement en lien les informations tirées de l'image et les données contenues dans les flores.

2.5.3 Apprentissage automatique de nombreuses classes

La comparaison des descripteurs extraits d'une image avec ceux de toute une base serait un moyen bien trop lourd pour parvenir à reconnaître l'espèce, surtout dans un contexte où toutes les données sont embarquées et tous les traitements effectués sur un processeur de téléphone. Il semble donc indispensable de réaliser un apprentissage hors-ligne de façon à rendre la phase de reconnaissance plus légère et procéder à une véritable classification.

Toutes les méthodes de classification ne seront pas adaptées à un problème qui considère un nombre de classes très élevé (plus de 100 espèces d'arbres natives). Il faut de plus prendre en compte certaines spécificités du problème, en particulier la grande variabilité qui peut exister au sein d'une même classe. Le fait est que pour une espèce, certaines formes vont être peu significatives car trop variables, alors que la donnée d'une forme en particulier va permettre d'apporter énormément de certitude quant à l'identification.

La prise en compte de l'incertitude liée à l'extraction des descripteurs et à la qualité de la segmentation est aussi un élément important pour la décision. Parvenir à déterminer efficacement quelle masse de certitude est représentée par chaque élément tiré de l'image et arriver à reporter cette information dans la reconnaissance de l'espèce permettrait d'affiner la décision.

Bibliographie

- [AGB⁺12] Akhil Arora, Ankit Gupta, Nitesh Bagmar, Shashwat Mishra, and Arnab Bhattacharya. A plant identification system using shape and morphological features on segmented leaflets. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [AMK97] Sadegh Abbasi, Farzin Mokhtarian, and Josef Kittler. Reliable classification of chrysanthemum leaves through curvature scale space. In *International Conference on Scale-Space Theory in Computer Vision*, pages 284–295, 1997.
- [BCF⁺08] P. Belhumeur, D. Chen, S. Feiner, D. Jacobs, W. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. Sheorey, S. White, and L. Zhang. Searching the world’s herbaria : A system for visual identification of plant species. In *European Conference on Computer Vision*, 2008.
- [BCMB09] André Ricardo Backes, Dalcimar Casanova, and Odemir Martinez Bruno. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 42(1) :54 – 67, 2009.
- [BM92] Serge Beucher and Fernand Meyer. The morphological approach to segmentation : the watershed transformation. *Mathematical Morphology in Image Processing*, 1992.
- [BMP01] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 :509–522, 2001.
- [BTVG06] Herbert Bay, Tinne Tuytelaars, and Luc J. Van Gool. SURF : Speeded up robust features. In *European Conference on Computer Vision, Lecture Notes in Computer Science*, pages 404–417, 2006.
- [BYM⁺12] Vera Bakic, Itheri Yahiaoui, Sofiène Mouine, Saloua Litayem, Wajih Ouertani, Anne Verroust-Blondet, Hervé Goëau, and Alexis Joly. Inria imedia2’s participation at imageclef 2012 plant identification task. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.

- [CA10] Carlos Caballero and M. Carmen Aranda. Plant species identification using leaf image retrieval. In *Proceedings of the ACM International Conference on Image and Video Retrieval, CIVR 2010*, pages 327–334, 2010.
- [CBFMB11] Dalcimar Casanova, João Batista Florindo, and Odemir Martinez Bruno. Ifsc/usp at imageclef 2011 : Plant identification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
- [CBFNGMB12] Dalcimar Casanova, João Batista Florindo, Wesley Nunes Gonçalves, and Odemir Martinez Bruno. Ifsc/usp at imageclef 2012 : Plant identification task. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [CDF⁺04] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.
- [Che93] Chaur-Chin Chen. Improved moment invariants for shape discrimination. *Pattern Recognition*, 5(26) :683–686, 1993.
- [CNMJS06] J. Camargo Neto, G. Meyer, D. Jones, and A. Samal. Plant species identification using Elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture*, 50(2) :121–134, 2006.
- [CRBW10] James S. Cope, Paolo Remagnino, Sarah Barman, and Paul Wilkin. Plant texture classification using gabor co-occurrences. In *Advances in Visual Computing*, volume 6454 of *Lecture Notes in Computer Science*, pages 669–677, 2010.
- [DWZ07] Ji Xiang Du, Xiao Feng Wang, and Guo-Jun Zhang. Leaf shape based plant species recognition. *Applied Mathematics and Computation*, 185 :883–893, 2007.
- [FS11] Stefan Fiel and Robert Sablatnig. Automated identification of tree species from images of the bark, leaves and needles. In *Proceedings of the 16th Computer Vision Winter Workshop*, 2011.
- [GBJ⁺11] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The CLEF 2011 plant images classification task. In *CLEF 2011 Working Notes*, 2011.
- [GBJ⁺12] H. Goëau, P. Bonnet, A. Joly, D. Barthelemy, N. Boujemaa, and J.-F. Molino. The ImageCLEF 2012 plant image identification task. In *ImageCLEF 2012 Working Notes*, 2012.
- [GBJ⁺13] H. Goëau, P. Bonnet, A. Joly, V. Bakic, D. Barthelemy, N. Boujemaa, and J.-F. Molino. The ImageCLEF 2013 Plant Identification Task. In *CLEF 2013 Working Notes*, 2013.

- [GJY⁺11] Hervé Goëau, Alexis Joly, Itheri Yahiaoui, Pierre Bonnet, and Elise Mouysset. Participation of inria & pl@ntnet to imageclef 2011 plant images classification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
- [HS88] C. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*, pages 147–151, 1988.
- [Hu62] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2) :179–187, 1962.
- [HZD⁺06] Zhi-Kai Huang, Chun-Ho Zheng, Ji-Xiang Du, Zhong-Hua Quan, and Sheng-Bo Guo. Bark classification based on gabor filter features using rbpnn neural network. In *International Conference on Neural Information Processing*, pages 80–87, 2006.
- [ICL10] Chinese Academy of Sciences Intelligent Computing Laboratory. ICL plant leaf image dataset, <http://www.intelengine.cn/english/dataset>, 2010.
- [INK98] Cholhong Im, Hirobumi Nishida, and Tosiyasu L. Kunii. A hierarchical method of recognizing plant species by leaf shapes. In *IAPR Workshop on Machine Vision Applications*, pages 158–161, 1998.
- [KBB⁺12] Neeraj Kumar, Peter Belhumeur, Arijit Biswas, D. Jacobs, I. Kress, W. Lopez, and Soares Joao V. B. Leafsnap : A computer vision system for automatic plant species identification. In *European Conference on Computer Vision*, volume 2, pages 502–516, 2012.
- [KPP10] D. Knight, J. Painter, and M. Potter. Automatic plant leaf classification for a mobile field guide. Technical report, Standford University, 2010.
- [KYÜ11] Hanife Kebapci, Berrin A. Yanikoglu, and Gözde B. Ünal. Plant image retrieval using color, shape and texture features. *The Computer Journal*, 54(9) :1475–1490, 2011.
- [LC06] Yan Li and David Dagan Chi, Zheru Feng. Leaf vein extraction using independent component analysis. In *IEEE International Conference on Systems, Man and Cybernetics*, pages 3890–3894, 2006.
- [LJ07] Haibin Ling and D.W. Jacobs. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2) :286–299, 2007.
- [Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60 :91–110, 2004.
- [MA04] F. Mokhtarian and S. Abbasi. Matching shapes with self-intersections : Application to leaf classification. *IEEE Transactions on Image Processing*, 13(5) :653–661, 2004.

- [MB95] E.N. Mortensen and W.A. Barrett. Intelligent scissors for image composition. In *SIGGRAPH*, pages 191–198, 1995.
- [MM92] F. Mokhtarian and A.K. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14 :789–805, 1992.
- [MYB12] Olfa Mzoughi, Itheri Yahiaoui, and Nozha Boujemaa. Extraction of leaf parts by image analysis. In *International Conference on Image Analysis and Recognition (ICIAR)*, pages 348–358, 2012.
- [MYVB12] Sofiène Mouine, Itheri Yahiaoui, and Anne Verroust-Blondet. Advanced shape context for plant species identification using leaf image retrieval. In *International Conference on Multimedia Retrieval*, page 49, 2012.
- [NZ06] Maria-Elena Nilsback and Andrew Zisserman. A visual vocabulary for flower classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1447–1454, 2006.
- [NZ07] M-E. Nilsback and A. Zisserman. Delving into the whorl of flower segmentation. In *British Machine Vision Conference*, volume 1, pages 570–579, 2007.
- [NZ08] Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes. In *ICVGIP*, pages 722–729, 2008.
- [Ots79] Nobuyuki Otsu. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 :62–66, 1979.
- [RKB04] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut" : interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23(3) :309–314, 2004.
- [SK00] Takeshi Saitoh and Toyohisa Kaneko. Automatic recognition of wild flowers. In *International Conference on Pattern Recognition*, pages 2507–2510, 2000.
- [SK04] T. Saitoh and T. Kaneko. Automatic recognition of blooming flowers. In *International Conference on Pattern Recognition*, volume 1, pages 27–30, 2004.
- [Söd01] Oskar Söderkvist. Computer vision classification of leaves from swedish trees. Master's thesis, Linköping University, 2001.
- [SZ03] Josef Sivic and Andrew Zisserman. Video google : A text retrieval approach to object matching in videos. In *IEEE International Conference on Computer Vision*, pages 1470–1477, 2003.

- [Tea80] Michael Reed Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America*, 70(8) :920–930, 1980.
- [TKC09] C.-H. Teng, Y.-T. Kuo, and Y.-S. Chen. Leaf segmentation, its 3d position estimation and leaf classification from a few images with very close viewpoints. In *Proceedings of the 6th International Conference on Image Analysis and Recognition, ICIAR '09*, pages 937–946, 2009.
- [WBX⁺07] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yuxuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. A leaf recognition algorithm for plant classification using probabilistic neural network. In *IEEE International Symposium Signal Processing and Information Technology*, 2007.
- [WCF03] Z. Wang, Z. Chi, and D. Feng. Shape based leaf image retrieval. *IEE Proceedings - Vision, Image and Signal Processing*, 150 :34–43, 2003.
- [WCFW00] Z. Wang, Z. Chi, D. Feng, and Q. Wang. Leaf image retrieval with shape features. In *Advances in Visual Information Systems*, volume 1929 of *Lecture Notes in Computer Science*, pages 41–52. 2000.
- [WHD⁺08] X.-F. Wang, D.-S. Huang, J.-X. Du, X. Huan, and L. Heutte. Classification of plant leaf images with complicated background. *Applied Mathematics and Computation*, 205(2) :916–926, 2008.
- [YAT11] Berrin A. Yanikoglu, Erchan Aptoula, and Caglar Tirkaz. Sabanci-okan system at imageclef 2011 : Plant identification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
- [YAT12] Berrin A. Yanikoglu, Erchan Aptoula, and Caglar Tirkaz. Sabanci-okan system at imageclef 2012 : Combining features and classifiers for plant identification. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [Zer34] F. Von Zernike. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode (Diffraction theory of the cut procedure and its improved form, the phase contrast method). *Physica*, 1(7-12) :689–704, 1934.

Chapitre 3

Représentations paramétriques de feuilles

Des modèles pour comprendre les formes foliaires



Cercis siliquastrum L.

Résumé du Chapitre

Ce chapitre détaille les modèles paramétriques mis au point pour représenter efficacement les différentes formes que peuvent prendre les feuilles d'arbres. Il détaille dans un premier temps le modèle polygonal destiné à représenter les formes globales des feuilles simples, et son extension aux feuilles palmatilobées. Le cas un peu plus complexe des feuilles composées est ensuite abordé, et les deux modèles nécessaires à leur représentation décrits. Enfin, des modélisations pour les formes plus locales à la base et à l'apex sont également proposés. La construction géométrique de chacun des modèles est présentée dans le détail, et le chapitre est conclu par une brève étude des utilisations possibles des modèles introduits.



LA DIVERSITÉ des formes couvertes par les feuilles rend la définition d'un modèle unique, capable de rendre compte de toutes les structures et combinaisons possibles de formes locales comme globales, complètement illusoire. Cependant, synthétiser la connaissance *a priori* disponible sur les formes foliaires, sur leurs configurations vraisemblables, sur leur pouvoir discriminant, présente un intérêt évident dans le cadre d'un processus de traitement automatique.

La segmentation d'objets naturels dans des images en environnement complexe est une tâche ardue où l'introduction de connaissances extérieures ne peut être qu'un atout. En effet connaître à l'avance les formes que peut prendre l'objet que l'on cherche à isoler constitue un avantage certain, même si leur variabilité rend cette connaissance difficile à exploiter. L'utilisation de contraintes très flexibles de façon à permettre une large expressivité des formes, mais dans le même temps suffisamment restrictives pour interdire des formes en dehors du spectre des possibles, serait évidemment bénéfique pour un algorithme de segmentation.


De manière analogue, le fait de savoir au préalable quels sont les critères morphologiques des feuilles qui vont s'avérer discriminants pour identifier les espèces végétales permet d'envisager une description intuitivement plus efficace, en tout cas porteuse d'une certaine charge sémantique. En utilisant des modèles spécifiquement dédiés à la représentation de ces formes discriminantes, la caractérisation de la feuille que l'on obtient se rapprocherait de celle que pourrait faire un observateur humain, et pourrait simuler son pouvoir de discernement.

C'est dans cette optique d'aider à la fois à l'extraction et à la description que nous avons développé des modèles pour les différentes formes et structures que peuvent prendre les feuilles d'arbres. Ces modèles relativement simples sont pensés pour être à la fois capables de s'adapter à la vaste diversité des formes rencontrées au fil des espèces d'arbres, mais aussi suffisamment délimités dans leur construction pour ne correspondre qu'à l'objet de leur définition : les formes de feuilles.



3.1 Modèle polygonal de feuille simple

Les formes les plus évidentes à représenter sont sans doute les formes globales des feuilles simples. Les termes utilisés pour les décrire s'apparentent en effet à des formes géométriques plutôt intuitives (cercle, ellipse, ovale, triangle, cœur) qui semblent à première vue pouvoir être modélisées. Le fait est que, dans la nature, les formes prises par les feuilles vont rarement pouvoir être rattachées à un terme unique, représentatif d'une forme canonique, et tout l'intérêt d'un modèle de forme globale serait justement de pouvoir représenter des formes se situant entre plusieurs de ces jalons.

L'idée qui préside donc à la définition d'un tel modèle est de chercher à quantifier, par le biais de paramètres significatifs, la localisation de la forme représentée dans l'univers des formes référencé par les termes botaniques. Une rapide observation de ce qui caractérise ces formes, d'après les descriptions sommaires qui peuvent en être données, permet de dégager plusieurs grands axes schématisés sur la Figure 3.1, selon lesquels les divers qualificatifs botaniques vont venir s'échelonner :

► **Circularité** :  : de **Linéaire** à **Orbiculaire** les formes varient en étant de plus

en plus larges et de plus en plus arrondies.

- ▶ **Ovalité** :  : de **Obovale** à **Ovale** les formes varient en étant de plus en plus larges près de leur base et de plus en plus fines près de leur sommet.
- ▶ **Rectangularité** :  : de **Rhomboidale** à **Oblongue** les formes varient en étant de plus en plus allongées, les cotés de plus en plus longuement parallèles.

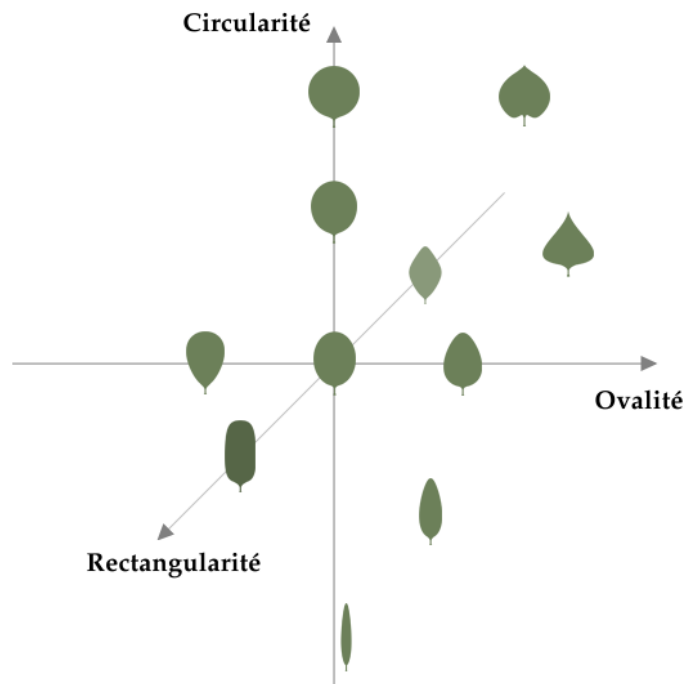


FIGURE 3.1 – Axes principaux de variations des formes globales des feuilles simples : circularité, ovalité, rectangularité

On imagine ainsi aisément tout le spectre des formes possibles entre des feuilles linéaires très allongées et des feuilles orbiculaires presque circulaires. La continuité de ce spectre est quelque chose qui n'est pas véritablement capturé par les termes botaniques, mais qui peut parfaitement l'être par des paramètres numériques mesurant la position d'une forme sur cet axe. C'est cette réflexion qui nous a poussé à utiliser trois paramètres fondamentaux, assimilables à ces trois axes, pour décrire la forme globale des feuilles simples : la **largeur maximale** de la feuille, la **position verticale** où cette largeur est atteinte, et la **longueur latérale** sur laquelle les bords vont être approximativement parallèles.

La manière la plus simple de mettre en pratique ces grandeurs pour l'instant descriptives dans un modèle géométrique de forme est de construire un polygone à partir de ces paramètres. C'est cette idée de simplicité pour passer de la description à la forme et *vice versa* qui a guidé la suite du travail présenté. Les mesures considérées pouvant toutes être prises relativement à la longueur de la feuille, il sera plus pratique de bâtir un polygone en se basant sur un axe représentant la longueur de

la feuille. Préalablement à la construction, on suppose donc connus deux points B et A représentant respectivement la base et l'apex de la feuille. À ce stade, à partir de deux points et de trois paramètres numériques, on peut déjà construire une forme sommaire :

Le paramètre de **position verticale** p , défini relativement à la longueur $BA = h$, et généralement entre 0 et 1, permet de définir le point central C de la feuille. On notera \vec{a} le vecteur unitaire de l'axe de la feuille, tel que $\vec{BA} = h \cdot \vec{a}$, et C vérifie alors :

$$\vec{BC} = p \cdot \vec{BA} = p \times h \cdot \vec{a}$$

Le paramètre de **largeur maximale** l , lui aussi la plupart du temps entre 0 et 1 (même s'il peut arriver que la feuille soit plus large que longue !) va définir un segment $[C_g C_d]$ perpendiculaire à l'axe (BA) , centré sur C , et de longueur $l \times h$. Si l'on note \vec{n} le vecteur normal unitaire ($\|\vec{n}\| = 1$ et $\vec{n} \cdot \vec{a} = 0$) on peut définir les deux points C_g et C_d de la façon suivante :

$$\vec{CC_g} = \frac{l}{2} \times h \cdot \vec{n}$$

$$\vec{CC_d} = -\frac{l}{2} \times h \cdot \vec{n}$$

Enfin, la **longueur latérale** L permet de construire sur chacun des côtés un segment parallèle à l'axe principal, de longueur $L \times h$, centré sur le point du bord respectif. Ces segments latéraux $[T_g^B T_g^A]$ et $[T_d^B T_d^A]$ font en fait office de tangentes donnant une forme plus ou moins étirée. Pour le côté gauche par exemple, ils sont simplement définis par :

$$\vec{C_g T_g^A} = \frac{L}{2} \times h \cdot \vec{a}$$

$$\vec{C_g T_g^B} = -\frac{L}{2} \times h \cdot \vec{a}$$

Cette première phase de construction est illustrée dans la Figure 3.2.

Si l'on relie alors les points ainsi définis, on obtient une forme polygonale capable de varier en suivant la plupart des différences possibles entre les formes foliaires. Cependant, les formes

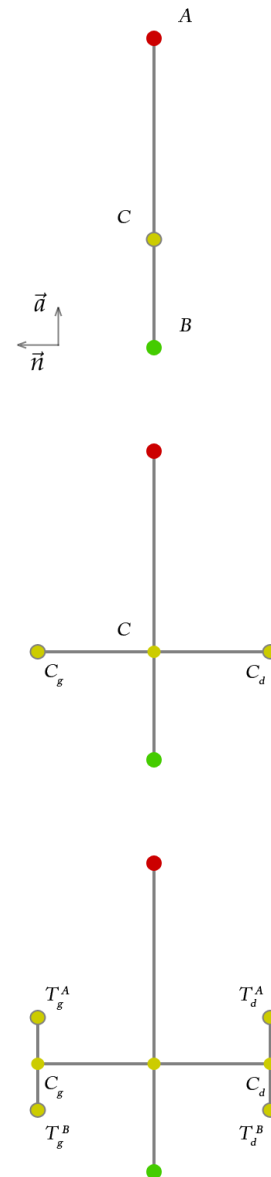


FIGURE 3.2 – Construction du modèle polygonal paramétrique de forme globale (I)

obtenues sont très rudimentaires et les angles trop marqués, notamment au sommet et à la base pour correspondre aux formes bien souvent arrondies des feuilles. De plus, le caractère pointu ou arrondi du limbe va souvent aller de pair avec la largeur de la feuille à ces emplacements, et une feuille obovale sera généralement plus arrondie à son sommet et plus en pointe en direction de sa base. C'est une spécificité de forme, certes plus locale, mais qu'il convient de bien prendre en compte si l'on veut produire un modèle capable de suivre fidèlement la forme d'une feuille.

On rajoute ainsi deux paramètres numériques supplémentaires pour définir la forme globale de la feuille au niveau de la base et de l'apex. Il s'agit bien là de forme globale, à savoir à quel point la forme se rétrécit à ses extrémités ou reste au contraire ouverte, ce qui diffère des formes apicales et basales bien plus locales qui servent également à la caractérisation des feuilles, et qui seront abordées dans la Section 3.5. Les grandeurs considérées correspondent à des **angles d'ouverture** à la **base** et au **sommet** du polygone et permettent une construction intuitive :

Les paramètres d'**angle à la base** α_B et d'**angle à l'apex** α_A sont utilisés pour construire deux triangles isocèles $B^B B_g B_d$ et $A^A A_g A_d$ placés au niveau des points B et A et dont les côtés égaux sont de longueur constante. Ils auront alors respectivement pour angles au sommet α_B et α_A et les milieux de leurs bases seront B et A . La définition géométrique des points s'écrit alors :

$$\overrightarrow{AA^A} = \frac{h}{5} \times \cos\left(\frac{\alpha_A}{2}\right) \vec{n}$$

$$\overrightarrow{AA_g} = \frac{h}{5} \times \sin\left(\frac{\alpha_A}{2}\right) \vec{a}$$

$$\overrightarrow{AA_d} = -\frac{h}{5} \times \sin\left(\frac{\alpha_A}{2}\right) \vec{a}$$

$$\overrightarrow{BB^B} = -\frac{h}{5} \times \cos\left(\frac{\alpha_B}{2}\right) \vec{n}$$

$$\overrightarrow{BB_g} = \frac{h}{5} \times \sin\left(\frac{\alpha_B}{2}\right) \vec{a}$$

$$\overrightarrow{BB_d} = -\frac{h}{5} \times \sin\left(\frac{\alpha_B}{2}\right) \vec{a}$$

Les **dix points** que l'on obtient alors permettent de tracer le polygone $\Pi = A^A A_d T_d^A T_d^B B_d B^B B_g T_g^B T_g^A A_g$ représentatif de la forme globale de la feuille. Cette seconde phase de construction est illustrée dans la Figure 3.3.

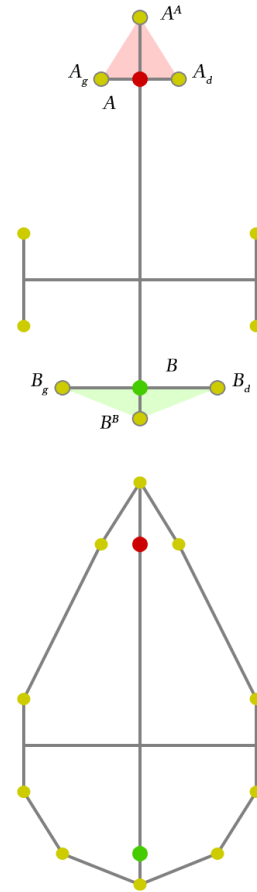


FIGURE 3.3 – Construction du modèle polygonal paramétrique de forme globale (II)








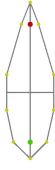
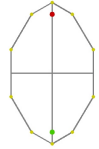
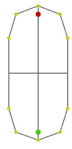

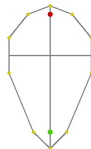







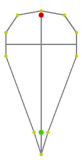
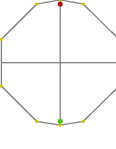
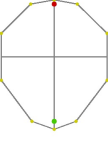
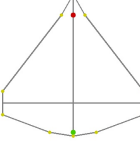
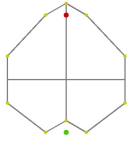
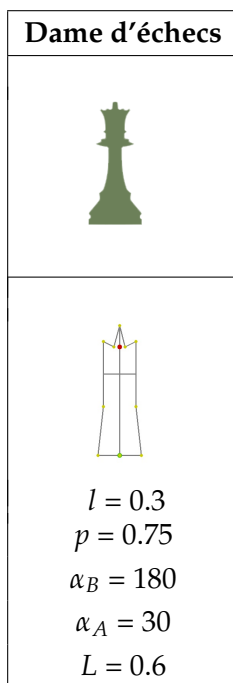
Linéaire	Lancéolée	Elliptique	Oblongue	Ovale	Obovale
					
					
$l = 0.2$ $p = 0.5$ $\alpha_B = 40^\circ$ $\alpha_A = 40^\circ$ $L = 0.4$	$l = 0.4$ $p = 0.4$ $\alpha_B = 90^\circ$ $\alpha_A = 60^\circ$ $L = 0.35$	$l = 0.7$ $p = 0.5$ $\alpha_B = 120^\circ$ $\alpha_A = 120^\circ$ $L = 0.4$	$l = 0.5$ $p = 0.5$ $\alpha_B = 140^\circ$ $\alpha_A = 140^\circ$ $L = 0.6$	$l = 0.7$ $p = 0.35$ $\alpha_B = 140^\circ$ $\alpha_A = 90^\circ$ $L = 0.3$	$l = 0.7$ $p = 0.65$ $\alpha_B = 90^\circ$ $\alpha_A = 140^\circ$ $L = 0.3$
Rhomboidale	Spatulée	Orbiculaire	Suborbiculaire	Triangulaire	Cordée
					
					
$l = 0.7$ $p = 0.5$ $\alpha_B = 90^\circ$ $\alpha_A = 90^\circ$ $L = 0.2$	$l = 0.6$ $p = 0.75$ $\alpha_B = 35^\circ$ $\alpha_A = 160^\circ$ $L = 0.2$	$l = 1.0$ $p = 0.5$ $\alpha_B = 160^\circ$ $\alpha_A = 160^\circ$ $L = 0.4$	$l = 0.9$ $p = 0.55$ $\alpha_B = 140^\circ$ $\alpha_A = 160^\circ$ $L = 0.4$	$l = 1.2$ $p = 0.25$ $\alpha_B = 160^\circ$ $\alpha_A = 60^\circ$ $L = 0.2$	$l = 1.0$ $p = 0.45$ $\alpha_B = 240^\circ$ $\alpha_A = 120^\circ$ $L = 0.4$

TABLE 3.1 – Adéquation visuelle des formes globales canoniques du limbe foliaire et de leur modélisation paramétrique

En jouant alors sur les valeurs des différents paramètres ($l; p; L; \alpha_B; \alpha_A$), on parvient sans problème à recréer manuellement l'ensemble des formes canoniques de feuilles, comme le montre la Table 3.1. Le résultat est un ensemble de formes polygonales visuellement très proches des représentations de la forme foliaire, bien que manquant quelque peu de flexibilité dans le cas des formes les plus arrondies.

Un tel modèle possède plusieurs propriétés qui le rendent particulièrement intéressant. Tout d'abord il parvient à condenser en seulement 5 valeurs numériques un nombre conséquent de formes. Ce jeu pourtant réduit de paramètres permet en effet, à partir de deux points, de couvrir le spectre notablement étendu des formes de feuilles simples, et d'en fournir une description numérique directement interprétable. Cette description compacte et puissante à la fois par son pouvoir d'expression et sa dimension presque sémantique constitue de fait un atout pour la caractérisation des objets.



Cette expressivité est en revanche à double tranchant puisque, pour embrasser de façon assez exhaustive l'ensemble des formes possibles pour une feuille, le modèle polygonal n'en demeure pas moins capable de créer des formes qui ne correspondent en aucun cas à des feuilles vraisemblables. Il y aurait donc nécessité de restreindre les valeurs pouvant être prises par les paramètres pour être certain de ne modéliser que des formes foliaires. Un intervalle autorisé pour chaque paramètre permet déjà d'éviter quelques formes aberrantes :

$$l \in [0.1, 2.5], p \in [0.1, 0.9], \alpha_B \in [20, 300], \alpha_S \in [20, 300], L \in [0.1, 0.6]$$

On voit déjà également émerger des corrélations évidentes entre les différents paramètres (lorsque la position verticale p augmente, alors l'angle à l'apex α_A aussi, les feuilles plus larges à leur sommet y étant aussi généralement plus arrondies) qui tendent à suggérer que qu'il n'y a qu'un domaine délimité de valeurs pouvant être prises en même temps par les paramètres du modèle, de façon à obtenir des formes de feuille.

La construction du modèle est malgré tout en elle-même porteuse d'une certaine quantité de connaissance botanique *a priori* puisqu'elle inclut une forte notion de symétrie axiale (qui est présente chez quasiment l'intégralité des espèces d'arbres, à quelques exceptions près) qui l'empêche tout de même de représenter n'importe quelle forme. De même, sa définition par dix points la contraint à des formes très compactes et n'autorise que peu de concavités (principalement à la base et à l'apex là où les formes globales de feuilles peuvent elles-mêmes en présenter) ou d'irrégularité.

Le dernier avantage est le peu de possibilité d'obtenir des formes similaires avec des jeux de paramètres différents. En effet, à part peut-être dans le cas de formes allongées où la longueur des tangentes latérales ne va finalement pas changer le profil de la forme, la rigidité des segments qui forment tous des angles entre eux assure que la variation des paramètres se répercute de façon visible sur la forme résultante, garantissant ainsi que la description fournie par les valeurs des paramètres est une caractérisation pertinente de la forme.

La modélisation proposée des formes globales de feuilles simples permet donc d'offrir une description numérique capable de représenter les formes canoniques définies par la botanique mais aussi et surtout une quantification de toutes les formes intermédiaires. Cette flexibilité lui permet de s'adapter sans difficulté aux formes globales de feuilles simples réelles comme le montre la Figure 3.4

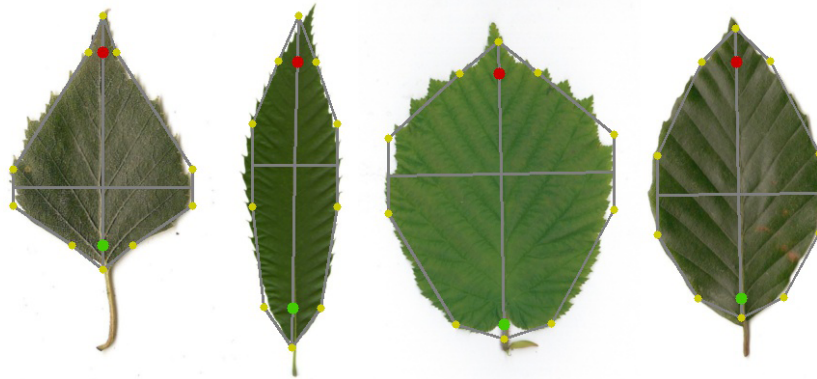


FIGURE 3.4 – Exemples de modèles de feuille simple ajustés à la main sur différentes feuilles réelles

Les valeurs nécessaires à sa définition sont en nombre très réduit ce qui garantit à la fois une limitation dans l'expressivité (malgré tout suffisante pour couvrir les formes foliaires simples, voire plus) et une légèreté dans la manipulation. Le résultat est donc un modèle porteur de connaissance *a priori* mais extrêmement flexible, et permettant une réversibilité entre forme et description numérique. Ces qualités en font un objet potentiellement utile pour de la segmentation avec *a priori* de forme, la description morphologique des feuilles, et la reconstruction de formes.

3.2 Modèle polygonal de feuille palmatilobée

Le modèle polygonal introduit pour représenter les formes globales des feuilles simples ne permet malheureusement pas de reconstruire tous les types de feuilles, et notamment les feuilles lobées. Si les feuilles pennatilobées peuvent souvent être vues comme des feuilles simples dont le bord présente des sinuosités très marquées (on en revient ici à l'artificialité et la subjectivité évidente des termes botaniques) les feuilles palmatilobées, dont les lobes se rattachent tous à la base de la feuille, voient leur forme structurée de manière véritablement différente. Il en résulte que le modèle de feuille simple pourra nettement moins bien être utilisé pour les feuilles palmatilobées, ainsi que l'illustre la Figure 3.5

Aucun terme de forme globale ne pourrait donc s'appliquer à une feuille palmatilobée. En revanche, ces termes ne manquent pas d'être utilisés pour décrire la forme des lobes, composantes de la forme de la feuille. Ces lobes sont invariablement en nombre impair, répartis symétriquement de part et d'autre d'un lobe principal axé sur la nervure principale de la feuille, et présentent en effet dans l'immense majorité des cas tous des formes similaires. Un modèle capable de représenter les formes des feuilles simples peut donc être utilisé pour représenter les formes des lobes des feuilles lobées.

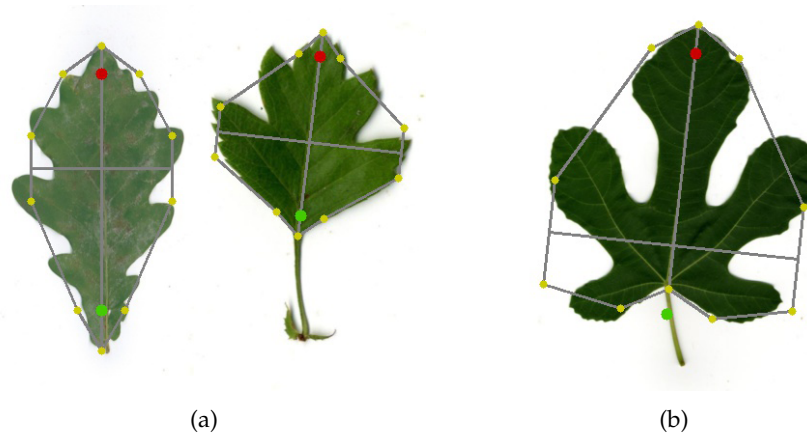


FIGURE 3.5 – Adaptation du modèle de feuille simple aux feuilles pennatilobées (a) et impossibilité d'adéquation avec les feuilles palmatilobées (b)

De façon plus précise, afin de modéliser les feuilles palmatilobées, nous avons fait l'hypothèse (sans doute un peu approximative) qu'une feuille palmatilobée comme celle de la Figure 3.6 était assimilable à une superposition d'un **nombre impair** de lobes simples, partageant tous une **forme globale identique**, répartis **symétriquement** par rapport à la nervure principale, et divergeant tous à partir du **point d'insertion** du pétiole.

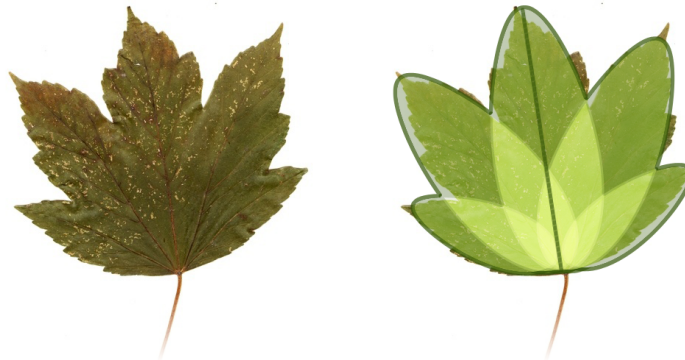


FIGURE 3.6 – Exemple d'approximation d'une feuille palmatilobée par une superposition symétrique de lobes simples identiques

La définition d'un modèle polygonal pour les feuilles palmatilobées va ainsi pouvoir s'appuyer sur celui des feuilles simples, avec cette idée de lobes symétriques reliés au même point et d'une même forme (donc des mêmes paramètres) pour tous les modèles de lobes. La construction d'un tel modèle va donc reposer sur un certain nombre de paramètres numériques supplémentaires. Le premier n'est autre que le **nombre de lobes** de la feuille, et puisque l'on suppose ce nombre impair et les lobes répartis équitablement par rapport au lobe principal on peut ne considérer que le **nombre n_L de lobes** additionnels sur chaque côté de la feuille. Pour une feuille simple, ce paramètre vaudra donc 0, puis 1 pour une feuille à 3 lobes, 2 pour une feuille à 5 lobes, et ainsi de suite. Pour construire ensuite les modèles correspondant à chacun des lobes, il est alors simplement nécessaire de déterminer les deux points constituant leur axe.

De façon à ce que tous se rattachent bien au même point censé être le point d'insertion du pétiole, c'est le point extrême du polygone B_B qui servira d'attache et non le point de base B . Tous les axes des lobes auront ainsi la même base. Pour définir l'équivalent du point A pour chaque lobe, il suffit alors de connaître son orientation et sa distance au point d'insertion, et un modèle polygonal simple pourra alors être bâti autour de cet axe.

Etant donnée la contrainte de symétrie par rapport à l'axe principal, on se limite à définir pour chacune des n_L paires de lobes secondaires numérotée l deux paramètres θ_l et h_l correspondant respectivement à l'angle et à la **longueur relative** de l'axe du lobe par rapport à l'axe principal. Ces valeurs permettent de construire les points A_g^l et A_d^l situés au sommet de l'axe des deux lobes symétriques :

$$\overrightarrow{B_B A_g^l} = h_l \times h (\cos \theta_l \cdot \vec{a} + \sin \theta_l \cdot \vec{n})$$

$$\overrightarrow{B_B A_d^l} = h_l \times h (\cos \theta_l \cdot \vec{a} - \sin \theta_l \cdot \vec{n})$$

La construction de chaque modèle Π_l est ensuite la même que pour le modèle simple Π . Elle est basée sur les mêmes valeurs de paramètres, avec la seule particularité que le point d'insertion B_B sert de point extrême à tous les lobes (le point correspondant n'est donc pas calculé pour les polygones des lobes secondaires) de façon à ce qu'il soit bien situé sur le contour de la forme $\cup \Pi_l$ définie par le modèle, obtenue en superposant les modèles polygonaux. La Figure 3.7 illustre ce processus.

La modélisation ainsi obtenue est relativement rigide. Les hypothèses de symétrie et d'identité des formes notamment sont des contraintes fortes, qui risquent de ne pas toujours être vraies dans les cas de forme naturelles. La variabilité naturelle que l'on peut retrouver entre les différentes feuilles d'une même espèce peut en effet se voir répercutée entre les différents lobes d'une même feuille. Autrement dit, on considère comme régulières des formes qui par nature sont tout sauf immuables. La forme générale botanique est par ailleurs une notion assez floue, et il y a une différence importante entre dire qu'un même terme s'applique à tous les lobes d'une feuille (ce qui est bien souvent le cas dans les descriptions des flores) et prétendre que la forme de ces lobes va pouvoir être représentée par exactement les mêmes valeurs numériques, ce qui apparaît sur certains exemples peu réguliers de la Figure 3.8.

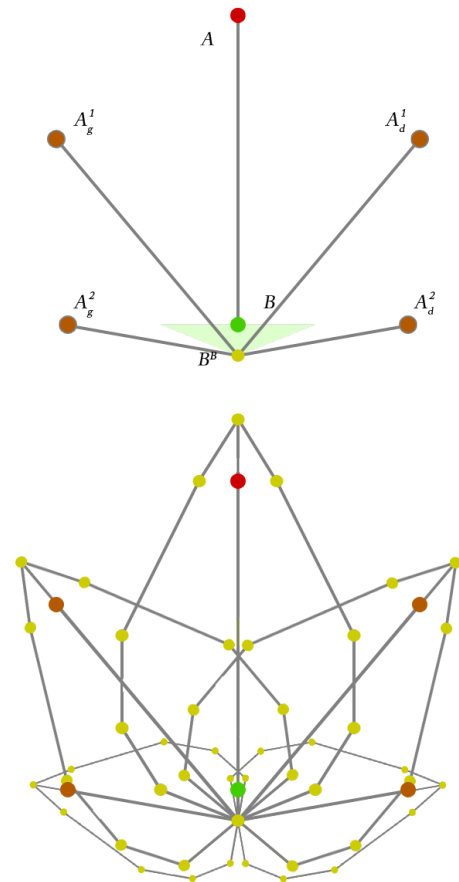


FIGURE 3.7 – Construction du modèle polygonal paramétrique de forme palmatilobée

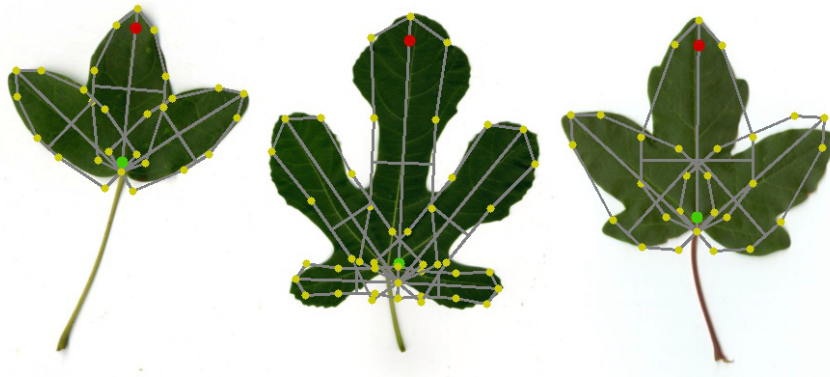


FIGURE 3.8 – Exemples de modèles de feuille palmatilobée ajustés à la main sur différentes feuilles réelles, et problèmes dus à l’hypothèse de symétrie

Cependant, on peut voir cette contrainte comme une limite à l’expressivité du modèle qui, si elle empêche donc le polygone de s’adapter avec exactitude à la forme de n’importe quelle feuille, définit un domaine de formes peu éloigné de l’ensemble des formes possibles, avec moins de formes aberrantes que si l’on avait, par exemple, autorisé chacun des lobes à avoir son propre jeu de paramètres de forme simple, ou permis à deux lobes correspondants d’avoir des longueurs et des directions différentes. Le compromis entre l’expressivité d’un modèle (l’ensemble des formes pouvant être représentées) et sa vraisemblance (le fait de ne pouvoir prendre que des formes plausibles pour l’objet modélisé) est difficile à trouver et il vaut sans doute mieux ne pas donner trop de latitude à un modèle déjà capable d’une grande variété, vu son nombre de degrés de liberté.

L’autre limite de cette modélisation vient du fait que l’on considère une intersection de formes. De ce fait, la forme de chacun des lobes dans cette zone où il se recouvrent ne va pas être très significative, et par conséquent les paramètres numériques qui l’influent perdent de leur représentativité. Il est en effet probablement plus facile de créer des formes proches avec des valeurs très différentes pour certains paramètres que dans le cas du modèle simple, comme le montre la Figure 3.9. C’est un écueil difficilement évitable si l’on souhaite conserver la même simplicité de construction, mais dont il faut être conscient si l’on veut utiliser cette représentation comme une description de la forme modélisée.

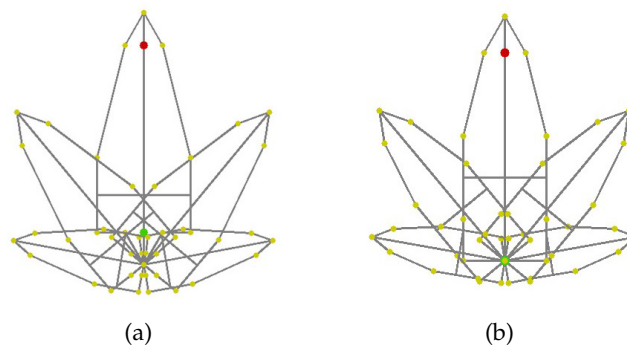


FIGURE 3.9 – Formes très proches obtenues avec des jeux de paramètres différents : $l=0.5$, $p=0.2$, $\alpha_B=60^\circ$, $\alpha_A=60^\circ$, $L=0.4$, $\theta_1=40^\circ$, $h_1=0.9$, $\theta_2=80^\circ$, $h_2=0.6$ (a) $l=0.4$, $p=0.4$, $\alpha_B=180^\circ$, $\alpha_A=60^\circ$, $L=0.4$, $\theta_1=40^\circ$, $h_1=0.85$, $\theta_2=80^\circ$, $h_2=0.55$ (b)

Ce modèle constitue en fait un modèle unique regroupant à la fois les feuilles simples et les feuilles

palmatilobées puisqu'il suffit d'avoir un paramètre n_L à 0 pour retrouver un modèle polygonal simple. On se retrouve donc avec une description sémantique très forte où la valeur d'un seul paramètre va permettre d'établir une différence entre les espèces à feuilles simples et celles à feuilles palmatilobées (rares étant les cas d'espèces pouvant présenter à la fois ces deux types quoique présents - le lierre *Hedera helix* L. pour n'en citer qu'un).

On dispose donc au final d'une modélisation légère, nécessitant la définition de deux points et la donnée de $6 + 2 \times n_L$ paramètres numériques ($l; p; L; \alpha_B; \alpha_A$) et $(n_L; \theta_1; h_1; \dots; \theta_{n_L}; h_{n_L})$ permettant de recréer de manière approximative (du fait de la définition polygonale et des fortes contraintes de symétrie) l'ensemble des formes prises par les feuilles non-composées. Les valeurs des différents paramètres permettent une description d'assez haut niveau de la forme de la feuille, directement interprétable, à condition de se rappeler que tous ne seront pas aussi pertinents selon les cas de figure (α_B et p dans un modèle à plusieurs lobes en particulier). La capacité de ce modèle à n'exprimer des formes que dans un domaine fini mais de façon extrêmement flexible permet d'envisager son introduction dans une méthode de segmentation basée modèle, en cherchant à optimiser automatiquement ses différents paramètres pour modéliser le mieux possible la feuille présente dans une image.

3.3 Modèle synthétique de feuille composée

Il existe malgré tout des types de feuilles qui ne rentrent pas dans la modélisation proposée. C'est évidemment le cas des feuilles composées, où le limbe n'est plus d'un seul tenant mais constitué de plusieurs **folioles** indépendantes, reliées au rachis, prolongement du pétiole. La feuille est alors bien l'ensemble de ces folioles, même si cela est quelque peu contraire au bon sens, et c'est cet objet qu'il faut représenter si l'on souhaite modéliser la feuille.

La difficulté avec ce type de feuille, outre le fait que sa définition soit assez peu intuitive pour le non-initié, est que la variété des formes résultantes est encore plus importante que dans le cas des feuilles simples. En effet, plus que de la diversité des formes de folioles, les variations les plus importantes vont venir des différences de nombre et d'organisation des folioles qui vont produire des formes extrêmement éloignées. Il faut ajouter la particularité que, au contraire d'un limbe unique, la feuille composée contient des parties flexibles qui vont pouvoir faire apparaître des formes globales assez peu prévisibles. C'est cette première diversité qu'il va falloir essayer de capturer dans un modèle avant de s'intéresser aux folioles prises séparément.

Les folioles peuvent s'organiser suivant plusieurs dispositions données dans la Figure 3.11, ainsi qu'on l'a évoqué, selon que la feuille est pennée (folioles échelonnées sur l'axe, en nombre pair ou impair), digitée (folioles divergeant du même point), trifoliée (trois folioles) ou bipennée (deux niveaux de composition). En laissant de côté les feuilles doublement composées dont la modélisation serait une gageure, on se retrouve tout de même avec des organisations bien différentes. Toutefois, introduire un modèle pour chacune de ces possibilités irait quelque peu à l'encontre du but de synthèse de la connaissance *a priori* dans un modèle. S'il faut déterminer le type de feuille avant de pouvoir utiliser la connaissance pour la modéliser, cela perd un peu de son intérêt.

Nous avons donc fait le choix d'utiliser un modèle unique pour synthétiser les différentes organisations possibles. Il s'inspire en premier lieu de la structure **imparipennée** qui est la plus commune, mais peut sans trop de difficultés s'adapter aux autres types d'organisations. Comme le modèle de feuille simple, il suppose connus deux points notés cette fois P et T correspondant respectivement à l'imparipennée et au centre de la imparipennée. L'idée est ensuite de représenter la position des folioles par des disques dont les centres et les rayons vont pouvoir être déterminés par un ensemble de paramètres. Le premier paramètre est bien entendu le **nombre de folioles** que comporte la feuille, ou sous l'hypothèse de symétrie axiale (valable pour les feuilles pennées comme digitées) du **nombre n_F de paires de folioles** secondaires.

Pour chacun de ces n_F couples de folioles numéroté f , on associe une **position verticale** p_f correspondant à sa position relative sur l'axe $[PT]$. Cela permettrait de placer le point centre C_f de la paire sur l'axe ; cependant celui-ci pouvant être courbe, nous introduisons un paramètre de **courbure** k permettant de placer les centres sur une courbe parabolique :

$$\overrightarrow{PC_f} = p_f \times h \cdot \vec{a} + k (p_f(1 - p_f)) \times h \cdot \vec{n}$$

Les points F_f^g et F_f^d définissant chaque foliole seront disposés de part et d'autre de cet axe courbe en suivant la direction normale à la tangente à la courbe en C_f soit $\vec{n}_f = \vec{N}_f / \|\vec{N}_f\|$ avec $\vec{N}_f = k(2p_f - 1) \cdot \vec{a} + \vec{n}$. La **distance** à laquelle ils se trouvent de l'axe fait l'objet d'un paramètre unique δ exprimé relativement à la hauteur :

$$\overrightarrow{C_f F_f^g} = \delta \times h \cdot \vec{n}_f \quad \overrightarrow{C_f F_f^d} = -\delta \times h \cdot \vec{n}_f$$

Enfin, on trace en chacun de ces points ainsi qu'au point terminal T un cercle représentant la foliole. Le rayon de ces cercles est le même pour tous et dépend d'un paramètre r de **rayon relatif**. On obtient donc finalement un ensemble de $2 \times n_F + 1$ cercles éventuellement superposés représentant l'organisation d'une feuille composée. Cette construction est illustrée dans la Figure 3.10.

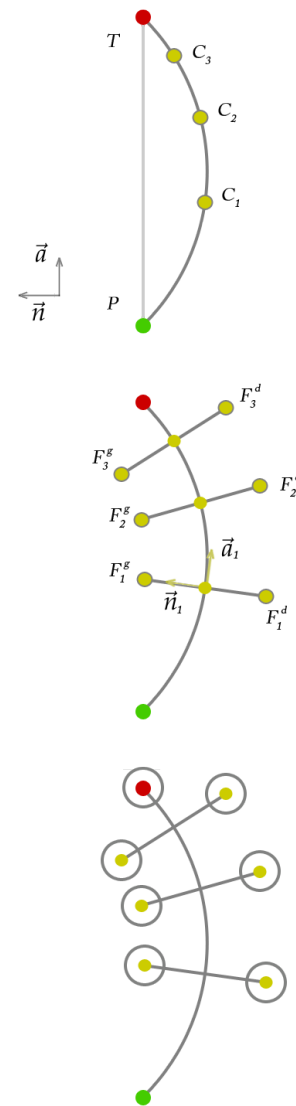


FIGURE 3.10 – Construction du modèle d'organisation de feuille composée

Ce modèle permet de capturer de façon satisfaisante le nombre et la localisation des folioles, y compris pour des feuilles qui intuitivement ne correspondraient pas à un tel modèle. Cette capacité d'adaptation est illustrée en Figure 3.11. Il ne faut pas oublier que le but n'est pas nécessairement de déterminer la structure véritable de la feuille, et on peut obtenir une représentation parfaitement acceptable en plaçant la foliole terminale entre les deux folioles supérieures d'une feuille paripennée, ou encore en jouant sur la distance d pour bien obtenir un cercle par foliole d'une feuille digitée.

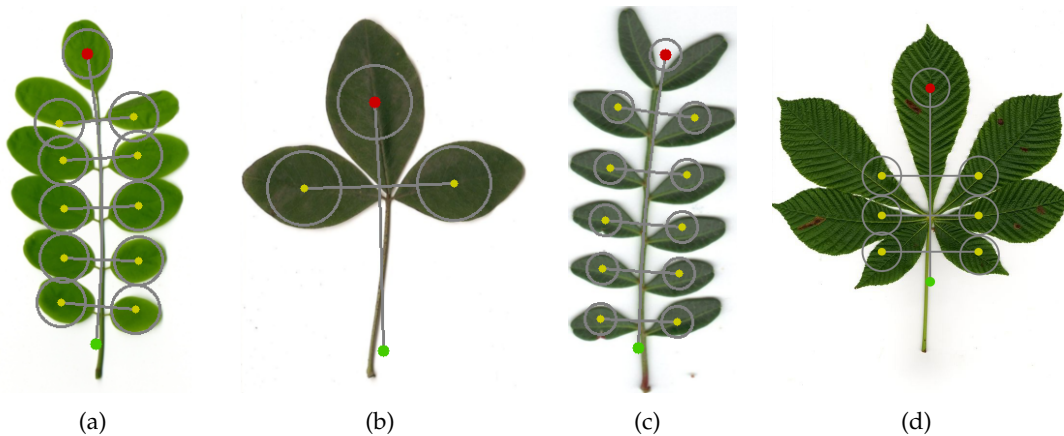


FIGURE 3.11 – Adaptation aux différents types de feuilles composées : imparipennée (a), trifoliée (b), paripennée (c) et digitée (d)

Des suppositions assez fortes sont toutefois faites sur la symétrie axiale des feuilles, ainsi que sur la régularité des folioles qui se trouvent ici toutes de la même taille, qui ne seront pas toujours vérifiées. Le modèle tient aussi pour acquise la répartition échelonnée des folioles le long du rachis, toutes à la même distance de l'axe de la feuille, ce qui constitue souvent une approximation.

Un certain nombre de paramètres auraient pu être rajoutés pour rendre le modèle capable de s'adapter avec plus de précision aux différentes formes possibles, par exemple en rendant les rayons différents selon les folioles (avec la possibilité d'un rayon nul pour la foliole terminale), en autorisant aussi des distances variables pour les différentes paires de folioles ou même en ajoutant un paramètre d'angle de la branche d'une foliole avec l'axe principal. Il faut bien voir que ce seraient là autant de degrés de libertés supplémentaires qui, s'ils amélioreraient l'adaptabilité du modèle, le rendraient moins significatif puisque trop peu contraint pour ne représenter que des feuilles et trop libre pour assurer une réversibilité entre forme et paramètres, sans compter une complexité accrue. C'est encore une fois le compromis expressivité/vraisemblance qui préside aux choix faits pour la définition de ce modèle.

Tel quel, nous disposons d'une modélisation capable de rendre compte efficacement de grandes différences dans l'organisation des feuilles composées, contenant notamment des informations sur le nombre de folioles, une information capitale pour l'identification de l'espèce. Il possède surtout l'avantage d'être suffisamment général pour pouvoir s'adapter à différents types de feuilles composées sans nécessité d'évaluer dans un premier temps la structure de la feuille pour ensuite choisir un modèle adapté. Il conserve malgré cela suffisamment de rigidité pour que la description fournie par ses paramètres possède un sens exploitable.

3.4 Modèles polygonaux joints de folioles

Une fois connu l'agencement des folioles, c'est évidemment leur forme individuelle qui va devoir être examinée. D'un point de vue du vocabulaire botanique, la forme globale des folioles peut être décrite de la même façon que celle des feuilles simples, il est donc parfaitement envisageable de tirer parti de la modélisation des feuilles simples pour représenter chaque foliole d'une feuille composée.

De la même façon que pour les feuilles palmatilobées, nous avons fait des hypothèses fortes quant à la régularité de la feuille d'un point de vue de la forme. La principale supposition consiste à considérer que toutes les folioles ont une **forme identique** et que celle-ci est **simple non-palmatilobée**. La base de la modélisation des folioles va donc être un ensemble de modèles polygonaux simples (avec un paramètre n_L fixé à 0) bâtis à partir de points différents mais possédant les mêmes paramètres de forme.

On pourrait à ce stade définir un modèle extrêmement rigide plaçant les points nécessaires à la construction des polygones des folioles de façon homogène. Les paramètres nécessaires à ce modèle seraient d'abord les mêmes que pour l'organisation des feuilles composées (à savoir courbure k , position verticale p_f et distance δ) pour construire les points B_f correspondant à la base des modèles de la même façon que les centres des cercles :

$$\overrightarrow{PC_f} = p_f \times h \cdot \vec{a} + k (p_f(1 - p_f)) \times h \cdot \vec{n}$$

$$\overrightarrow{C_f B_f} = (-1)^{f-1} \delta \times h \cdot \vec{n}_f$$

Pour placer les points d'apex A_f il suffirait de définir une longueur de folioles h_F et une orientation θ_F qui seraient les mêmes pour tous les modèles polygonaux. La seule différence concerne les points B_0 et A_0 correspondant à la foliole puisqu'on pose $C_0 = B_0 = T$. La définition des apex devient alors :

$$\overrightarrow{B_f A_f} = h_F \times h \left(\cos(\theta_F) \cdot \vec{n}_f + (-1)^{f-1} \sin(\theta_F) \cdot \vec{a}_f \right)$$

$$\overrightarrow{B_0 A_0} = h_F \times h \cdot \vec{a}_0$$

Une fois définis pour chaque foliole les points B_f et A_f on utilise les mêmes paramètres l , p , α_B , α_A en L pour construire les modèles polygonaux Π_f basés sur chacun de ces axes. Cette construction est illustrée dans la Figure 3.12.

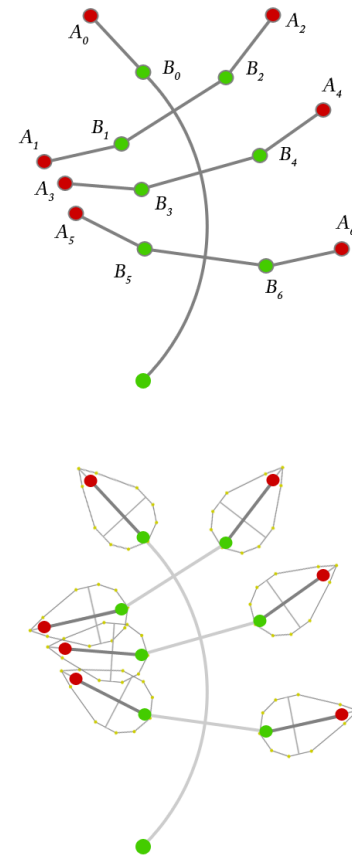


FIGURE 3.12 – Construction du modèle rigide de folioles

Cependant, si l'on compare un modèle construit de cette façon rigide avec une feuille réelle (Figure 3.13) il apparaît clairement que la modélisation n'est pas satisfaisante. La flexibilité impliquée par la nature des feuilles composées fait que l'orientation des folioles les unes par rapport aux autres est très difficilement prévisible, et la considérer comme un paramètre global ne permet pas de s'adapter à cette variabilité. De plus les variations de la taille relative des folioles peuvent être conséquentes, et il faudrait donc faire varier cette longueur indépendamment pour chaque foliole. De même, la symétrie des paires n'est pas toujours rigoureuse, une foliole pouvant venir s'insérer sur le rachis avec un petit décalage sur son homologue, ou les pétioles (pétioles des folioles) ayant des longueurs un peu différentes.

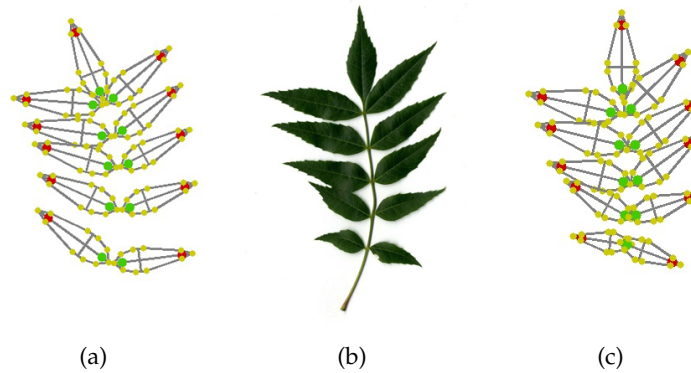


FIGURE 3.13 – Comparaison entre un modèle de folioles rigide (a) et une image de feuille composée réelle (b) par rapport à un modèle complètement flexible aux points libres (c)

Le cas des feuilles composées est donc plus épineux que celui des feuilles palmatilobées, où la variabilité entre les lobes, si elle est bel et bien présente, est atténuée par la continuité du limbe, ainsi que par le nombre relativement limité de lobes. Pour capturer ces variations et pouvoir s'adapter avec précision aux formes des différentes folioles, il faut donc, en lieu et place des grandeurs globales d , h_F et θ_F , paramétrer pour le modèle de chaque foliole numérotée f sa position verticale p_f , sa distance à l'axe d_f , sa longueur h_f et son orientation θ_f , et conserver la même construction.

Chaque modèle de foliole requiert alors la donnée de 4 paramètres numériques pour déterminer les coordonnées de 2 points. Seulement, avec deux degrés de liberté par point, il est parfaitement équivalent de considérer directement comme variables les 4 coordonnées elles-mêmes. La paramétrisation n'a dès lors plus tellement d'intérêt et on peut se contenter de placer ces points librement. Une telle modélisation où les points définissant les polygones des folioles sont libres permet une flexibilité bien plus grande et va pouvoir s'adapter aux variations naturelles de taille et d'orientation des folioles qu'on ne manquera pas de rencontrer, comme l'illustre la Figure 3.13 (c). Les variables nécessaires sont donc directement les points $(B_0; A_0; \dots; B_{2 \cdot n_F}; A_{2 \cdot n_F})$ auxquels il faut rajouter les paramètres des modèles des folioles $(l; p; L; \alpha_B; \alpha_A)$.

La seule contrainte d'un tel modèle réside alors dans l'obligation qu'ont les polygones à avoir la même forme (à un facteur de taille près). Il s'agit bien entendu d'une approximation, mais qui permet de dégager la forme dominante des folioles comme caractéristique représentative de la feuille. Le fait

de relâcher la contrainte de régularité sur les points permet d'aligner les axes des modèles de manière plus exacte sur les folioles, ce qui donnera par conséquent une approximation plus précise de la forme globale moyenne des folioles, puisque les polygones seront plus proches de la feuille.

Le principal problème que pose la forme apparente d'une feuille composée est la possibilité de recouvrement entre folioles. Lorsque toutes se superposent, il devient très difficile d'être certain de l'adéquation du modèle avec la feuille. Le problème vient du fait que les polygones peuvent alors se recouper les uns les autres, et de la même façon que dans la zone de superposition du modèle palmatilobé, les variations de certains paramètres (largeur l , position p notamment) risquent de ne pas influencer sur la forme émergente, comme le souligne la Figure 3.14. La conséquence en est bien évidemment une diminution de la représentativité des valeurs des paramètres.

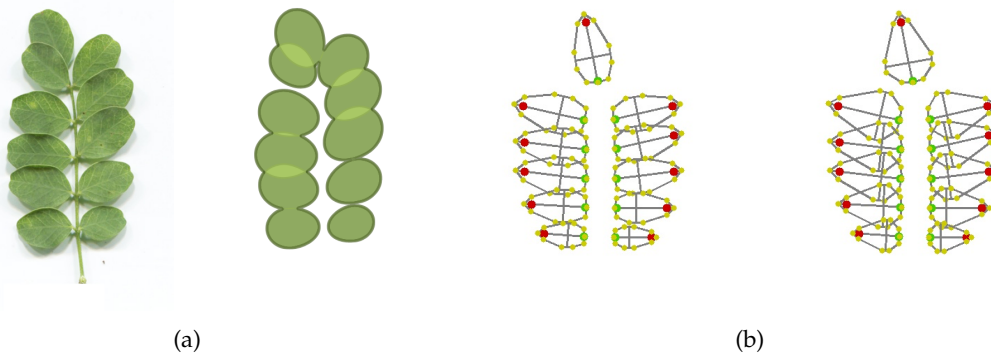


FIGURE 3.14 – Exemple de recouvrement des folioles d'une feuille composée (a) et perte d'influence des paramètres du modèle sur la forme apparente (b)

Pour contrecarrer ce risque, nous avons simplement choisi de considérer comme forme résultante du modèle non pas la superposition des modèles polygonaux que l'on pourrait noter $\cup \Pi_f$, mais leur union disjointe $\sqcup \Pi_f$, de façon à ce que les intersections de plusieurs polygones ne fassent pas partie de la forme définie par le modèle. Dans l'éventualité où des folioles se recouvrent en grande partie, la forme risque tout de même d'être mal évaluée, mais le fait que les paramètres doivent être les mêmes pour toutes les folioles assure un lissage de ce problème local. Cela reste ainsi un moyen simple mais très efficace de s'assurer que les paramètres de forme des folioles conservent une influence majeure sur la forme produite, et donc une pertinence comme description.

En modélisant en deux temps les feuilles composées, d'abord pour capturer leur organisation globale, puis en s'attachant à évaluer du mieux possible la forme dominante de leurs folioles, on extrait une information de très haut niveau. La description produite caractérise le nombre et l'agencement des folioles, et leurs tailles et orientations relatives peuvent ensuite être examinées (par exemple pour différencier feuilles pennées et digitées). Quant à leur forme, elle peut être modélisée de façon très robuste grâce à des contraintes implicites contenues dans la définition du modèle. La contrainte d'identité de forme comme celle de non-intersection garantit un point de vue global sur la forme des folioles. Cette représentation jointe évite ainsi les problèmes que poserait une modélisation séparée (un modèle polygonal indépendant par foliole) et fournit une description synthétique et compacte de la forme globale des folioles.

3.5 Modèles de formes basales et apicales

En plus des formes globales et structurelles de la feuille, il est intéressant de modéliser des formes plus locales, qui font elle aussi l'objet d'une description poussée chez les botanistes. En particulier, les formes du limbe autour de sa base et de son apex présentent une diversité importante que l'on retrouve dans le vocabulaire botanique. Une fois encore, les termes employés correspondent simplement à des archétypes d'une réalité géométrique plus continue, et en pratique moins cloisonnée. Transposer cette caractérisation dans un domaine numérique plus libre présente donc un certain attrait.

À la différence des formes globales où, pour dégager quelques caractéristiques de haut niveau, on peut tolérer une certaine approximation, une modélisation polygonale ne pourra pas suffire à rendre compte des formes assez fines du bord de la feuille dans ces zones précises. Il faut donc imaginer un modèle capable de déformations moins rigides mais paramétrable par un nombre réduit de paramètres potentiellement significatifs. Lorsque l'on dissèque les spécificités des formes canoniques, les critères apparaissant comme porteurs de sens sont malheureusement plus difficiles à mettre en évidence que les mesures simples utilisées pour les formes globales.

Dans le cas de la base comme de l'apex, ce qui semble caractériser la forme est en premier lieu l'**angle** formé par les deux bords du limbe à cet endroit précis, qui va pouvoir être plus ou moins aigu (sommet acuminé, base en coin) voire plat (forme tronquée) ou rentrant (base en cœur). Il y a en revanche une grosse part d'incertitude quand à l'échelle à laquelle il faut considérer la feuille pour observer ces formes. Le contexte local va être déterminant, et on ne peut considérer les formes basales et apicales qu'en se basant sur un point précis (base ou apex) du contour de la feuille qui servira de référence pour évaluer la forme. L'angle dont il est question pourra alors être vu comme l'angle formé par les directions prises par le bord à partir de ce point-là, soit un angle entre tangentes.

En plus de cette valeur d'angle, les termes font référence à une notion d'arrondi plus ou moins marqué, ou au contraire de pointe effilée présentant donc un profil plutôt concave. La courbure, ou du moins l'**inflexion** du contour à partir du point de référence est donc à prendre en compte. La **taille** d'une éventuelle pointe va aussi séparer plusieurs catégories (pour l'apex, acuminée, cuspidée, mucronée renvoient à des pointes de tailles décroissantes) bien que plus difficile à quantifier du fait de sa relativité à l'échelle. Enfin, certaines **dissymétries** vont pouvoir être rencontrées, et elles sont bien plus importantes à capturer que lorsqu'on considère la forme globale, notion plus synthétique. Les différences marquées entre les deux côtés de la feuille (base oblique, sommet recourbé) vont d'ailleurs s'avérer extrêmement caractéristiques de certaines espèces.



FIGURE 3.15 – Diversité des courbes formées par la feuille au niveau de son apex et de sa base

L'idée est donc de construire un modèle de courbes pouvant, avec un ensemble de paramètres s'inspirant des critères évoqués, reproduire la diversité des formes présentée sur la Figure 3.15. Nous proposons d'utiliser pour cela des courbes de Bézier qui, pouvant être définies simplement à l'aide de quelques points, semblent à première vue suffisamment expressives pour convenir.

La construction de modèles de formes basales et apicales passera donc par la définition de deux courbes de Bézier se rejoignant en un même point, respectivement B^* et A^* , représentant dans un cas le point d'insertion du pétiole, et dans l'autre le point du bord correspondant à l'extrémité de la nervure principale de la feuille (à noter que dans le cas d'une feuille composée, l'objet d'étude devient la foliole, et qu'on a alors autant de points B_f^* et A_f^* que de folioles). Dans la suite, nous détaillerons uniquement la construction pour le modèle d'apex, la méthode pour la base étant identique.

On commence par donner un axe au modèle à l'aide un paramètre d'orientation o_A qui permettra un alignement différent de l'axe de la forme globale. Il faut ensuite fixer la taille du modèle, chose que nous avons choisi de faire uniformément en fixant $h_A = h/5$. Enfin, un paramètre d'angle à l'apex α_A^* pas nécessairement égal à celui du modèle polygonal α_A permet de construire les points extrémités des courbes des bords droit et gauche A_d^* et A_g^* comme la base d'un triangle isocèle d'angle α_A en son sommet A^* :

$$\overrightarrow{A^*A_d^*} = h_A \left(\cos(o_A + \frac{\alpha_A^*}{2}) \cdot \vec{a} + \sin(o_A + \frac{\alpha_A^*}{2}) \cdot \vec{n} \right)$$

$$\overrightarrow{A^*A_g^*} = h_A \left(\cos(o_A - \frac{\alpha_A^*}{2}) \cdot \vec{a} + \sin(o_A - \frac{\alpha_A^*}{2}) \cdot \vec{n} \right)$$

Restent à placer pour chaque côté deux points de contrôle de la courbe de Bézier. Puisque ceux-ci définissent des tangentes, il est intéressant de les construire en faisant intervenir des paramètres d'angles relatifs à la direction du côté qui permettront notamment de trouver l'angle effectif de la courbe au point A^* et de caractériser son inflexion. On exprime ces angles à partir du sommet A^* (θ_A^d et θ_A^g) pour les premiers et du point extrémité de chaque côté (θ_d^A et θ_g^A) pour les seconds, de façon à ce qu'une mesure positive place toujours le point vers l'extérieur de la forme.

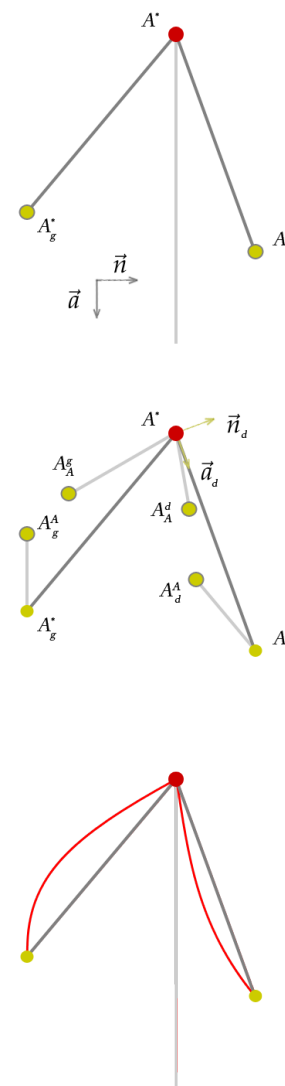


FIGURE 3.16 – Construction du modèle paramétrique de forme apicale

Il faut aussi ajouter pour chacun un paramètre de distance à son point référence (δ_A^d et δ_A^s , puis δ_d^A et δ_g^A) exprimé relativement à la longueur h_A :

$$\overrightarrow{A^*A_A^d} = \delta_A^d \times h_A \left(\cos(\theta_A^d) \cdot \vec{a}_d + \sin(\theta_A^d) \cdot \vec{n}_d \right)$$

$$\overrightarrow{A_d^*A_d^A} = \delta_d^A \times h_A \left(-\cos(\theta_d^A) \cdot \vec{a}_d + \sin(\theta_d^A) \cdot \vec{n}_d \right)$$

$$\overrightarrow{A^*A_A^s} = \delta_A^s \times h_A \left(\cos(\theta_A^s) \cdot \vec{a}_g - \sin(\theta_A^s) \cdot \vec{n}_g \right)$$

$$\overrightarrow{A_g^*A_g^A} = \delta_g^A \times h_A \left(-\cos(\theta_g^A) \cdot \vec{a}_g - \sin(\theta_g^A) \cdot \vec{n}_g \right)$$

On obtient ainsi pour chaque côté une séquence de 4 points $A^*A_A^dA_d^AA_d^*$ et $A^*A_A^sA_g^AA_g^*$ qui constituent chacun le polygone de contrôle d'une courbe de Bézier représentant alors la forme de ce côté du bord de la feuille. La construction du modèle Ψ_A constitué de ces deux courbes est illustrée sur la Figure 3.16.

La possibilité d'avoir des paramètres d'angles et de tailles différents selon le côté permet de représenter des formes largement dissymétriques, qui peuvent se présenter dans la réalité. Elle va surtout autoriser de petites différences locales qui permettront de mieux s'adapter aux formes réelles de la base et de l'apex. En effet si l'on observe un peu des feuilles réelles, on voit que bien souvent il y a peu de véritable symétrie, notamment en ce qui concerne la courbure des bords. Dès qu'une pointe apparaît (c'est donc surtout vrai pour l'apex) comme sur les feuilles de la Figure 3.17, un bord va presque toujours être plus convexe que l'autre, témoin d'une légère inclinaison naturelle plus que d'une forme véritablement irrégulière.

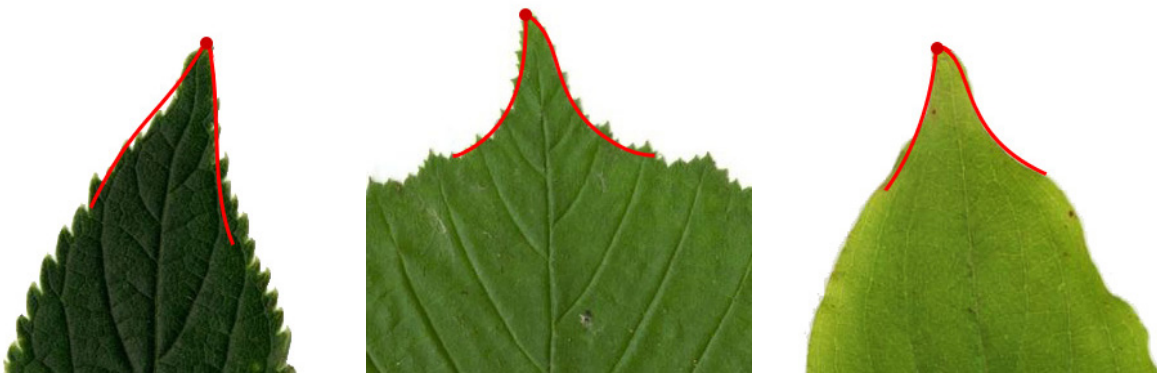


FIGURE 3.17 – Asymétrie visible des formes réelles d'apex pour des espèces décrites comme régulières

Les termes du vocabulaire botanique renvoient pourtant eux à des formes idéales, donc symétriques, à des projections de l'esprit qui redresse naturellement les imperfections dues à une variabilité naturelle. Il est donc indispensable de laisser au modèle la latitude nécessaire pour s'adapter à ces











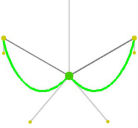
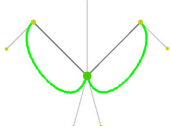
Atténuée	En Coin	Arrondie	Tronquée	En Cœur	Echancrée
					
					

TABLE 3.2 – Exemples de formes de bases générées à l’aide d’un modèle paramétrique de courbes

petites irrégularités, mais il sera par conséquent très difficile d’interpréter les différences entre les valeurs des paramètres pour les deux côtés comme une mesure de la dissymétrie de la forme. En revanche, l’opération qui consiste à régulariser la forme pourra être effectuée simplement en calculant la valeur moyenne des paramètres des bords gauche et droit.

D’un point de vue de l’expressivité, un tel modèle sans contrainte pourrait prendre des formes très variées, dont celles de la Table 3.2, mais aussi certaines totalement impossibles pour des formes foliaires. C’est par exemple le cas si le placement des points de contrôle rend les courbes tellement concaves qu’elles en viennent à se croiser, ou encore si les forces des tangentes sont telles qu’une boucle se forme sur un des côtés. On peut régler simplement ces problèmes en bornant le domaine des valeurs prises par les angles des tangentes ($\theta_A^g, \theta_g^A \in [-\theta_{max}, \theta_{max}]$), ainsi qu’en contraignant les tangentes à ne pas s’entrecroiser ($\delta_A^d + \delta_d^A < 1$).

La description de forme que l’on obtient sous forme de valeur d’angles et de rapports de distances est cette fois-ci beaucoup moins évidente à interpréter directement à haut niveau. Elle n’en demeure pas moins très caractéristique de la forme, et pour le coup parfaitement réversible. En effectuant une moyenne gauche-droite des descripteurs, on parvient tout de même à condenser l’information de forme locale de la base ou de l’apex dans un nombre très réduit de paramètres qui peuvent même être manipulés pour tirer une information d’acuité ou de concavité de la feuille à cet endroit. Sa construction flexible en revanche va permettre de s’adapter sans aucun problème aux formes des feuilles, mais cette grande élasticité est aussi le signe que le modèle est moins porteur de connaissance *a priori* que ne pouvaient l’être les modèles dédiés à des formes plus globales.

Ainsi, il apparaît que dès lors qu’on s’intéresse à des variations locales, il faut être capable d’une grande adaptabilité pour conserver une cohérence avec les formes réellement observées. Il est donc légitime de s’interroger sur la pertinence d’introduire une modélisation pour des détails aussi fins ; cela aurait par exemple peu de sens de construire un modèle *a priori* pour la forme des dents de la feuille tant celle-ci semble variable, et plus appropriée à une analyse *a posteriori*. La base et l’apex se situent entre deux, et l’introduction d’un modèle porteur d’une somme certes réduite de connaissance sur la relative symétrie, les courbures possibles de la feuille à ces emplacements se justifie par la description de plus haut niveau fournie. Celle-ci sera vraisemblablement pertinente puisqu’inspirée

de critères discriminants, et aura au moins l'avantage en conservant un minimum de rigidité de tolérer un certain niveau de bruit.

3.6 Utilités des modélisations de formes

Les modélisations des différentes formes foliaires proposées dans ce chapitre offrent plusieurs possibilités d'utilisation, qui se résument toutes à l'apport d'une connaissance *a priori* dans le traitement des feuilles. En effet, ces modèles contiennent tous dans leur définition et leur construction une part importante de savoir botanique, puisqu'ils sont pensés pour recréer les formes distinguées dans les ouvrages spécialisés, en s'appuyant essentiellement sur les éléments qui permettent de différencier ces formes.

Il semble évident qu'une chaîne de traitement d'images ne peut que bénéficier d'un apport extérieur de connaissance sur son objet d'étude qui devrait permettre à chaque étape de prendre de meilleures décisions. La segmentation ne pourra produire que des formes en adéquation avec ce qu'il est possible de rencontrer chez l'objet en question, la description qui utilise les critères les plus déterminants sera probablement plus compacte et plus riche en informations, la classification pourra tirer avantage de probabilités antérieures relatives à la coexistence de plusieurs formes, à l'importance à donner à tel ou tel critère.

3.6.1 Segmentation avec *a priori* de forme

L'emploi le plus évident de tels modèles est la prise en compte d'un *a priori* de forme pour la segmentation des feuilles dans des images éventuellement complexes. Au cours de cette phase, on va chercher à isoler une région de pixels dans l'image qui devront correspondre à la feuille d'intérêt. C'est une tâche complexe qui sera rendue un peu plus aisée si l'on connaît à l'avance les formes que peut prendre l'objet que l'on cherche. On peut ainsi envisager, dans le cas d'une feuille simple par exemple, d'adapter un modèle de forme globale à l'image pour avoir une première idée de la région qui sera retenue, tout en étant certain que sa forme s'accorde avec ce que la connaissance botanique définit comme possible.

L'avantage commun à toutes ces modélisations est de reposer sur un nombre de paramètres relativement faible pour créer un vaste éventail de formes. Elles sont donc faciles à déformer en jouant sur les valeurs de ces paramètres, et la dimensionnalité de l'espace des formes créées et forcément plus réduite que si la forme était parfaitement libre. Par conséquent on peut envisager d'en faire des modèles déformables avec l'idée d'en optimiser itérativement les paramètres de construction en examinant tour à tour toutes les variations possibles, variations qui demeureront en nombre raisonnable. Le nombre réduit de paramètres rend en tout cas ce processus d'optimisation relativement léger, permettant d'envisager son utilisation sur des plateformes mobiles. Leur légèreté et leur capacité à refléter une connaissance botanique en fait ainsi des outils qui se prêtent bien à une extraction, à partir de l'image brute, d'un objet qui s'apparente à une forme foliaire, constituant une base pour une méthode de segmentation avec *a priori* de forme.

3.6.2 Pertinence de la description

L'autre atout de cette définition paramétrique est qu'une fois optimisé un modèle afin de s'ajuster à la forme concernée de la feuille, les valeurs présidant à sa construction fournissent une description pertinente de la forme représentée. Ces paramètres renvoient de fait à des grandeurs apparentes et mesurables qui permettent surtout de faire la différence entre les formes canoniques employées comme références dans les ouvrages de flores. Ils constituent ainsi quelque part une caractérisation de la position de la forme dans un univers jalonné par ces archétypes botaniques

Les paramètres des différents modèles sont donc porteurs non seulement d'un sens géométrique du fait de leur correspondance directe avec la forme qu'ils permettent de bâtir, mais aussi d'un lien avec une connaissance botanique par leur définition comme facteurs délimitant les différentes formes possibles. Cette corrélation entre construction et description n'est possible que parce que la définition des modèles est basée sur l'idée de recréer toutes les formes canoniques avec un nombre limité de paramètres. Le but est que les modèles puissent être à la fois suffisamment adaptables pour représenter toute la palette des formes existantes et dans le même temps assez rigides pour que différentes valeurs de paramètres ne puissent pas produire la même forme.

3.6.3 Interprétation sémantique

La rigidité des modèles permet également de s'assurer de la cohérence sémantique entre la valeur des paramètres et les grandeurs qu'ils sont censés représenter. Le fait que le paramètre de largeur du modèle polygonal par exemple soit le seul moyen d'élargir la forme permet au final d'être certain que sa valeur, une fois la forme ajustée à la feuille, correspondra bien à la largeur maximale de la feuille, telle qu'un observateur aurait pu la mesurer. Cette adéquation entre la modélisation idéale d'une forme foliaire et les grandeurs réelles qu'elle vise à capturer la rend directement interprétable à un haut niveau sémantique.

Le fait que les paramètres renvoient à ce qui permet de distinguer les termes botaniques et que leur valeur soit représentative de ces mesures laisse envisager une interprétation sémantique de la description fournie par le modèle, s'appuyant sur le vocabulaire descriptif des formes foliaires. La conversion des valeurs numériques en mots intelligibles pour un utilisateur (pas nécessairement en un seul mot puisque l'espace décrit par les modèles est, on l'a vu, plus continu que celui délimité par les termes de formes) est un objectif parfaitement accessible qui présenterait un grand intérêt à la fois du point de vue de l'interactivité du traitement que de la pédagogie par l'identification immédiate de formes visuelles avec des termes par ailleurs peu accessibles.

3.6.4 Reconstruction de formes

Toujours dans une optique d'interaction visuelle, la modélisation paramétrique permet par sa nature réversible de reproduire des formes à partir de valeurs numériques, pas nécessairement

extraites d'une image. Par exemple on peut envisager de régulariser quelque peu certains paramètres (suppression de la courbure de la feuille composée, symétrie de la base ou de l'apex) de façon à créer une visualisation pertinente de ce que capture le modèle au cours de l'analyse.

Il va également être possible de visualiser sous forme de modèles des formes apprises par ailleurs, comme la forme moyenne prise par une espèce, ou bien la forme correspondant le plus exactement à un terme botanique. Cette capacité à rendre visuellement compréhensibles de manière immédiate des notions qui peuvent être plus difficiles à expliciter offre l'opportunité d'une expérience enrichie pour l'utilisateur qui aura accès de façon intuitive au déroulement du traitement, à l'opposé d'une boîte noire qui se contenterait d'une réponse finale sans éléments d'interprétation en cours de route.

3.6.5 Cohérence de la connaissance

Enfin, l'intérêt plus global d'utiliser une même modélisation d'un bout à l'autre d'un processus d'identification des feuilles est de s'assurer de la cohérence de la connaissance *a priori* introduite au long du traitement. Il est primordial que les sources d'information ne soient pas conflictuelles si l'on veut pouvoir s'en servir pour guider et interpréter les résultats des différentes étapes intermédiaires.

Représenter les formes foliaires en utilisant le même outil pour la segmentation des feuilles, pour leur description morphologique et pour leur reconnaissance constitue une garantie que de tels conflits ne se présenteront pas. Les grandeurs modélisées renvoient constamment aux mêmes réalités objectives, et l'introduction de connaissances sur ces grandeurs pourra se faire sans risque d'introduire une erreur dommageable pour la performance du système.

S'appuyer à ce point sur des modèles pose en revanche une forte pression sur la robustesse de leur évaluation, et si tout est parfaitement cohérent lorsque le modèle est idéalement ajusté, l'erreur sera tout aussi cohérente du début à la fin si celui-ci est erroné. C'est là le double tranchant de l'introduction d'une modélisation à fort contenu sémantique, le risque venant du fait que toute interprétation à haut niveau fait fatalement l'hypothèse que la modélisation est en accord avec la réalité. La contrepartie positive en est néanmoins que toute erreur pourra vraisemblablement être remontée à une modélisation incorrecte dont la discordance sera de toute façon immédiatement visible et explicable.

La modélisation géométrique des diverses formes foliaires offre ainsi un intérêt non-négligeable, même si elle exige d'axer tout le processus d'identification sur l'évaluation de modèles. La somme de connaissance *a priori* transmise par ces représentations permet en tout cas d'envisager un traitement facilité des images de feuilles d'arbres, où les contraintes de formes seront précieuses pour extraire l'objet feuille d'images complexes, et où les descripteurs de haut niveau associés à la définition paramétrique des modèles ouvrent la possibilité d'une interprétation sémantique de l'ensemble de l'information extraite des images.

Chapitre 4

Segmentation des images de feuilles

Modèles déformables pour la segmentation d'images naturelles



Fagus sylvatica L.

Résumé du Chapitre

Ce chapitre détaille les méthodes de segmentation s'appuyant sur les modélisations de formes foliaires utilisées dans le cadre du traitement d'images contenant des feuilles en environnement naturel. Les modèles correspondant aux différents types de feuilles sont tout d'abord employés dans une approche de type modèle déformable, avec pour but de s'ajuster au mieux à la feuille contenue dans les images. Cette optimisation se fait en se basant essentiellement sur la couleur, et une partie importante du chapitre détaille les moyens mis au point pour modéliser la couleur de la feuille dans le cas d'images complexes et peu contrôlées. Enfin l'algorithme de segmentation par contours actifs basés sur les modèles de forme est présenté, et ses performances longuement évaluées et comparés à des méthodes de l'état de l'art en segmentation. *

La Section 4.4 présente des travaux menés en collaboration avec Julien MILLE, maître de conférences au LIRIS.

La Section 4.5 présente des résultats obtenus par Manuel GRAND-BROCHIER, post-doctorant au sein du projet.

*. Toutes les images de feuilles présentées dans ce chapitre proviennent d'appareils photos de smartphone et ont été prises dans un contexte d'utilisation réelle.

SEGMENTER les images de façon à isoler la feuille de son environnement constitue le point de départ de toute la chaîne de traitement, et donc une étape cruciale qui va conditionner le reste du processus, jusqu'à l'identification finale. Il s'agit d'une tâche extrêmement difficile, du fait de la complexité des images à traiter et du manque de certitudes sur leur contenu. La robustesse de cette phase va en tout cas être déterminante pour la performance de la méthode de reconnaissance dans son ensemble.

Les images de feuilles en environnement naturel posent un certain nombre de problèmes qui rendent la segmentation particulièrement délicate. Le fait que la couleur seule ne suffise généralement pas à distinguer la feuille d'intérêt de son environnement, qui peut lui aussi être en grande majorité végétal, rend caduc l'emploi de méthodes globales de seuillage ou de clustering. Cela pousse à utiliser un algorithme de segmentation plus complexe, prenant en compte différentes informations issues de l'image, ou laissant l'opportunité d'une interaction avec l'utilisateur.

Parmi les méthodes existantes, toutes ne s'appliquent pas à ce cadre complexe. Il est cependant intéressant d'étudier comment elles abordent le problème de façon générale, et comment les informations de couleur, de contour ou encore de texture sont généralement prises en compte. La question de l'introduction de connaissances *a priori* sur la forme est également un point déterminant, et les solutions déjà proposées pour la traiter sont multiples. Il convient donc de s'intéresser de plus près à l'état de l'art de la segmentation d'images naturelles pour déterminer les apports pertinents dans notre cadre précis.

4.1 Segmentation d'objets en environnement complexe

Les méthodes utilisées pour la segmentation d'objets dans les images numériques peuvent être regroupées en trois grandes catégories, esquissées dans le Chapitre 3 : les méthodes globales basées sur la valeur des pixels, les méthodes basées sur l'évolution de régions et les méthodes basées sur la détection de contours. La première classe de méthodes repose sur une vision globale de l'image pour prendre une décision, et se prête difficilement aux environnements naturels complexes où les variations très locales de couleur et de luminosité rendent ces algorithmes inopérants. Dans la suite seront donc essentiellement examinées les deux catégories restantes.

4.1.1 Estimation de gradient et détection de contours

La détection de contours dans les images correspond à la recherche de changements brutaux de l'intensité et se rapporte donc à l'étude de la dérivée de l'image. Dans une formulation classique, pour une image en niveaux de gris I , caractérisée donc par son intensité $I(x, y)$ en tout pixel, la dérivée de l'image correspond à un vecteur de dimension 2, le gradient de l'image :

$$\vec{G}(x, y) = \vec{\nabla}I(x, y) = \left(\frac{\partial I}{\partial x}(x, y); \frac{\partial I}{\partial y}(x, y) \right)$$

L'évaluation de ce gradient sur une image numérique qui n'est évidemment pas continue passe généralement par un filtrage de l'image, par des masques dont les plus communs sont ceux de Prewitt et de Sobel, tous deux de taille 3x3, ce qui offre une moindre précision mais une plus grande robustesse que des différences de pixels voisins comme le filtre de Roberts. De façon à limiter l'influence du bruit, ces masques sont souvent appliqués sur une l'image préalablement lissée par un filtrage, ce qui permettra de ne pas s'attacher à des petites structures de bruit qui génèrent un fort gradient. Les points de contour correspondent alors aux maxima du gradient de l'image, et une façon de trouver le contour d'un objet est de localiser les maxima de sa magnitude :

$$G(x, y) = |\vec{G}(x, y)| = \left(\frac{\partial I}{\partial x}(x, y)^2 + \frac{\partial I}{\partial y}(x, y)^2 \right)^{\frac{1}{2}}$$

Autrement dit, l'approche la plus simple est de localiser les passages à zéro de la dérivée du gradient (avec une dérivée seconde négative). C'est donc en observant la dérivée seconde de l'image, ou plus exactement le laplacien L que l'on peut détecter facilement les contours :

$$L(x, y) = \left(\frac{\partial^2 I}{\partial x^2}(x, y) + \frac{\partial^2 I}{\partial y^2}(x, y) \right)$$

La phase de lissage et de calcul du laplacien passant tous deux par une convolution de l'image, elles peuvent être combinées de façon à ne convoluer l'image qu'une seule fois par un laplacien de gaussienne, et chercher les passages par zéro dans le résultat de ce filtrage. Cette approche historique [MH80] constitue l'un des premiers essais de détection des contours d'objets, mais sa sensibilité au bruit et son uniformité à l'échelle de l'image ne garantit pas de se débarrasser de contours non pertinents ou de détecter des contours peu marqués.

Ces problèmes sont en revanche traités par des approches moins directes comme la détection de contour de Canny [Can86] qui, après filtrage par des dérivées de gaussiennes, cherche à supprimer les points qui ne sont pas des maxima du gradient dans la direction locale du gradient de l'image, puis relie les portions de contour de façon à obtenir des courbes fermées. Cette approche analytique est tout à fait efficace mais elle détecte tous les contours présents dans l'image, sans réelle distinction, ce qui n'est pas l'idéal pour détecter une forme en particulier.

4.1.2 Méthodes de segmentation basées région

A l'inverse, on peut voir la segmentation comme le découpage de l'image en régions homogènes selon un critère particulier, le plus souvent la couleur. Sans plus d'informations spatiales, on se rapproche de méthodes de *clustering* visant à isoler des classes de pixels en minimisant la variabilité intra-classe. Les méthodes les plus populaires étant les *k-means* et les *fuzzy-c-means* [Bez81].

Dès lors que l'on considère des régions connexes en plus d'être homogènes, les méthodes de segmentation ne peuvent plus être vues comme un simple problème de segmentation. L'approche

la plus simple est la croissance de région, qui, à partir de points sources vont ajouter aux régions les pixels voisins satisfaisant le critère d'homogénéité (relativement à un seuil Θ) jusqu'à ce que la croissance couvre toute l'image. Par exemple si l'on considère l'image I , un pixel P pourrait être ajouté à la région Ω_r d'intensité moyenne \bar{I}_{Ω_r} si :

$$\frac{1}{|\Omega_r| + 1} \int_{\Omega_r \cup \{P\}} (I(x, y) - \bar{I}_{\Omega_r})^2 dx dy < \Theta$$

Pour se passer de l'initialisation des points sources, l'approche *split & merge* [HP74] part de l'image complète, en subdivisant progressivement les régions non-homogènes selon une grille régulière, avant de fusionner les régions voisines suffisamment proches. Là encore l'homogénéité en termes d'intensité ou de couleur peut-être utilisé, en calculant une variance.

Les méthodes de lignes de partage des eaux (ou *watershed*) forment aussi une croissance de région, où l'image est cette fois considérée d'un point de vue topologique, l'idée étant de rechercher les bassins versants du relief défini par l'image [BL79, BM92]. La méthode la plus répandue est celle de la montée des eaux, qui simule une élévation de niveau à partir du point le plus bas de l'image. Les bassins d'abord indépendants, et étiquetés à l'aide de marqueurs, définissent lorsqu'ils se rejoignent les frontières des régions de l'image.

Une modélisation probabiliste de l'image et la représentation des pixels comme les sommets d'un graphe permet de définir la segmentation comme la recherche d'une coupe optimale de graphe [GPS89]. Ces méthodes dites de *graph-cuts* [BJ01, FH04, RKB04] permettent la représentation efficace de régions dans l'image, en offrant également la possibilité d'une interaction pour faciliter la segmentation.

Enfin, l'algorithme estimateur de densités *MeanShift* [Che95] bien qu'apparenté au *clustering*, peut être adapté à la segmentation basée région en prenant en compte la dimension spatiale dans la distribution des pixels à représenter [CM99].

Dans l'ensemble, ces méthodes utilisent peu l'information de transition qui peut s'avérer des plus intéressantes pour localiser précisément le contour de la feuille, et surtout rendent difficile l'inclusion de connaissance *a priori* sur la forme de l'objet dans le processus de la segmentation. Il semble donc plus pertinent de s'orienter vers des méthodes qui travaillent en déformant des objets qui laissent la possibilité des contraintes de forme dans leur évolution.

4.1.3 Modèles de contours actifs

La détection du contour d'un objet peut être vue comme la recherche d'une courbe fermée optimale, au sens d'une certaine mesure dépendant de l'image, mais pas uniquement. C'est le cas avec les contours actifs ou *snakes* [KWT88] et toutes les méthodes qui en découlent. Le contour est ici une courbe Γ qui, à partir d'une position initiale, va se déformer de façon à atteindre une position optimale qui minimise ce que l'on va nommer une **énergie**, fonction composée dans sa définition initiale de

trois termes :

$$E^*(\Gamma) = E_{interne}^*(\Gamma) + E_{attache}^*(\Gamma) + E_{externe}^*(\Gamma)$$

La première composante de cette fonction est une **énergie interne** qui agit comme contrainte sur la forme du contour et garantit d'obtenir une courbe relativement lisse. Elle fait appel à des dérivées du premier et du deuxième ordres de la courbe, jouant ainsi sur son élasticité et sur sa rigidité. C'est une énergie d'abord définie localement avant d'être intégrée le long du contour. Ainsi, si l'on effectue une paramétrisation de la courbe par son abscisse curviligne u , on peut noter son contour comme $\Gamma = \{\Gamma(u) = (x(u); y(u)), u \in [0, 1]\}$ et l'énergie interne s'exprime de la façon suivante :

$$E_{interne}(\Gamma) = \int_{\Gamma} E_{interne}^*(u) du = \int_0^1 \alpha \left(\frac{d\Gamma}{du} \right)^2 + \beta \left(\frac{d^2\Gamma}{du^2} \right)^2 du$$

Le deuxième terme est une **énergie** dite d'attache aux données de l'image, qui va pousser la courbe à se rapprocher des contours des objets. Elle repose donc sur l'information locale de l'image permettant de définir les contours, et donc en premier lieu le gradient. Une fois encore, on intègre un terme local sur la longueur du contour pour produire la fonction qui devra être minimisée :

$$E_{attache}(\Gamma) = \int_{\Gamma} E_{attache}^*(u) du = \int_0^1 -|\vec{G}(x(u), y(u))| ds$$

Pour s'affranchir de l'information de gradient, le terme d'attache peut être défini autrement, en prenant en compte uniquement l'homogénéité des régions définies par le contour [Ron94, CV01], se rapprochant plus des approches basées région. L'énergie est alors définie sur l'ensemble de la région $\Omega(\Gamma)$ délimitée par la courbe, ainsi que sur son complémentaire $\bar{\Omega}(\Gamma)$ le but étant que le contour sépare deux régions globalement homogènes. On s'appuie alors sur la distance des pixels à la couleur moyenne de chaque région :

$$E_{attache}^*(\Gamma) = \int_{\Omega(\Gamma)} \left(I(x, y) - \bar{I}_{\Omega(\Gamma)} \right)^2 dx dy + \int_{\bar{\Omega}(\Gamma)} \left(I(x, y) - \bar{I}_{\bar{\Omega}(\Gamma)} \right)^2 dx dy$$

Enfin le dernier terme est une **énergie externe** qui laisse la place pour la définition de contraintes extérieures optionnelles qui vont pouvoir influencer l'évolution et la forme du contour. Par exemple, il est possible d'utiliser ce terme pour forcer l'ancrage du contour à des points particuliers de l'image, dans le cadre éventuellement d'un processus partiellement interactif.

Concernant la représentation des contours, l'approche originelle se base sur une paramétrisation directe de la courbe par son abscisse curviligne [KWT88], mais d'autres représentations paramétriques peuvent être envisagées, en utilisant notamment des splines qui permettent en particulier de réduire le nombre de calculs et de rendre la contrainte de lissage implicite [BHU00, ZLMY⁺02]. On cherche dans tous les cas à faire varier Γ jusqu'à atteindre un minimum de la fonctionnelle d'énergie où l'équation

$\partial E^*(\Gamma)/\partial\Gamma = 0$ est vérifiée, ce qui se fait typiquement par une descente de gradient.

Cependant, une approche populaire passe par une formulation du contour par une **courbe de niveau** ou *level-set* [OS88, MSV95, CV01, SK08], où le contour est vu comme la ligne de niveau 0 d'une fonction auxiliaire. Une telle formulation implicite du problème tolère particulièrement bien les changements de topologie. Elle nécessite dans son évolution la résolution d'une équation aux dérivées partielles pour calculer la vitesse de déplacement de la courbe en chaque point, qui peut parfois s'avérer assez complexe.

4.1.4 Modèles déformables avec *a priori* de formes

La prise en compte de la connaissance *a priori* sur la forme de l'objet recherché au cours de l'évolution de formes que l'on déforme progressivement peut se faire de plusieurs manières. La définition de *templates* paramétriques déformables, permet d'introduire plus de contraintes en lien avec un objet bien précis, par exemple les parties du visage [YHC92]. Les contraintes que l'on peut faire porter sur un contour actif en comparaison sont bien moins précises et spécifiques. L'énergie à minimiser par le modèle peut contenir différents termes selon les caractéristiques de l'objet d'étude, et l'évolution se fait par descente de gradient sur les valeurs des paramètres.

Une autre solution consiste à considérer un contour défini par un ensemble réduit de points et d'en apprendre automatiquement l'agencement et la répartition pour produire un *active shape model* [CTCG95]. Ce modèle très spécifique peut ensuite être déformé pour s'adapter au contenu de l'image, en utilisant un ensemble de paramètres appris via l'étude statistique des formes, qui varient dans un domaine limité de façon à produire des formes acceptables.

Enfin, on peut représenter une forme par un ensemble assez rigide de triangles adjacents [Fel04] pouvant chacun être déformé. Ces modèles peuvent être eux aussi appris automatiquement, avec l'agencement des triangles et les variations qui leurs sont autorisées. Ces méthodes de modèles déformables sont assez flexibles pour s'adapter à un objet connu dans l'image, et elles conservent une importante rigidité qui leur permet par exemple de détecter les formes modélisées même lorsque la scène est encombrée ou que l'objet est occulté.

4.1.5 Adaptation aux images de feuilles en environnement naturel

Le cas des feuilles d'arbres se situe en quelque sorte à la croisée des problèmes classiques traités par les méthodes de segmentation. Les objets qu'elles constituent présentent en effet un problème à plusieurs facettes, qu'illustre la Figure 4.1 :

- Mal caractérisées par l'information de gradient, du fait de structures fortement contrastées à l'intérieur de la feuille (reflets, nervures, potentiellement plus marqués que le bord) elles se prêtent mal à une détection de contour, et un contour actif classique nécessiterait une initialisation très proche du bord pour être pertinent.

- ▶ Globalement trop peu homogènes du point de vue de la couleur, ou trop proches des couleurs de l'environnement, les approches basées sur un critère d'homogénéité de région risquent de produire une sur-segmentation, ou de déborder de l'objet d'intérêt
- ▶ Trop variables dans leurs formes, un modèle rigide pouvant surpasser les difficultés du contenu de l'image en contraignant fortement la forme obtenue se montrerait trop restrictif pour représenter l'ensemble des formes possibles.

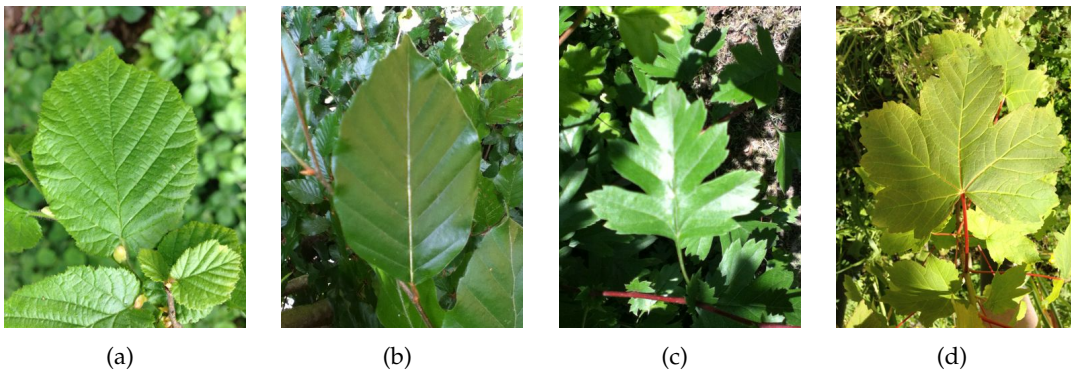


FIGURE 4.1 – Exemples de problèmes liées aux images naturelles de feuilles : fort gradient interne (a), non-homogénéité des couleurs (b), reflets marqués (c), proximité des couleurs du fond (d) et variabilité des formes

Pourtant, la connaissance *a priori* sur la forme est bel et bien présente, mais elle pêcherait du manque de flexibilité des méthodes évoquées ci-avant. Même si le choix des modèles déformables est visiblement pertinent, il semble peu réaliste d'espérer une représentation suffisamment adaptable pour à la fois capturer les grandes variations de forme globale et s'ajuster précisément au contour de la feuille, tout en ne prenant que des formes plausibles. L'idée qui découle de ce constat est alors d'effectuer la segmentation en deux temps, en commençant par s'appuyer sur la connaissance de la forme pour estimer la forme globale de la feuille, puis, dans un second temps, abandonner un peu de rigidité pour permettre à la forme d'aller localement s'adapter au contour de la feuille.

Décomposer le processus en deux étapes distinctes permet en outre de pouvoir focaliser à chacune d'elles sur l'information de l'image qui sera la plus pertinente. Pour évaluer la forme globale de la feuille, on va ainsi pouvoir se baser sur un critère de similarité de couleur qui, sans pour autant être uniformément vérifié, devrait permettre d'adapter un modèle de forme avec un regard assez global. Une fois connue l'allure générale, l'information de contour devient nettement plus pertinente puisque l'on ne cherche plus que des contours dans le voisinage d'une forme donnée. Une telle méthode de segmentation permettra ainsi de considérer les choses dans l'ordre, tout en introduisant la quantité de connaissance *a priori* nécessaire pour contourner certaines des difficultés posées par les images naturelles de feuilles.

4.2 Modélisation de la couleur et cartes de dissimilarité

Comme on l'a vu, les régions définies par les feuilles dans des images prises en environnement naturel, où l'éclairage peu maîtrisé vient s'ajouter à une irrégularité naturelle, sont rarement homogènes du point de vue de la couleur. Il est pourtant évident qu'il s'agit d'une information clé pour séparer la feuille d'intérêt et le second plan. Simplement, le critère à optimiser ne peut pas se contenter d'être celui de la distance à une couleur moyenne, comme c'est souvent le cas dans les méthodes de segmentation basées région.

4.2.1 Distribution des couleurs et critère d'homogénéité

Lorsqu'on observe la distribution des couleurs, il apparaît clairement qu'une seule couleur moyenne ne suffirait pas à en capturer la forme. Elle ne forme en effet jamais une distribution compacte, organisée uniformément autour d'un point de l'espace des couleurs, mais des formes bien plus étirées, recourbées comme l'illustre la Figure 4.2. De plus, si l'on compare la distribution des couleurs de l'arrière plan, on remarque évidemment une intersection non-négligeable. La conséquence en est que la distance d'un pixel à la couleur moyenne de la feuille ne constitue pas une mesure fiable de son appartenance à la feuille.

Pour s'en rendre compte, on peut produire une carte affichant pour chaque pixel (x, y) représenté par sa couleur $I_{RGB}(x, y) = (R(x, y); G(x, y); B(x, y))$ la distance $d_{\overline{RGB}}$ qui le sépare de la couleur moyenne des pixels situés dans la région $\Omega_{\mathcal{F}}$ correspondant à la feuille :

$$d_{\overline{RGB}}(x, y, \Omega_{\mathcal{F}}) = \left((I_{RGB}(x, y) - \bar{I}_{RGB}(\Omega_{\mathcal{F}}))^T (I_{RGB}(x, y) - \bar{I}_{RGB}(\Omega_{\mathcal{F}})) \right)^{\frac{1}{2}} \quad (4.1)$$

$$\bar{I}_{RGB}(\Omega_{\mathcal{F}}) = \frac{1}{|\Omega_{\mathcal{F}}|} \sum_{(x', y') \in \Omega_{\mathcal{F}}} I_{RGB}(x', y')$$

Sur une telle carte, les pixels homogènes à la feuille du point de vue de la couleur apparaissent en gris foncé, alors que les intensités plus claires soulignent les pixels distants, qui seraient considérés comme extérieurs à la région si elle devait être rigoureusement homogène. Pour qu'un critère d'homogénéité, ou plutôt une mesure de dissimilarité comme l'est $d_{\overline{RGB}}$ soit pertinente, il faut que sur cette carte, la feuille d'intérêt apparaisse intégralement sombre et que le reste de l'image se détache assez nettement. Dans le cas de la distance euclidienne à la couleur RGB moyenne, la Figure 4.2 montre bien que ce n'est que très rarement le cas.

L'espace RGB où la couleur est représentée par ses composantes brutes fortement corrélées n'est de toute façon pas l'idéal pour modéliser l'information de couleur. On préfère généralement se placer dans des espaces colorimétriques dont la définition permet de décorrélérer l'information de luminance de celle de chrominance. C'est le cas par exemple des espaces teinte-luminosité-saturation (HLS) et teinte-saturation-valeur (HSV), l'espace YUV, ou encore l'espace perceptuel CIE L*a*b*. La distance entre les couleurs dans ces espaces correspondent généralement mieux à ce qui est perçu par l'œil humain, ce qui les rend plus pertinentes pour mesurer l'homogénéité de la couleur.

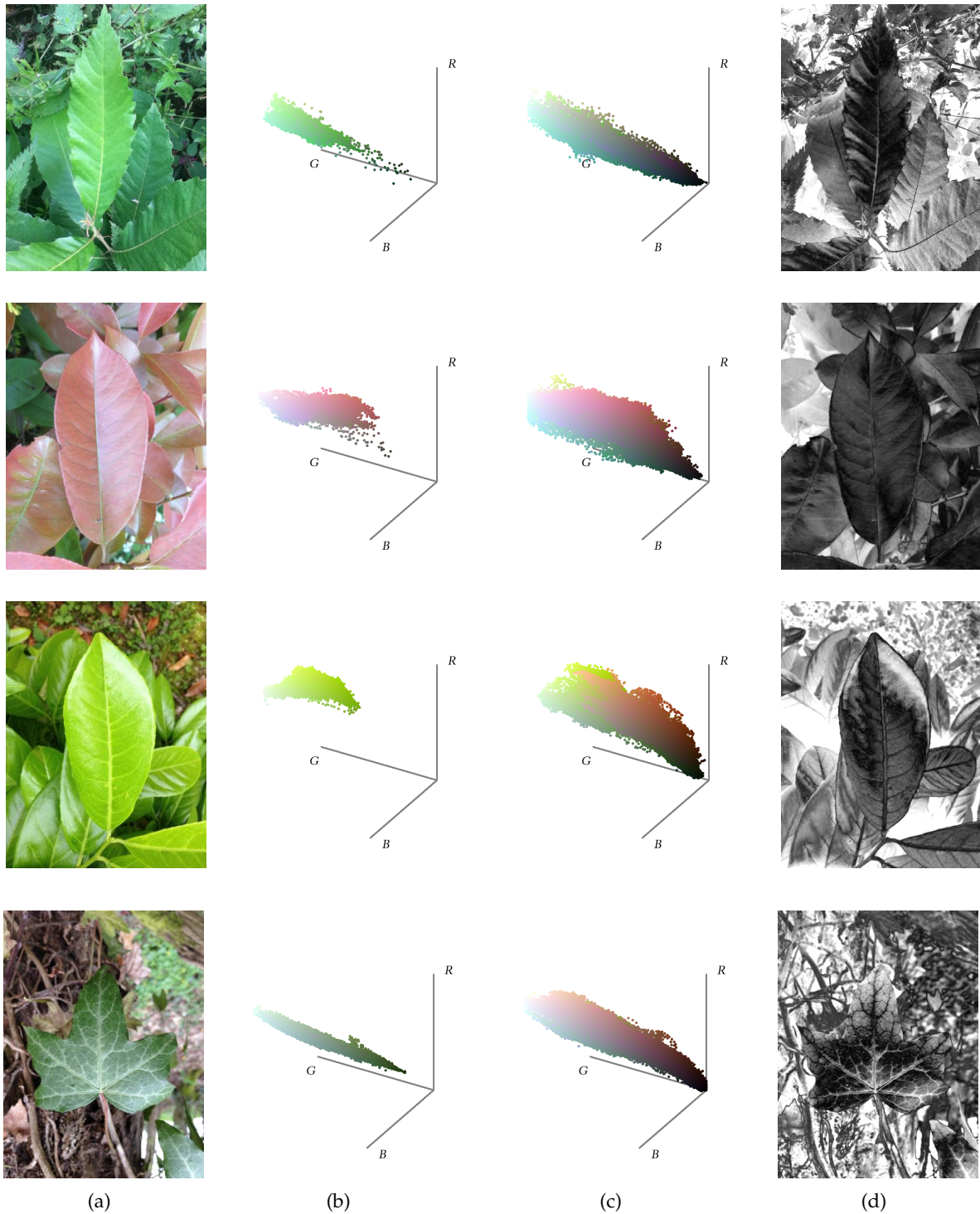


FIGURE 4.2 – Distribution des couleurs de la feuille (b) et du fond (c) dans l'espace RGB et carte de distance à la couleur moyenne de la feuille (d)

La distribution des couleurs de la région formée par la feuille apparaît effectivement plus compacte, notamment dans les espaces HSV et surtout $L^*a^*b^*$, ainsi qu'on le voit sur la Figure 4.3. Par conséquent, lorsque l'on calcule les cartes de distance, en prenant cette fois-ci la distance euclidienne du pixel dans l'espace concerné, par exemple $I_{Lab}(x, y)$ à la couleur moyenne $\bar{I}_{Lab}(\Omega_{\text{feuille}})$, elles sont alors visiblement meilleures, la feuille contenant surtout moins de pixels distants.

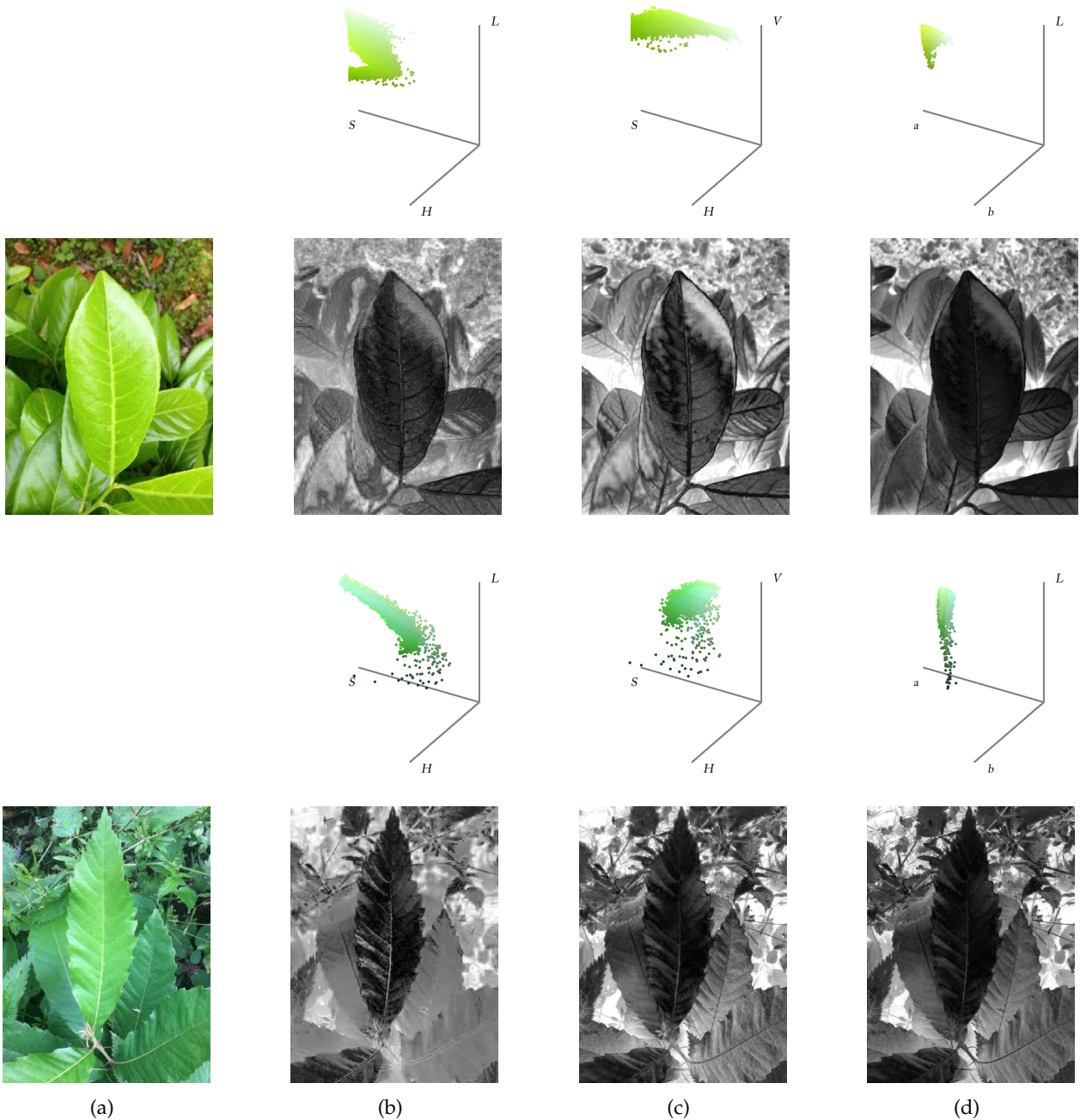


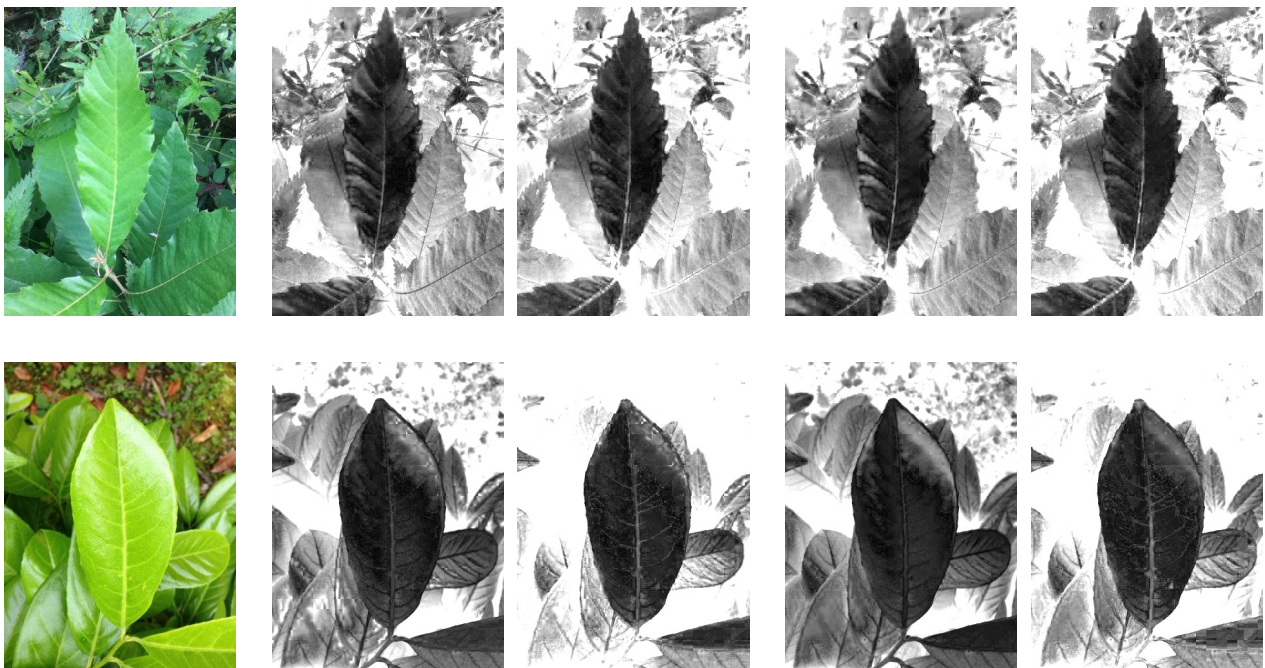
FIGURE 4.3 – Distribution des couleurs de la feuille et carte de distance à la couleur moyenne dans les espaces colorimétriques HLS (b), HSV (c) et $L^*a^*b^*$ (d)

L'homogénéité n'est pourtant pas parfaite, loin s'en faut, et même dans ces espaces plus adaptés, représenter la couleur de la région à segmenter par sa moyenne sera insuffisant pour établir le critère permettant de discerner l'ensemble des pixels de l'objet. La distribution peut être modélisée de manière plus satisfaisante en prenant en compte en plus de la moyenne la variabilité des couleurs selon les différentes dimensions de l'espace de couleur. On va ainsi pouvoir représenter la couleur de la feuille dans un espace \mathcal{C} non plus par un vecteur unique mais par un modèle gaussien calculé sur une région Ω composé d'un point moyen $\bar{I}_{\mathcal{C}}(\Omega)$ et d'une matrice de variance-covariance $\Sigma_{\mathcal{C}}(\Omega)$ témoignant de la forme de la distribution.

Le critère d'homogénéité pourra alors être défini comme une distance à ce modèle gaussien, de sorte que la variabilité des couleurs sur la surface de la feuille soit prise en compte. On peut ainsi utiliser une distance euclidienne normalisée où seule la variance selon chaque dimension est prise en compte, ou une distance de Mahalanobis qui va aussi considérer l'information de covariance. La distance euclidienne normalisée correspond en fait à une distance de Mahalanobis calculée à partir d'une matrice $\Sigma_{\mathcal{C}}(\Omega)$ diagonale :

$$d_{M,\mathcal{C}}(x,y,\Omega) = \left((I_{\mathcal{C}}(x,y) - \bar{I}_{\mathcal{C}}(\Omega))^T \Sigma_{\mathcal{C}}(\Omega)^{-1} (I_{\mathcal{C}}(x,y) - \bar{I}_{\mathcal{C}}(\Omega)) \right)^{\frac{1}{2}} \quad (4.2)$$

La distribution est dans les deux cas modélisée par un ellipsoïde dont les dimensions (et l'orientation dans le cas de la distance de Mahalanobis) sont données par la matrice de variance-covariance. Ce sont visiblement les espaces HSV et L*a*b* qui se prêtent le mieux à cette modélisation, au vu de la forme des distributions de couleur. De façon évidente, la feuille apparaît avec ce critère comme une région bien plus homogène, comme le montrent les cartes de distance de la Figure 4.4, puisque sa variabilité est quelque part prise en compte dans la mesure de son homogénéité.



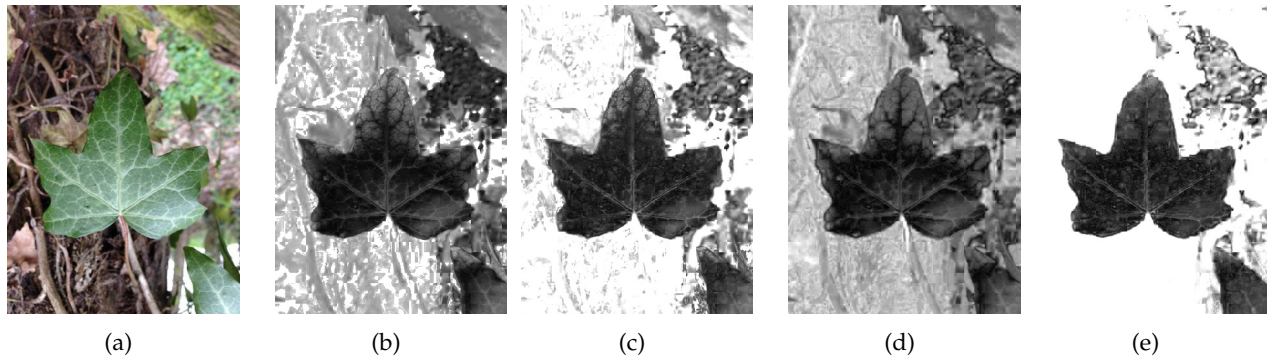


FIGURE 4.4 – Cartes de distances à une modélisation gaussienne des couleurs de la feuille : distances normalisée et de Mahalanobis respectivement dans les espaces HSV (b) et (c) et $L^*a^*b^*$ (d) et (e)

Cela pose d'ailleurs quelques problèmes, puisque toute région va alors pouvoir être considérée comme homogène du moment que sa distribution de couleurs dans l'espace choisi est suffisamment en adéquation avec une modélisation gaussienne. Ainsi, si l'on se place dans un cadre typique de croissance de région où l'homogénéité est constamment réévaluée au fur et à mesure que des pixels sont inclus dans la région, il va être très difficile de faire la part des choses avec un critère d'homogénéité qui, bien que visiblement adapté, risquerait de se montrer trop tolérant.

Plutôt que de miser sur une réévaluation constante de l'homogénéité, il nous a alors semblé bien plus pertinent de fonder la segmentation sur la similarité des pixels aux couleurs d'une région initiale, notée Ω_0 , justement pour pouvoir inclure la variabilité naturelle des couleurs de la feuille sans prendre le risque de pouvoir considérer comme homogène toute région trop variable. Cette région initiale doit bien évidemment se trouver à l'intérieur de la feuille de façon à ce que les cartes de distances produites à partir de la modélisation des couleurs, calculées une fois pour toutes, correspondent à la dissimilarité des pixels avec les couleurs de la feuille.

La définition de Ω_0 va dès lors avoir une influence considérable sur le calcul de ces cartes de dissimilarité, et donc sur le reste de la segmentation. L'idéal est naturellement que la distribution des couleurs dans Ω_0 soit identique à celle de la feuille elle-même Ω_{feuille} . Ce sera rarement le cas, et il faut donc que la modélisation des couleurs soit la plus robuste possible vis à vis des écarts entre ces distributions. C'est là que la représentation par une gaussienne montre ses limites, étant par nature limitée aux couleurs correspondant à l'ellipsoïde et incapable d'extrapoler au-delà. La Figure 4.5 montre que la modélisation par une gaussienne peine à considérer comme proches des pixels absents de la zone initiale, ici définie manuellement.



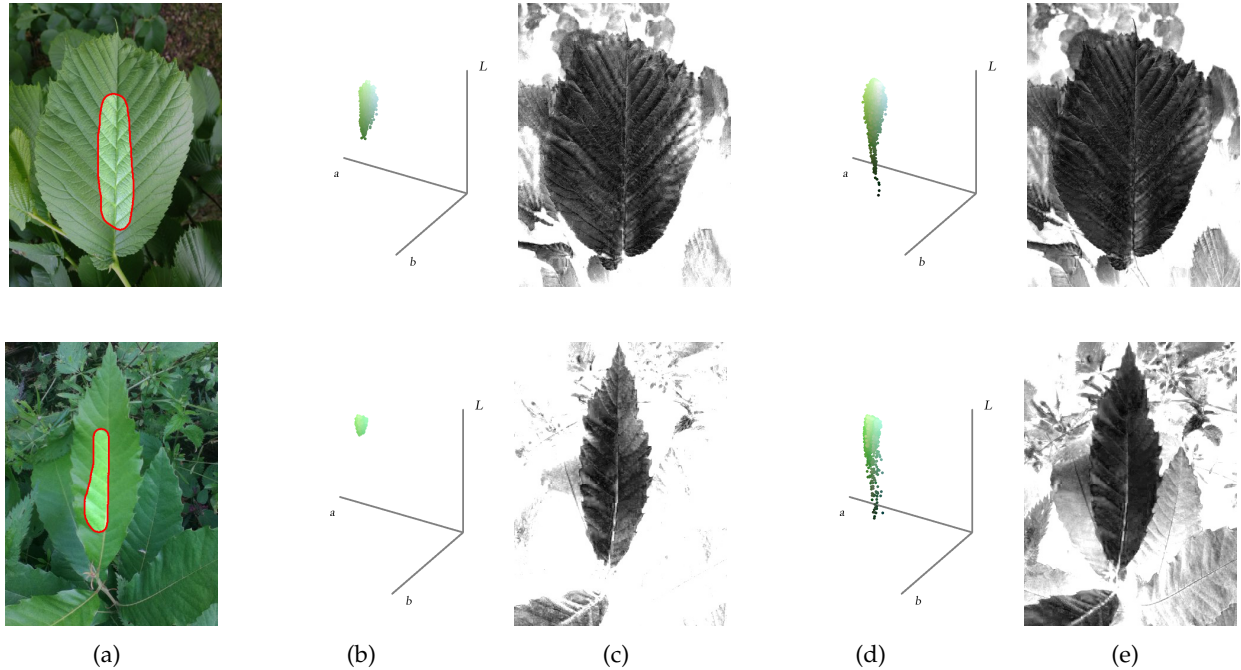


FIGURE 4.5 – Comparaison des distributions de couleurs dans l'espace $L^*a^*b^*$ et des dissimilarités à une modélisation gaussienne des couleurs respectivement de la région initiale (b) et (c) et de la feuille entière (d) et (e)

L'idée va donc être de produire une représentation de la couleur de la feuille telle que la dissimilarité des pixels puisse être approximativement la même selon qu'on calcule cette représentation sur Ω_0 ou sur Ω_{ent} . Puisque les distributions respectives des couleurs ne sauraient être les mêmes, il faut trouver un moyen d'extrapoler la distribution de Ω_{ent} à partir de celle de Ω_0 . En observant la répartition des couleurs dans les différents espaces, il nous a semblé que c'était dans l'espace $L^*a^*b^*$ qu'une telle généralisation était la plus envisageable.

La distribution des couleurs de la feuille dans cet espace semble en effet former de façon quasi systématique un cylindre assez étroit dont la direction principale semble suivre celle de la luminosité, avec une inclinaison plus ou moins visible. En particulier, les couleurs très claires des reflets, ou celles plus sombres des ombres propres de la surface de la feuille sont généralement en continuité de ce cylindre, de manière sensiblement rectiligne. La direction de l'axe du cylindre semble ainsi caractéristique de la distribution des couleurs, et permet à partir d'une certaine plage de luminosité de prévoir la couleur dans des luminosités plus faibles ou plus élevées.

Pour mettre en pratique cette idée, nous avons choisi d'effectuer une régression linéaire de la distribution des couleurs en choisissant la luminosité L comme variable explicative. A partir des couleurs contenues dans une région Ω , que l'on notera $(I_{Lab}(x, y))_{(x; y) \in \Omega}$ on obtient des coefficients réel a permettant la prédiction d'une couleur $\hat{I}_{Lab}(L, \Omega) = (L; \hat{a}(L, \Omega); \hat{b}(L, \Omega))$ à partir d'une valeur de la luminosité comme fonction affine :

$$\hat{a}(L, \Omega) = a_0(\Omega) + a_1(\Omega)L \quad \hat{b}(L, \Omega) = b_0(\Omega) + b_1(\Omega)L \quad (4.3)$$

On peut alors définir la dissimilarité d'un pixel (x, y) à ce modèle global de couleur comme la distance euclidienne entre sa couleur $I_{Lab}(x, y)$ et la couleur prédite par le modèle avec la valeur de la variable explicative en ce pixel $L(x, y)$:

$$d_{LinReg, Lab}(x, y, \Omega) = \left((I_{Lab}(x, y) - \hat{I}_{Lab}(L(x, y), \Omega))^T (I_{Lab}(x, y) - \hat{I}_{Lab}(L(x, y), \Omega)) \right)^{\frac{1}{2}} \quad (4.4)$$

Cette mesure permet, lorsque la régression linéaire est effectuée sur Ω_0 , de quantifier la dissimilarité des couleurs de l'image à un modèle des couleurs de la feuille, y compris en dehors de la plage de luminosité comprise dans Ω_0 . Celle-ci doit malgré tout être suffisamment large pour que l'estimation de l'axe du cylindre soit fiable, mais le modèle offre alors des capacités d'extrapolation qui manquent à la modélisation gaussienne. La Figure 4.6 met en évidence cet apport en comparant les cartes de dissimilarité obtenues à partir d'une région initiale et de la feuille complète.

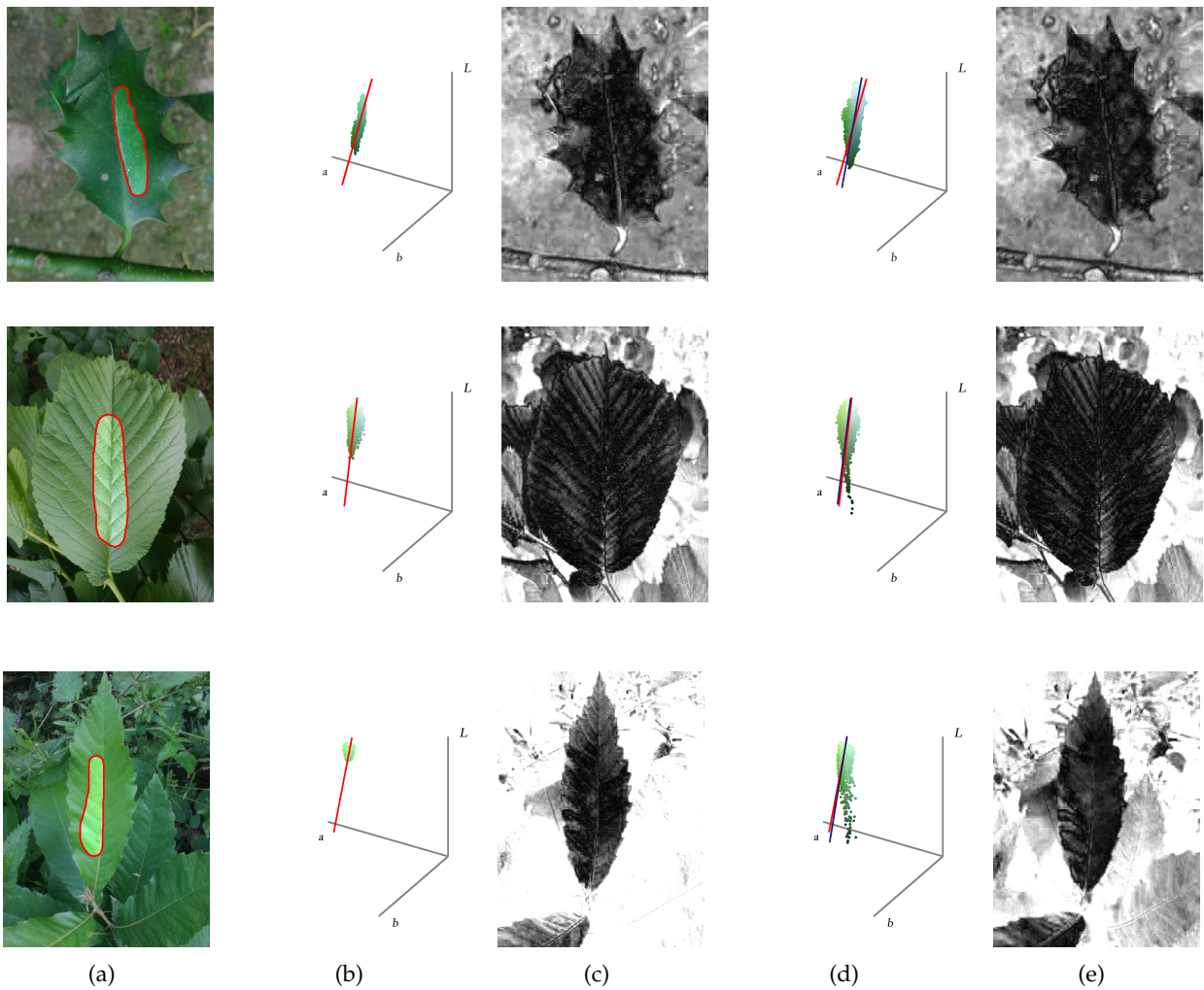


FIGURE 4.6 – Regression linéaire dans l'espace $L^*a^*b^*$ et cartes de dissimilarité au modèle linéaire des couleurs respectivement de la région initiale (b) et (c) et de la feuille entière (d) et (e)

On voit bien qu'en plus de distinguer efficacement la feuille de son environnement, cette modélisation s'avère plus stable relativement à la région d'initialisation Ω_0 . Cela est notamment dû aux propriétés de l'espace $L^*a^*b^*$, et il est intéressant de remarquer qu'une même approche par régression linéaire ne donne pas des prédictions aussi stables de la distribution des couleurs de la feuille dans les autres espaces. La Figure 4.7 montre que l'espace $L^*a^*b^*$ est le plus approprié pour cette approche, et donc pour permettre d'extrapoler des couleurs cohérentes avec la distribution mais en dehors d'une zone initiale.

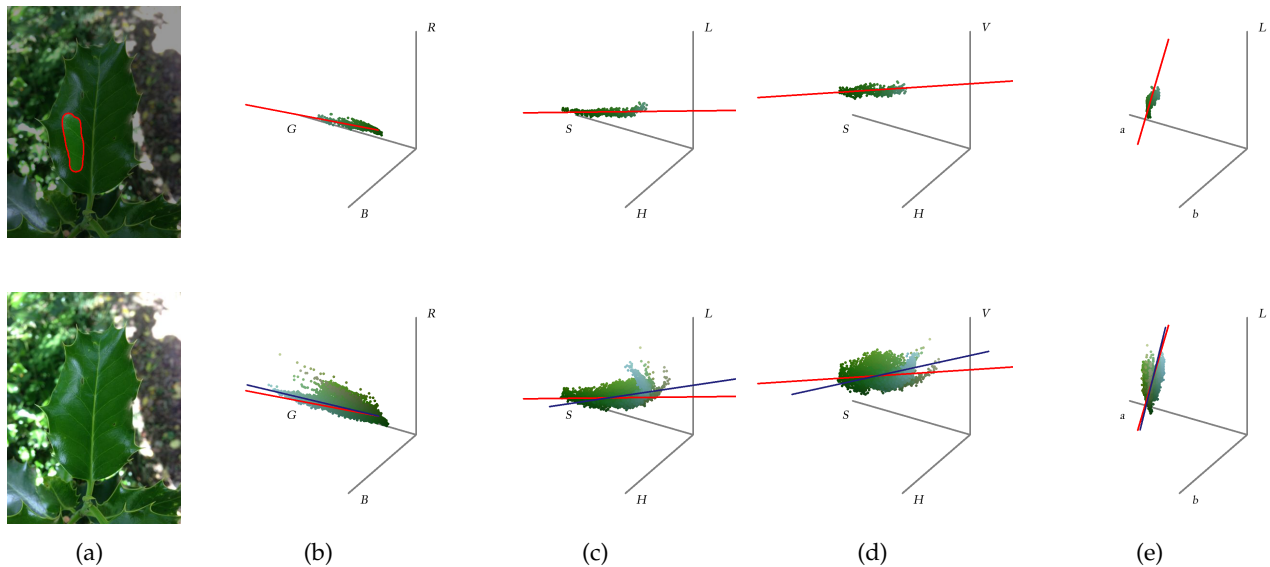


FIGURE 4.7 – Ecart entre les régressions linéaires des couleurs d'une région initiale (a) et de la feuille entière dans les espaces RGB (a), HLS (b), HSV (c) et L^*a^*b (d)

Cette modélisation de la couleur possède les caractéristiques à double tranchant d'une généralisation : étant capable de considérer comme proches des pixels appartenant à la feuille mais aux couleurs éloignées de celle de la zone initiale, elle risque dans le même temps d'inclure des pixels extérieurs à la feuille dont la couleur correspondrait au modèle. Il s'agit de toute manière d'un point de vue global sur l'image, qui reste par exemple incapable de discerner la feuille d'intérêt d'autres feuilles de mêmes couleurs situées à l'arrière plan. Et si elle s'adapte bien aux variations de luminance sur la surface de la feuille, elle est très sensible aux changements locaux de chrominance qui peuvent survenir.

4.2.2 Variations locales de la couleur

La couleur à l'intérieur d'une feuille peut varier. Les reflets, la nervation, les ombres ou les taches peuvent créer des changements de couleur brutaux, mais la principale variation le long de la surface de la feuille provient de son orientation et des différences d'illumination et constitue un changement progressif. Cette variation légère fait aussi bien intervenir la luminance que la chrominance, et renvoie à une réalité : si la feuille peut-être considérée comme localement homogène avec un critère approprié, elle n'est pour autant jamais véritablement plate.

Un modèle global de la couleur comme celui introduit ci-dessus pourra sans doute reconnaître ces couleurs variables comme couleurs de la feuille, mais il échouera forcément à capturer le fait que la feuille ne puisse présenter une couleur spécifique que dans une zone donnée de l'image. Et si cette couleur se retrouve dans une autre partie de l'image, peut-être qu'elle ne doit pas être considérée comme une couleur de la feuille ici. C'est notamment le cas si d'autres feuilles de la même couleur mais orientées différemment se trouvent à l'arrière plan. Un modèle global acceptera leurs couleurs même si elles sont différentes de la couleur locale de la feuille d'intérêt comme les exemples de la Figure 4.8. On se rend bien compte qu'un modèle global rigide n'est pas suffisant, et qu'il est également nécessaire de disposer d'un modèle qui s'adapte aux variations locales progressives de la couleur.



FIGURE 4.8 – Exemples de feuilles unies aux variations de couleurs locales (a) et difficultés du modèle global par regression linéaire L^*a^*b à les traiter (b)

L'idée va alors être simplement de représenter la couleur de la feuille en tout point de l'image par une moyenne qui dépend de l'endroit où l'on se trouve, rendant compte de ce que devrait être la couleur de la feuille à cet endroit de l'image. Le but est alors de faire que cette moyenne s'adapte aux changements progressifs de l'image, de façon à éviter qu'un modèle qui soit valable à une extrémité de la feuille ne le soit plus du tout à l'autre bout. C'est une conception qui se rapprocherait d'un modèle probabiliste de l'image, cherchant à estimer de façon heuristique en tout pixel la couleur la plus probable sous l'hypothèse que l'on se trouve à l'intérieur de la feuille.

Pour évaluer ce modèle, il est nécessaire d'utiliser une approche qui permet de propager progressivement l'information de couleur, présentée dans l'Algorithme 1. Nous avons choisi de le faire en explorant l'image, à partir d'une région initiale Ω_0 contenue dans la feuille, et en essayant de répondre à la question : quelle serait la couleur la plus probable à cet endroit si le pixel considéré faisait partie de la feuille ? Il va alors falloir se déplacer dans le reste de l'image et regarder si les couleurs rencontrées peuvent correspondre à une évolution progressive de la couleur de la feuille.

L'information que l'on va propager est une moyenne locale attendue de la couleur. Pour rester cohérent avec l'approche précédente, nous nous plaçons toujours dans l'espace $L^*a^*b^*$, et à chaque pixel $(x; y)$, nous associons une moyenne locale $\bar{I}_{Lab}(x, y)$. Ces moyennes sont initialement inconnues, excepté à l'intérieur de Ω_0 où l'on associe à chaque pixel la moyenne d'un voisinage carré de taille K autour de lui $V_{K \times K}(x, y)$:

$$\forall (x; y) \in \Omega_0, \bar{I}_{Lab}(x, y) = \frac{1}{K^2} \sum_{(x'; y') \in V_{K \times K}(x, y)} I_{Lab}(x', y') \quad (4.5)$$

Ensuite, l'idée est simplement de propager, et d'actualiser si nécessaire, cette moyenne à travers l'image. Tant que la variation de couleur entre le pixel considéré et la couleur de la feuille attendue en ce pixel reste limitée, on peut considérer que la feuille change progressivement de couleur, et prendre cet écart en compte dans la couleur que l'on s'attend à trouver un peu plus loin. En revanche, si le changement est trop brutal, il ne s'agit certainement pas d'une variation naturelle de la couleur de la feuille, il ne faudrait alors pas changer l'idée de ce qui est attendu pour les pixels suivants.

On définit alors plusieurs paramètres sur lesquels va reposer la méthode. Le premier est un paramètre de **tolérance** τ qui représente en fait un seuil de distance, la distance maximale entre la couleur d'un pixel et la moyenne propagée pour qu'un pixel soit pris en compte dans l'actualisation de la moyenne locale. Le deuxième paramètre est une valeur d'**adaptativité** α entre 0 et 1 qui quantifie la rapidité avec laquelle la moyenne locale va s'adapter aux changements locaux.

La propagation se fait par le biais d'une file à priorité P dans laquelle on place les pixels restant à explorer. Initialement, elle contient l'ensemble des pixels de la région Ω_0 qui sont dans le même temps marqués comme appartenant à l'ensemble E des pixels explorés. L'idée est que cette file soit triée selon la distance des pixels à leur moyenne locale :

$$d_{Locale}(x, y) = \left((I_{Lab}(x, y) - \bar{I}_{Lab}(x, y))^T (I_{Lab}(x, y) - \bar{I}_{Lab}(x, y)) \right)^{\frac{1}{2}} \quad (4.6)$$

Le pixel en tête de la file est traité en premier, de telle sorte que les pixels sont examinés dans l'ordre de leur vraisemblance en termes de couleur. La moyenne locale en chaque pixel n'est actualisée qu'une seule fois ; on s'en assure en propageant la moyenne du pixel considéré uniquement à ses voisins non-encore explorés, qui seront à leur tour marqués comme explorés et positionnés dans la file, en fonction de leur distance à la moyenne locale.

En considérant le pixel courant $(x; y)$, la première étape est de rechercher les pixels non explorés parmi ses 8 voisins que l'on notera $V_8(x, y)$. Si la couleur du pixel $I_{Lab}(x, y)$ est suffisamment proche de sa moyenne locale attendue $\bar{I}_{Lab}(x, y)$, la couleur que l'on propage aux voisins non-explorés sera une somme pondérée de la couleur et de la moyenne. Sinon, le pixel n'est pas pris en compte, et on se contente de propager la moyenne courante :

$$\forall (x', y') \in V_8(x, y) \cap \bar{E}, \bar{I}_{Lab}(x', y') = \begin{cases} \alpha I_{Lab}(x, y) + (1 - \alpha) \bar{I}_{Lab}(x, y), & \text{si } d_{Locale}(x, y) < \tau \\ \bar{I}_{Lab}(x, y), & \text{sinon} \end{cases} \quad (4.7)$$

Le calcul de la couleur locale attendue détaillé dans l'Algorithme 1 nécessite d'évaluer dans le même temps la distance de tout pixel $(x; y)$ à sa moyenne locale $\bar{I}_{Lab}(x, y)$. Cette mesure que l'on

note donc $d_{Locale}(x, y)$ constitue de fait une caractérisation de la dissimilarité d'un pixel au couleurs propagées depuis une région initiale, et peut donc être utilisée pour construire des cartes de dissimilarité similaires à celles obtenues avec le modèle global de couleur, mais prenant cette fois-ci en compte uniquement la variation locale de couleur. Le déroulement de la propagation et la carte de dissimilarité résultante sont illustrés dans la Figure 4.9.

Algorithme 1 Calcul de la moyenne locale attendue

Entrée : I_{Lab} l'image en $L^*a^*b^*$; Ω_0 une région initiale

Sortie : \bar{I}_{Lab} la moyenne locale attendue; d_{Locale} la carte de dissimilarité

```

P ← ∅
for all (x, y) ∈ Ω0 do
   $\bar{I}_{Lab}(x, y) \leftarrow \bar{I}_{Lab}(V_{K \times K}(x, y))$ 
   $d_{Locale}(x, y) \leftarrow \left( (I_{Lab}(x, y) - \bar{I}_{Lab}(x, y))^T (I_{Lab}(x, y) - \bar{I}_{Lab}(x, y)) \right)^{\frac{1}{2}}$ 
  insert (P, (x; y),  $d_{Locale}(x, y)$ )
  E ← E ∪ {(x; y)}
end for
while P ≠ ∅ do
  (x; y) ← front(P)
  delete (P, (x; y))
  if rand(0, 1) > Δ then
    insert (P, (x; y), ∞)
  else
    for all (x', y') ∈ V8(x, y) ∩  $\bar{E}$  do
      if  $d_{Locale}(x, y) < \tau$  then
         $\bar{I}_{Lab}(x', y') \leftarrow \alpha I_{Lab}(x, y) + (1 - \alpha) \bar{I}_{Lab}(x, y)$ 
      else
         $\bar{I}_{Lab}(x', y') \leftarrow \alpha I_{Lab}(x, y)$ 
      end if
       $d_{Locale}(x', y') \leftarrow \left( (I_{Lab}(x', y') - \bar{I}_{Lab}(x', y'))^T (I_{Lab}(x', y') - \bar{I}_{Lab}(x', y')) \right)^{\frac{1}{2}}$ 
      insert (P, (x'; y'),  $d_{Locale}(x', y')$ )
      E ← E ∪ {(x'; y')}
    end for
  end if
end while
return  $\bar{I}_{Lab}$ 

```

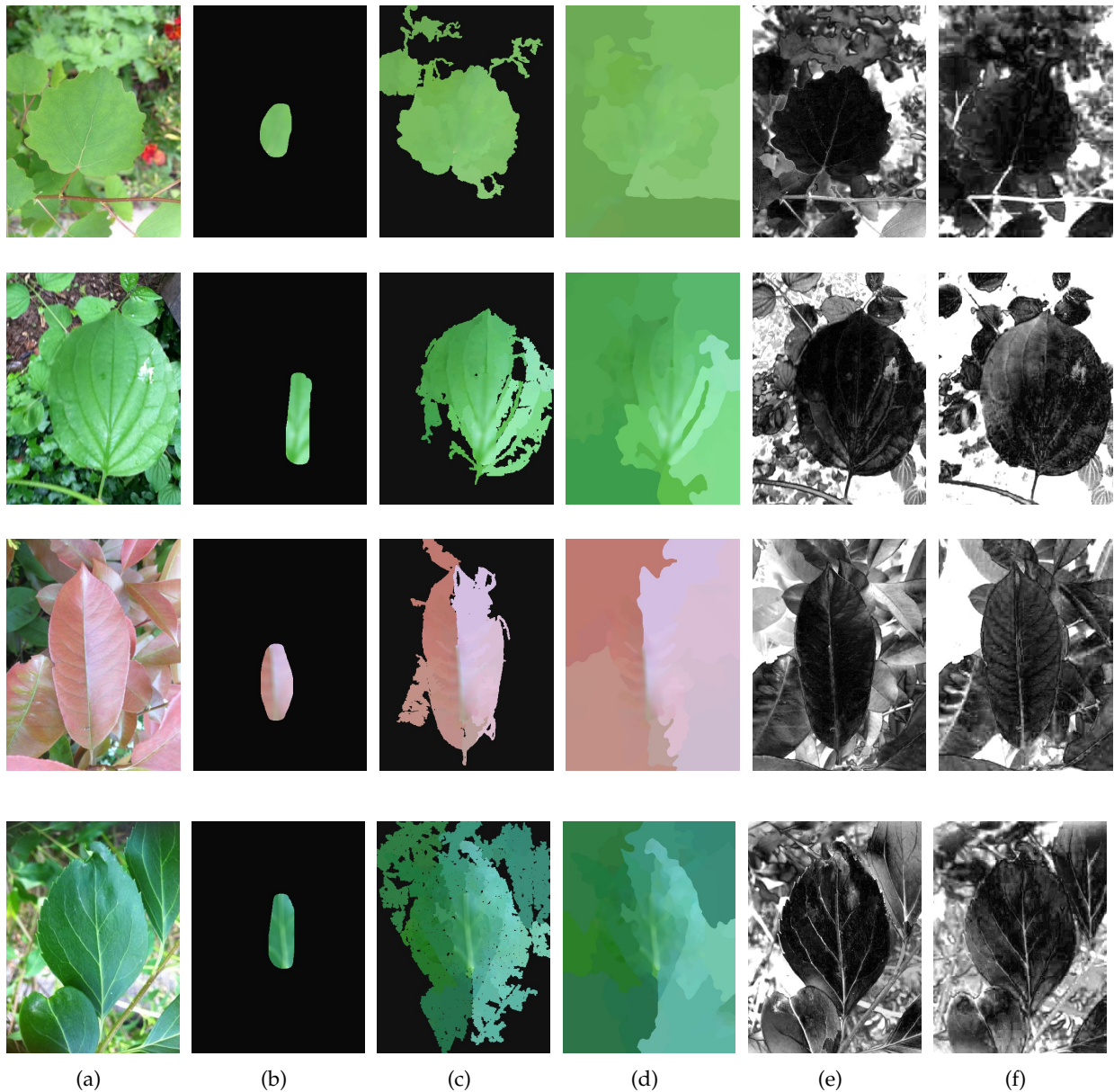


FIGURE 4.9 – Calcul de la dissimilarité à la moyenne locale : image de départ (a), région initiale (b), propagation (c), moyenne locale de couleur attendue (d) carte de dissimilarité locale (e) et comparaison avec la carte de dissimilarité au modèle global de couleur (f)

Le choix des valeurs des paramètres joue un rôle prépondérant dans le bon déroulement de l'algorithme. Le seuil de tolérance τ va être déterminant pour la pertinence de la moyenne. En effet, avec un seuil trop élevé, il sera impossible de faire la différence entre les couleurs de la feuille et celles du fond, et l'algorithme se limitera à propager la couleur locale de l'image (Figure 4.10 (f)). En revanche, une valeur de τ trop basse rendra le moindre saut de couleur un peu trop brusque de la feuille impossible à suivre pour la propagation, qui cessera dès lors d'adapter la moyenne et propagera une couleur uniforme (Figure 4.10 (b)).

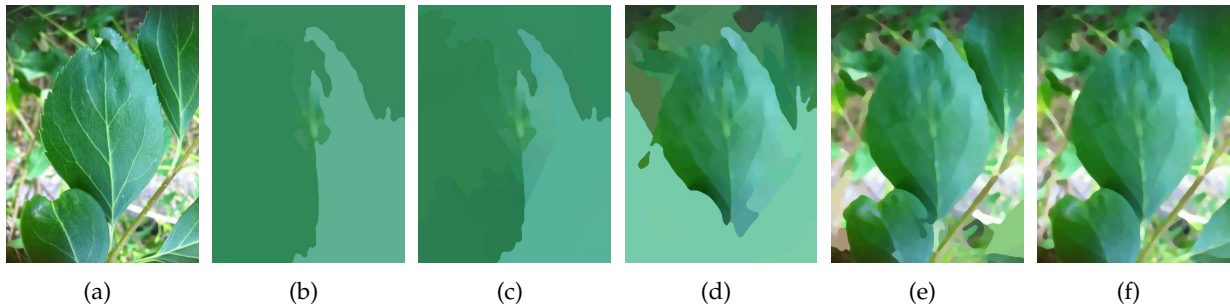


FIGURE 4.10 – Influence du paramètre de tolérance sur le calcul de la moyenne locale : $\tau = 0.001$ (b), $\tau = 0.01$ (c), $\tau = 0.02$ (d), $\tau = 0.05$ (e), et $\tau = 0.1$ (f) avec α fixé à 0.15

Concernant l'adaptativité α son influence est un peu plus fine à quantifier. Elle ne va jamais permettre de considérer comme proches les couleurs du fond, du moment que le seuil de tolérance est suffisamment bas. Ce qu'elle permet en revanche est de coller de plus ou moins près aux variations en surface de la feuille. Une adaptativité élevée va propager une moyenne qui suit fidèlement les variations locales de couleur (Figure 4.11 (f)) alors qu'une adaptativité faible sera plus prudente et conservera plus longtemps les couleurs initialement rencontrées (Figure 4.11 (b)).

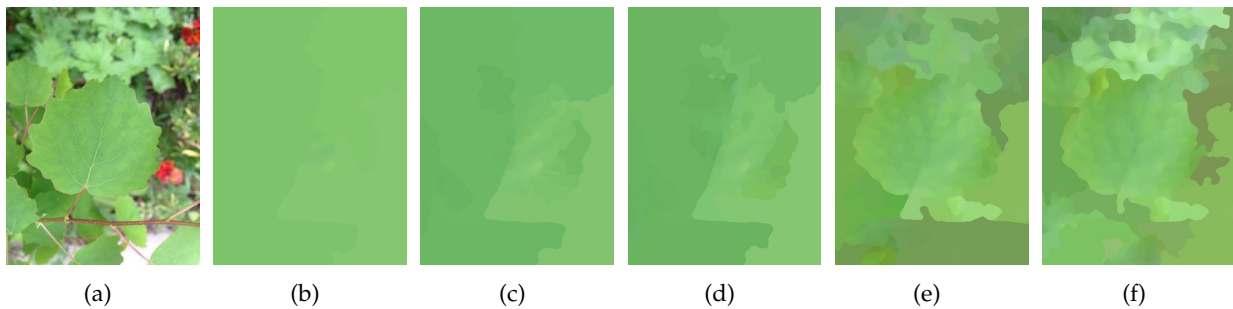


FIGURE 4.11 – Influence du paramètre d'adaptativité sur le calcul de la moyenne locale : $\alpha = 0$ (b), $\alpha = 0.1$ (c), $\alpha = 0.25$ (d), $\alpha = 0.5$ (e), et $\alpha = 0.75$ (f) avec τ fixé à 0.01

Le choix des deux valeurs n'est pas indépendant, et il faut le voir comme un compromis tolérance/adaptativité : soit on laisse un marge très faible aux couleurs par rapport à la moyenne, et dans ce cas, il faut être capable de s'adapter très vite pour ne pas être bloqué trop tôt, soit on laisse un peu de latitude aux couleurs, mais en propageant une information qui ne varie que lentement de façon à limiter le risque de sortir de la feuille. Les valeurs typiques employées sont de 0.01 pour la tolérance τ (les couleurs étant à valeurs entre 0 et 1) et de 0.15 pour l'adaptativité α , ce qui offre un compromis satisfaisant entre rigidité (pour ne pas risquer d'actualiser la couleur locale de la feuille avec des couleurs du fond) et flexibilité (pour ne pas être bloqué par des changements trop brusques de la couleur de la feuille).

La stratégie d'exploration décrite présente toutefois le risque de trouver une faille, un chemin de transition progressive de couleur entre la feuille et le fond, qui pourrait mener à considérer une large partie de l'environnement comme proche des couleurs de la feuille. C'est un problème qui peut le plus souvent être évité par un bon compromis tolérance/adaptativité qui limite l'existence de ces chemins,

mais cela fait partie des propriétés constitutives de la méthode de considérer comme proche de la feuille toute couleur atteignable par un chemin graduel à partir de l'intérieur de la feuille.

Pour contourner cet obstacle, nous rajoutons un paramètre de **directivité** Δ qui détermine à quel point l'exploration par plus proche couleur d'abord doit être privilégiée. Avec une valeur de Δ à 1, c'est toujours le pixel le plus proche en termes de couleur qui sera considéré en premier. Diminuer cette valeur rend le choix des pixels dans la file plus aléatoire, jusqu'à aller vers une propagation quasiment radiale. Ce petit ajustement va permettre de rendre les chemins entre la feuille et le fond plus difficiles à trouver, puisqu'il n'est plus garanti de suivre les chemins de distance minimale à coup sûr. En pratique, une directivité de 0.5 va suffire à avoir une propagation qui se fasse globalement en suivant les couleurs les plus proches mais en limitant la possibilité de trouver le chemin optimal vers l'extérieur.

Le calcul de la couleur locale attendue, et de la mesure de dissimilarité d_{Locale} qui en découle, permet de déterminer la cohérence des pixels de l'image avec une région initiale Ω_0 que l'on peut résumer à l'existence de chemins de variation progressive de couleur entre Ω_0 et le pixel considéré. La recherche de tels chemins n'est bien entendue ni exacte ni exhaustive, elle est simplement simulée par le mécanisme de propagation. Il n'en demeure pas moins que les pixels qui apparaîtront comme similaires à la feuille sont ceux pour lesquels la moyenne a su s'adapter, alors que ceux reconnus comme distants le sont à cause d'une discontinuité quelque part au cours de la propagation. Il est donc fort peu probable qu'un pixel distant appartienne à la feuille d'intérêt.

Le processus reste imparfait, étant donné le risque de voir des pixels de l'arrière-plan considérés comme similaires à la feuille. Il serait donc bon d'avoir un garde-fou qui puisse en améliorer la robustesse, y compris pour les images où l'on peut trouver une variation progressive de couleur de la feuille au fond.

4.2.3 Combinaisons de cartes de dissimilarité

Les approches précédentes constituent toutes deux des moyens adéquats de modéliser la couleur de la feuille, mais elles ont chacune leurs avantages et leurs inconvénients. Le modèle global de couleur par régression linéaire va bien reconnaître les couleurs de la feuille mais sans faire de grande différence entre la feuille d'intérêt et ses voisines de couleur proche, alors que le modèle local lui saura faire cette différence dès lors qu'il y a un changement brusque de couleur entre les deux feuilles. En revanche, il présente le risque de considérer le fond comme proche de la feuille si la transition n'est pas très marquée, quand la représentation globale permet de faire la différence entre des couleurs visiblement bien éloignées.

L'idéal consiste donc à combiner ces différents points de vue sur la couleur de l'image de façon à ne conserver que l'information la plus pertinente. Il faut cependant bien avoir en tête que les deux approches ne mesurent pas la même chose dans l'image si l'on ne souhaite garder que le meilleur de chacune. C'est un fait essentiel à prendre en compte pour les combiner de manière efficace, et une

combinaison uniforme comme une moyenne ou un minimum inclurait à la fois les qualités et les défauts et par conséquent ne serait pas vraiment suffisante.

Chaque représentation donne d'une certaine manière la probabilité selon son point de vue d'un pixel d'appartenir ou non à la feuille. Nous allons donc pouvoir raisonner en chaque pixel avec les probabilités des deux hypothèses complémentaires F et \bar{F} , respectivement que le pixel appartienne à la feuille et qu'il appartienne à l'arrière plan. La combinaison pourrait alors s'effectuer dans un cadre d'inférence Bayésienne, mais nous avons choisi d'utiliser la théorie évidentielle de Dempster-Shafer [Dem68, Sha76], qui peut être vue comme une généralisation de la théorie des probabilités. Elle présente notamment l'avantage de raisonner avec l'incertitude, et de modéliser explicitement l'absence de connaissance lorsque les sources ne fournissent aucune information tangible.

Lorsque l'on essaie d'analyser l'information apportée par chacune des méthodes, il semble qu'un pixel apparaissant comme proche dans le modèle de régression linéaire sera généralement un pixel de feuille mais pas nécessairement un de la feuille d'intérêt. Pour autant, certains pixels moyennement distants peuvent tout de même faire partie de la feuille s'ils se trouvent hors de la plage de luminosités de la zone initiale, et que la direction du cylindre a été mal évaluée. En revanche dès lors qu'on s'éloigne vraiment, il s'agit de couleurs très différentes en chrominance que l'on peut sans aucun doute exclure de celles possibles pour la feuille.

En ce qui concerne la distance à la moyenne locale estimée, la proximité d'un pixel va être un meilleur indicatif de son appartenance à la feuille, sans pour autant qu'on en soit tout à fait certain à cause de la possibilité de trouver un chemin continu et de variation graduelle de la feuille vers l'extérieur. Certaines parties de la feuille peuvent apparaître comme relativement lointaines si les changements se font trop brutaux et que la moyenne n'a pas le temps de s'adapter assez vite (c'est le cas des reflets par exemple). Enfin seuls les pixels très lointains, pour lesquels la moyenne espérée a en fait été propagée directement depuis la feuille, ne vont certainement pas en faire partie puisqu'aucun chemin n'a pu être trouvé pour les atteindre de façon progressive.

Intuitivement, on voit que les deux approches placent de la confiance différemment. L'approche globale reste mesurée quant à ce qui appartient à la feuille, mais catégorique pour ce qui en est très éloigné, alors que l'approche locale va au contraire permettre d'apporter une forte certitude d'appartenance à la feuille mais une défiance plus modérée lorsque les pixels sont vus comme distants. Et pour les distances intermédiaires, chacune permet une incertitude légitime, qui constitue le manque d'information qu'on espère combler par l'autre approche.

Pour formaliser ces notions, on se place dans un problème binaire en chaque pixel où l'on va affecter des **masses de croyance** aux événements qui vont pouvoir être combinés en utilisant la règle de Dempster. L'univers des possibles est alors réduit à l'ensemble $\mathcal{F} = \{F, \bar{F}\}$ correspondant aux événements opposés "feuille" et "non-feuille". La théorie de Dempster-Shafer opère sur un cadre de détermination formé par les parties de l'univers, nous travaillerons donc dans la suite dans l'ensemble des parties :

$$\mathcal{P}(\mathcal{F}) = \{\{F\}, \{\bar{F}\}, \{F, \bar{F}\}\}$$

On cherche donc à déterminer la dissimilarité du pixel $(x; y)$ à la couleur de la feuille, en effectuant une combinaison crédibiliste de sa dissimilarité $d_{LinReg, Lab}(x, y, \Omega_0)$ au modèle de couleur global et de sa distance $d_{Locale}(x, y)$ à la moyenne locale espérée. Dans la suite les coordonnées (x, y) seront oubliées, et on notera ces mesures respectivement d_{LinReg} et d_{Locale} de façon à alléger les notations, et parce que la décision ne se fait qu'au niveau du pixel.

Les deux valeurs de dissimilarité vont permettre de définir des affectations de masses de croyance (*basic belief assignments*) aux éléments de $\mathcal{P}(\mathcal{F})$. L'ensemble $\{F\}$ représente l'hypothèse selon laquelle le pixel fait partie de la feuille, $\{\bar{F}\}$ qu'il appartienne à l'arrière-plan, et l'ensemble $\{F, \bar{F}\}$ correspond à l'incertitude.

Les masses de croyance m_{LinReg} et m_{Locale} définies sur $\mathcal{P}(\mathcal{F})$ vont dépendre respectivement des dissimilarités d_{LinReg} et d_{Locale} , de façon à correspondre à l'apport d'information intuitivement esquissée ci-dessus. Nous avons choisi de les calculer comme des fonctions affines par morceau de la dissimilarité, à valeur entre 0 et 1 :

$$\begin{aligned}
 m_{LinReg}\{F\} &= \max(\lambda_{F,1}d_{LinReg} + \lambda_{F,0}, 0) & m_{Locale}\{F\} &= \max(\mu_{F,1}d_{Locale} + \mu_{F,0}, 0) \\
 m_{LinReg}\{\bar{F}\} &= \min(\lambda_{\bar{F},1}d_{LinReg} + \alpha_{\bar{F},0}, 1) & m_{Locale}\{\bar{F}\} &= \min(\mu_{\bar{F},1}d_{Locale} + \mu_{\bar{F},0}, 1) \\
 m_{LinReg}\{F, \bar{F}\} &= 1 - m_{LinReg}\{F\} - m_{LinReg}\{\bar{F}\} & m_{Locale}\{F, \bar{F}\} &= 1 - m_{Locale}\{F\} - m_{Locale}\{\bar{F}\}
 \end{aligned} \tag{4.8}$$

Les coefficients $\lambda_{F,1}$, $\lambda_{F,0}$, $\lambda_{\bar{F},1}$, $\lambda_{\bar{F},0}$, $\mu_{F,1}$, $\mu_{F,0}$, $\mu_{\bar{F},1}$, et $\mu_{\bar{F},0}$ ont été fixés empiriquement, de façon justement à modéliser l'apport d'information de chaque méthode. On aurait cependant pu envisager de les apprendre de façon automatique. Les Figures 4.12 et 4.13 représentent les fonctions de masses utilisées en fonction de la dissimilarité globale ou locale respectivement.

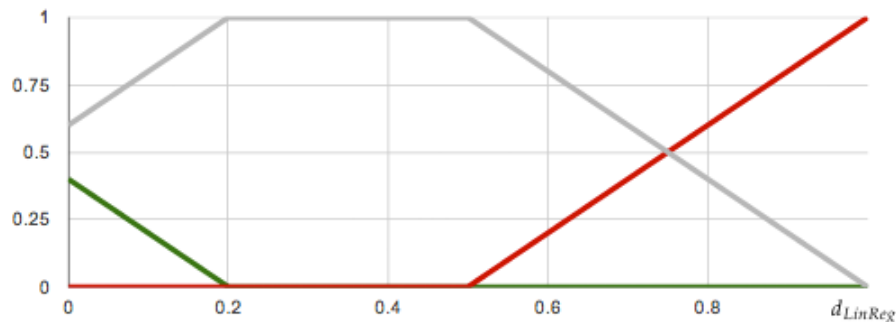


FIGURE 4.12 – Représentation des fonctions de masses utilisées en fonction de la dissimilarité au modèle global : ■ $m_{LinReg}\{F\}$, ■ $m_{LinReg}\{\bar{F}\}$, ■ $m_{LinReg}\{F, \bar{F}\}$

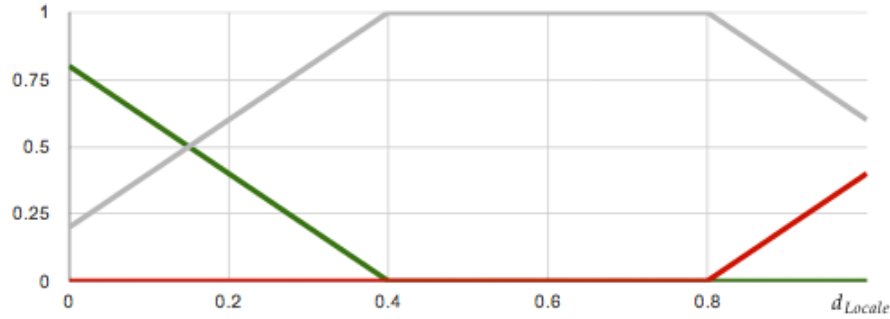


FIGURE 4.13 – Représentation des fonctions de masses utilisées en fonction de la dissimilarité à la moyenne locale : ■ $m_{Locale}\{F\}$, ■ $m_{Locale}\{\bar{F}\}$, ■ $m_{Locale}\{F, \bar{F}\}$

La règle de combinaison de Dempster offre un moyen de calculer la distribution jointe de masses de croyance $m_{LinReg \oplus Locale} = m_{LinReg} \oplus m_{Locale}$ résultant de la combinaison de l'information provenant des deux sources d'information. Elle utilise notamment une mesure K du conflit entre les deux distributions :

$$\begin{aligned} K &= \sum_{\substack{A, B \in \mathcal{P}(\mathcal{F}) \\ A \cap B = \emptyset}} m_{LinReg}(A) \cdot m_{Locale}(B) \\ &= m_{LinReg}\{F\}m_{Locale}\{\bar{F}\} + m_{LinReg}\{\bar{F}\}m_{Locale}\{F\} \end{aligned} \quad (4.9)$$

La règle de combinaison, que l'on peut voir comme une généralisation du théorème de Bayes, permet alors de mesurer l'accord entre les masses de croyance des différentes sources, et d'en éliminer le conflit en normalisant le résultat par $1 - K$:

$$\begin{aligned} m_{LinReg \oplus Locale}\{F\} &= \frac{1}{1 - K} \sum_{\substack{A, B \in \mathcal{P}(X) \\ A \cap B = \{F\}}} m_{LinReg}(A) \cdot m_{Locale}(B) \\ &= \frac{1}{1 - K} (m_{LinReg}\{F\}m_{Locale}\{F\} + m_{LinReg}\{F\}m_{Locale}\{F, \bar{F}\} + m_{LinReg}\{F, \bar{F}\}m_{Locale}\{F\}) \\ m_{LinReg \oplus Locale}\{\bar{F}\} &= \frac{1}{1 - K} \sum_{\substack{A, B \in \mathcal{P}(X) \\ A \cap B = \{\bar{F}\}}} m_{LinReg}(A) \cdot m_{Locale}(B) \\ &= \frac{1}{1 - K} (m_{LinReg}\{\bar{F}\}m_{Locale}\{\bar{F}\} + m_{LinReg}\{\bar{F}\}m_{Locale}\{F, \bar{F}\} + m_{LinReg}\{F, \bar{F}\}m_{Locale}\{\bar{F}\}) \\ m_{LinReg \oplus Locale}\{F, \bar{F}\} &= \frac{1}{1 - K} \sum_{\substack{A, B \in \mathcal{P}(X) \\ A \cap B = \{F, \bar{F}\}}} m_{LinReg}(A) \cdot m_{Locale}(B) \\ &= \frac{1}{1 - K} m_{LinReg}\{F, \bar{F}\}m_{Locale}\{F, \bar{F}\} \end{aligned} \quad (4.10)$$

La masse jointe permet de définir un couple de croyance et de plausibilité $\mathbf{bel}(F)$ et $\mathbf{pl}(F)$ que le pixel soit un pixel de la feuille. Leur définition est donnée par :

$$\begin{aligned} \mathbf{bel}_{LinReg\oplusLocale}(F) &= \sum_{A|A\subseteq\{F\}} m_{LinReg\oplusLocale}(A) = m_{LinReg\oplusLocale}\{F\} \\ \mathbf{pl}_{LinReg\oplusLocale}(F) &= \sum_{A|A\cap\{F\}\neq\emptyset} m_{LinReg\oplusLocale}(A) = m_{LinReg\oplusLocale}\{F\} + m_{LinReg\oplusLocale}\{F, \bar{F}\} \end{aligned} \quad (4.11)$$

La probabilité du pixel d'appartenir à la feuille se situe alors dans l'intervalle formé par la croyance et la plausibilité. Nous avons alors considéré comme approximation qu'elle se trouve juste entre les deux. Ce choix correspond dans ce cas à la probabilité pignistique, un compromis usuel dans le raisonnement avec les fonctions de croyance.

$$\begin{aligned} P_{\mathbf{bet}}(F) &= \frac{1}{2} \left(\mathbf{bel}_{LinReg\oplusLocale}(F) + \mathbf{pl}_{LinReg\oplusLocale}(F) \right) \\ &= m_{LinReg\oplusLocale}\{F\} + \frac{1}{2} m_{LinReg\oplusLocale}\{F, \bar{F}\} \end{aligned} \quad (4.12)$$

La probabilité pignistique permet alors de définir une nouvelle mesure de dissimilarité $d_{Croyance}$ qui correspond à la combinaison des deux dissimilarités initiales. A ce stade, pour obtenir une dissimilarité homogène sur l'image, on normalise les probabilités de chaque pixel par la probabilité maximale rencontrée dans l'image :

$$d_{Croyance}(x, y) = 1 - \frac{P_{\mathbf{bet}}(F)(x, y)}{\max_{(x', y') \in I} P_{\mathbf{bet}}(F)(x', y')} \quad (4.13)$$

Cette dissimilarité, par définition de la règle de combinaison, et vu que les fonctions utilisées pour définir les masses de base sont des fonctions affines, sera généralement une fonction rationnelle de degré 2 des deux dissimilarités d_{LinReg} et d_{Locale} , qui va permettre de réaliser une décision en prenant en compte plus d'information qu'une simple combinaison linéaire ou qu'une valeur minimale. La Figure 4.14 tente de montrer visuellement l'intérêt de cette combinaison crédibiliste.



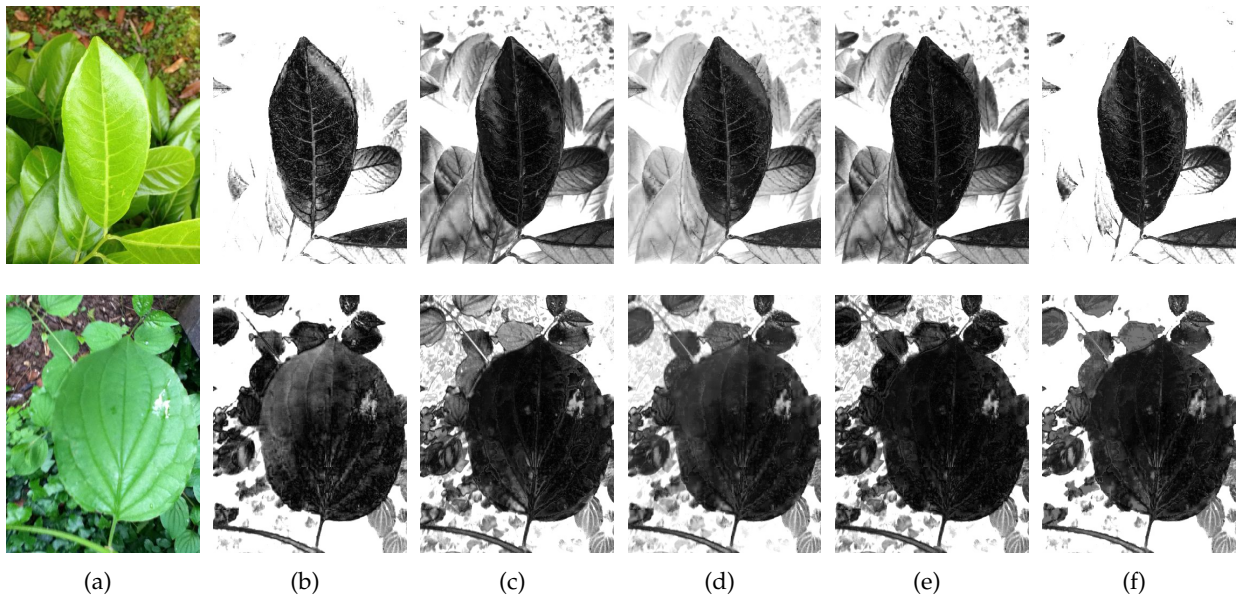


FIGURE 4.14 – Combinaison des cartes de dissimilarité issues du modèle global (b) et de la moyenne locale (c) à l’aide de la moyenne (d), du minimum (e) et des fonctions de croyance (f)

La combinaison des mesures de dissimilarité permet dans le cas où l’une des deux est erronée d’utiliser les informations provenant de l’autre pour prendre une meilleure décision. Le fait d’employer les fonctions de croyance permet une plus grande souplesse à ce niveau, en modélisant la fiabilité de l’information de dissimilarité selon la source, ce qui n’est évidemment pas le cas de combinaisons plus basiques.

Ce que l’on obtient au final est une carte de dissimilarité calculée une seule fois sur l’ensemble de l’image, et qui mesure pour chaque pixel sa cohérence du point de vue de la couleur avec une région initiale Ω_0 . Il s’agit d’une information assez riche qui prend en compte à la fois la couleur de la feuille considérée dans sa globalité et ses variations locales. Le fait d’évaluer le critère homogénéité une seule fois au début de la segmentation se justifie du point de vue du temps de calcul, mais aussi parce que la modélisation choisie est faite pour prendre en compte une certaine variabilité, et que le risque d’avoir une image entièrement homogène et de ne rien discriminer serait trop grand. Finalement, la région que l’on cherche à segmenter va alors être une région homogène à Ω_0 du point de vue de la couleur, et donc dont la dissimilarité totale est la plus faible possible.

4.3 Modèles paramétriques déformables de feuilles

La mesure de dissimilarité de couleur constitue une base pour la segmentation, résumant de façon condensée l’information de couleur de l’image, en rapport avec une région initiale qui doit être située à l’intérieur de la feuille. Cette donnée n’est pour autant pas suffisante pour parvenir à segmenter l’image. On voit nettement qu’un seuillage des cartes de dissimilarité, même une fois combinées, ne

produirait pas une segmentation satisfaisante. C'est là que la connaissance *a priori* sur la forme des feuilles va pouvoir être prise en compte pour, en s'appuyant sur l'information de couleur, localiser un forme globale de feuille dont les couleurs seront similaires avec celle de la région initiale. L'objectif est alors de réaliser une première phase de segmentation portant un regard assez global sur la forme pour isoler une région homogène en couleur qui, si elle ne correspond pas rigoureusement à la feuille, en esquisse les contours.

Les modèles de formes foliaires constituent un moyen flexible de représenter la connaissance sur la forme. Ils ont en plus l'avantage de reposer sur des paramètres numériques en nombre assez réduit, dont les variations permettent de créer la plupart des formes rencontrées dans la nature. L'idée est donc d'utiliser une approche de type modèle déformable avec pour but d'optimiser un modèle de forme sur l'image, en examinant les variations des différents paramètres et en se basant sur la dissimilarité pour ajuster au mieux leurs valeurs à la feuille dans l'image.

4.3.1 Minimisation d'énergie

La formulation employée pour résoudre le problème est très similaire à celle des contours actifs et autres modèles déformables où une région va subir une série de transformations de façon à minimiser une fonctionnelle d'énergie. De manière classique, cette énergie est décomposée en plusieurs termes représentant les différentes forces qui vont guider l'évolution de la forme, dont un terme d'attache qui définit la façon dont l'image est prise en compte.

Dans un modèle déformable basé région, une énergie d'attache classique [CV01] consiste à considérer la somme de l'hétérogénéité totale de la région Ω et de l'hétérogénéité totale de son complémentaire. Dans notre cas, on l'a vu, l'homogénéité de la région extérieure ne peut pas être une information pertinente étant donnée la complexité potentielle de l'arrière plan. Seul doit intervenir le contenu de la région d'intérêt, et plutôt que son homogénéité objective, nous préférons prendre en compte sa cohérence avec une région initiale à l'intérieur de la feuille, qui du point de vue de la couleur est estimée par la mesure de dissimilarité $d_{Croyance}$ (quoique d_{Locale} , d_{LinReg} , ou même $d_{M,Lab}$ constituent d'autres mesures elles-mêmes exploitables).

Cette dissimilarité est alors la principale force qui va guider l'évolution d'une région déformable. De la même façon que dans un modèle de contour actif basé région, on choisit comme principal critère à minimiser la dissimilarité totale de la région d'intérêt, ce qui permet de définir simplement une **énergie d'attache** sur la région Ω :

$$E_{couleur}(\Omega) = \int_{\Omega} d_{Croyance}(x, y, \Omega_0) dx dy \quad (4.14)$$

La définition de cette énergie basée sur la couleur a été pensée pour les images en environnement naturel. Dans le cas d'images plus simples, contenant une seule feuille sur fond uni, on peut se permettre d'utiliser une énergie plus simple. Nous avons vu en effet que la position par rapport à une valeur de luminosité constituait généralement un bon critère d'appartenance à la feuille permettant

l'application de seuillages, et que les méthodes de clustering fonctionnaient à merveille. Pour reprendre cette idée, tout en restant dans le même genre de formulation, on peut tout bonnement utiliser une formulation de la dissimilarité plus simple comme la distance de Mahalanobis (ce qui se rapprocherait de la mesure utilisée dans l'algorithme EM par exemple) :

$$E_{couleur}(\Omega) = \int_{\Omega} d_{M,Lab}(x, y, \Omega_0) dx dy \quad (4.15)$$

Cette énergie, prise telle quelle, n'est cependant pas tout à fait satisfaisante. En effet on voit bien que, la dissimilarité étant à valeur positive, la région minimisant cette énergie sera trivialement la région vide. Le fait de considérer la région complémentaire est ce qui permet de contourner ce problème dans la formulation classique, puisque rétrécir la région revient à élargir son complémentaire et potentiellement augmenter son hétérogénéité et donc l'énergie totale.

Pour contrecarrer cette tendance de l'énergie ainsi définie à réduire la région à sa plus simple expression, il faut rajouter une force qui pousse au contraire la région à s'étendre au maximum, c'est à dire à maximiser son aire. Cette **force ballon** va en quelque sorte combattre l'énergie d'attache jusqu'à atteindre un compromis, un équilibre au delà duquel elle ne pourra plus pousser la région à grandir. Elle est donc basée sur l'aire de la région, avec un signe négatif pour que la force s'exerce bien dans le bon sens :

$$E_{ballon}(\Omega) = -\omega_{ballon} \int_{\Omega} dx dy \quad (4.16)$$

Le poids ω_{ballon} est alors le paramètre qui détermine la force du ballon. On peut également le voir autrement, en rapport direct avec la mesure de dissimilarité. En effet si l'on additionne ces deux énergies, la grandeur que l'on somme sur la région est une différence entre la dissimilarité et le ballon. On peut donc voir ce poids comme la **valeur maximale d_{max} de dissimilarité** pour qu'un pixel puisse être ajouté à la région sans faire augmenter l'énergie. L'énergie d'attache à minimiser peut alors s'écrire :

$$\begin{aligned} E_{attache}(\Omega) &= E_{couleur}(\Omega) + E_{ballon}(\Omega) \\ &= \int_{\Omega} (d_{Croyance}(x, y, \Omega_0) - d_{max}) dx dy \end{aligned} \quad (4.17)$$

Cette formulation est intéressante, puisqu'elle renvoie à une idée de seuillage des cartes de dissimilarité, mais où le seuil d_{max} n'est pas appliqué de façon rigide. Il est en fait pris en compte implicitement dans la croissance d'une région ce qui permet un regard global et une certaine tolérance. Une région va ainsi pouvoir contenir certains pixels au delà du seuil si elle en contient par ailleurs d'autres dont la dissimilarité est très faible. Le résultat attendu en minimisant une énergie ainsi définie et de produire une région aussi grande que possible qui contienne au final le moins possible de pixels dissimilaires aux couleurs de la feuilles.

4.3.2 Un polygone actif pour les feuilles simples et palmatilobées

La région qui va devoir minimiser l'énergie définie ci-dessus peut être quelconque. En particulier on peut chercher à optimiser un modèle de forme de feuille simple eu égard à cette fonction. Trouver le plus grand polygone représentatif d'une forme globale de feuille qui s'adapte le mieux à la carte de dissimilarité de couleur permettrait d'avoir une première bonne idée de la région définie par la feuille dans l'image tout en s'assurant que la forme de cette région soit plausible.

En pratique, le modèle de feuille simple introduit dans le Chapitre 4 est un polygone complexe Π formé d'une superposition symétrique de polygones simples. Il repose à la fois sur la donnée de deux points B et A et sur un ensemble de paramètres que l'on peut décomposer en $\mathbf{p}_{Polygone} = (l; p; L; \alpha_B; \alpha_A)$ les paramètres permettant la construction des polygones simples et n_L et $\mathbf{p}_{Lobes} = (h_0; \theta_0; \dots; h_{n_L}; \theta_{n_L})$ les paramètres déterminant l'agencement des différents polygones formant les lobes. On peut alors écrire :

$$\begin{aligned} \Pi &= \Pi(B, A, \mathbf{p}_{Polygone}, \mathbf{p}_{Lobes}) \\ &= \bigcup_{l=-n_L}^{n_L} \Pi_l(B_l(B, A, \mathbf{p}_{Lobes}), A_l(B, A, \mathbf{p}_{Lobes}), \mathbf{p}_{Polygone}) \end{aligned} \quad (4.18)$$

Le but devient alors de trouver le polygone optimal Π^* qui minimise l'énergie $E_{attache}(\Pi)$ en cherchant à optimiser ses différents paramètres de construction de telle sorte que l'on obtienne en fait $\Pi^* = \Pi(B^*, A^*, \mathbf{p}_{Polygone}^*, n_L, \mathbf{p}_{Lobes}^*)$. Cette optimisation va donc pouvoir se faire en considérant les variations des différents paramètres de constructions et en cherchant à les faire varier de façon à diminuer l'énergie de la région Π .

Pour détailler la mise en place de ce processus, nous nous placerons dans un premier temps dans le cas d'une feuille simple non-palmatilobée, pour laquelle $n_L = 0$ et $\mathbf{p}_{Lobes} = (1; 0)$ sont constants et où l'on peut donc écrire $\Pi = \Pi_0(B, A, \mathbf{p}_{Polygone})$. Ce cadre restreint ne nécessite ainsi pas d'optimiser les paramètres \mathbf{p}_{Lobes} . Le nombre de paramètres à optimiser se réduit donc à 9, les 5 de $\mathbf{p}_{Polygone}$ et les 4 coordonnées des deux points A et B , comme on peut le voir sur la Figure 4.15.

On a vu au chapitre précédent que si l'on laissait le modèle polygonal de feuille simple totalement libre, certaines des formes qu'il était capable de prendre risquaient de ne pas du tout correspondre à des formes prises par des feuilles. Il faut donc s'assurer que cela ne se produise pas, en rajoutant des contraintes sur les valeurs des paramètres. Plutôt que de prendre des contraintes rigides, nous avons opté pour plus de flexibilité en les incluant dans un terme d'énergie. L'énergie interne du modèle peut-être vue comme étant en grande partie implicite, contenue dans sa construction, mais l'ajout de contraintes sur les valeurs pouvant être prises par ses paramètres en fait aussi partie.

Pour créer cette contrainte, nous avons appris les valeurs possibles des paramètres, à partir de modèles polygonaux ajustés à la main sur des exemples de feuilles suffisamment variés, ainsi qu'on a vu qu'il était possible de faire. La distribution de points dans l'espace des paramètres ainsi obtenue permet d'entraîner un modèle statistique en l'occurrence un modèle de mixture de gaussiennes à 10

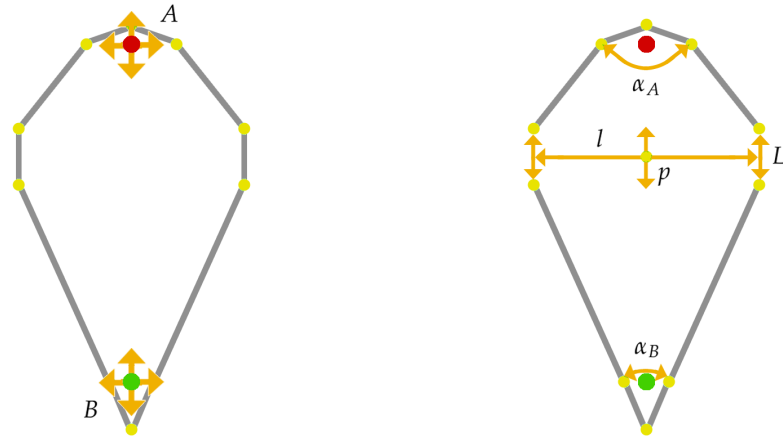


FIGURE 4.15 – Déformations possibles du modèle polygonal de feuille simple non-lobée : déplacements des points (a) et variations des paramètres (b)

composantes capable de représenter les différentes formes foliaires. Ce modèle une fois entraîné donne un ensemble de couples moyenne/covariance $\{\mu_{Polygone,g}; \Sigma_{Polygone,g}\}_{g=1..n_{Polygone}}$ qui vont permettre de définir une mesure de distance aux formes possibles de feuille comme une distance de Mahalanobis minimale. L'énergie interne qu'il faudra minimiser va alors simplement se reposer sur cette distance, de façon à obtenir une forme qui soit la plus proche possible d'une forme de feuille connue :

$$E_{interne}(\Pi) = D_{Feuille}(\mathbf{p}_{Polygone})$$

$$= \min_{g=1..n_{Polygone}} \left((\mathbf{p}_{Polygone} - \mu_{Polygone,g})^T \Sigma_{Polygone,g}^{-1} (\mathbf{p}_{Polygone} - \mu_{Polygone,g}) \right)^{\frac{1}{2}} \quad (4.19)$$

Au final, l'énergie totale à minimiser par le polygone au cours de son évolution va simplement être une somme de son énergie interne pondérée par un réel $\omega_{interne}$ et de son énergie d'attache :

$$E(\Pi) = E_{attache}(\Pi) + \omega_{interne} E_{interne}(\Pi) \quad (4.20)$$

L'approche choisie s'apparente alors à une descente de gradient de cette énergie en fonction des paramètres, une approche itérative visant à choisir à chaque étape la meilleure déformation à appliquer au polygone pour diminuer son énergie, en suivant la direction du plus fort gradient. On cherche donc à chaque étape à estimer la variation d'énergie induite par la variation des paramètres du polygone courant en estimant pour tout paramètre de $\mathbf{p}_{Polygone}$ (ainsi que pour les coordonnées des points B et A , de la même façon) des différences d'énergie, par exemple :

$$\Delta E_+(p) = E(\Pi(B, A, \mathbf{p}_{Polygone} + \Delta p)) - E(\Pi(B, A, \mathbf{p}_{Polygone}))$$

$$\Delta E_-(p) = E(\Pi(B, A, \mathbf{p}_{Polygone} - \Delta p)) - E(\Pi(B, A, \mathbf{p}_{Polygone}))$$

$$\Delta E_+(x_B) = E(\Pi(B + \Delta x, A, \mathbf{p}_{Polygone})) - E(\Pi(B, A, \mathbf{p}_{Polygone}))$$

$$\Delta E_-(x_B) = E(\Pi(B - \Delta x, A, \mathbf{p}_{Polygone})) - E(\Pi(B, A, \mathbf{p}_{Polygone}))$$

Ces différences d'énergies permettent d'effectuer une approximation des dérivées partielles de l'énergie en fonction des paramètres. En effet, le principal problème de la représentation par un modèle paramétrique vient de la difficulté d'appliquer des méthodes variationnelles, et le fait que les dérivées partielles en fonction des paramètres soient impossibles à calculer explicitement rend l'optimisation du processus peu envisageable. Il faut donc se contenter d'estimer localement ces dérivées locales à chaque itération :

$$\frac{\partial E}{\partial l} = \frac{\Delta E_+(l) - \Delta E_-(l)}{2\Delta l}$$

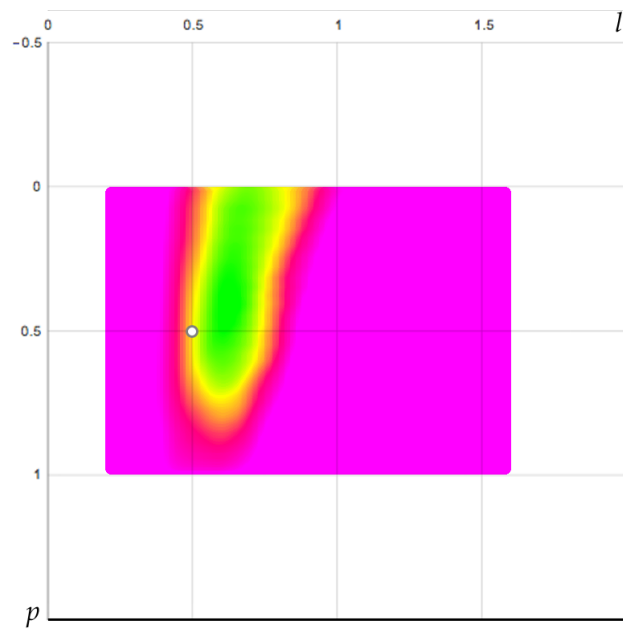
Les paramètres et les coordonnées des points pris indifféremment sont alors classés selon la valeur absolue de leur dérivée partielle de l'énergie. A ce stade deux approches sont possibles et se justifient : soit conserver uniquement la variable qui a la valeur absolue de la dérivée maximale ∂_{\max} , soit conserver toutes les variables et utiliser le vecteur formé par toutes les dérivées partielles et le considérer comme le gradient d'énergie dont il va falloir suivre la direction opposée. Dans ce cas on applique les transformations à toutes les variables du modèle à la fois pour passer de l'état (t) au suivant :

$$\begin{aligned} A^{(t+1)} &= A^{(t)} + \frac{1}{\partial_{\max}} \left(\frac{\partial E}{\partial x_A} \Delta x + \frac{\partial E}{\partial y_A} \Delta y \right) \\ \alpha_B^{(t+1)} &= \alpha_B^{(t)} + \frac{1}{\partial_{\max}} \left(\frac{\partial E}{\partial \alpha_B} \Delta \alpha_B \right) \end{aligned}$$

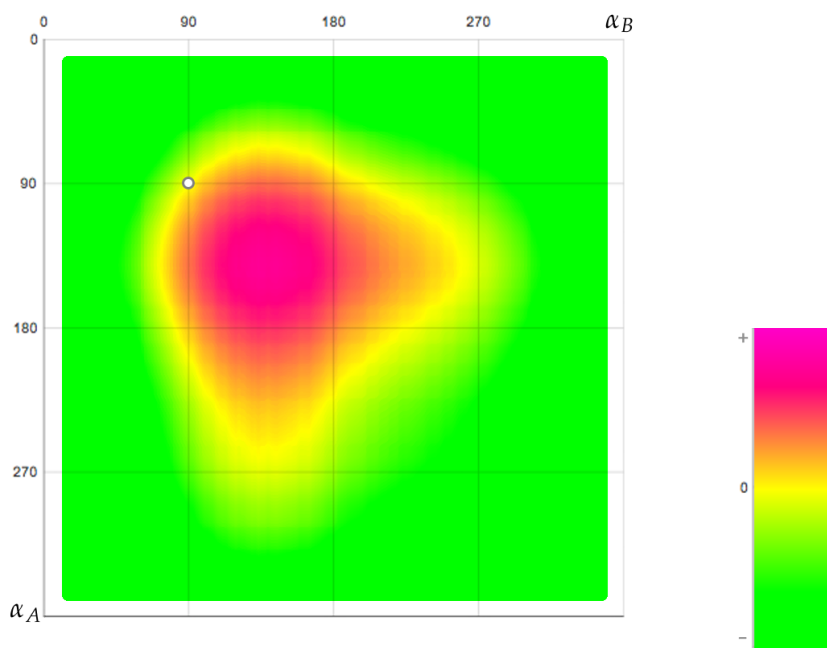
La méthode consiste alors, à partir d'une position initiale $\Pi^{(0)}$, à calculer les dérivées partielles et appliquer les transformations qui maintiennent les paramètres dans une plage acceptable et le modèle dans l'image, ce jusqu'à atteindre une convergence où l'énergie ne peut plus être diminuée. Tel quel, cet algorithme d'optimisation présente quelques risques, notamment celui de rester bloqué dans un minimum local d'énergie qui ne correspondrait pas véritablement à la forme optimale que l'on espère trouver. La Figure 4.16 permet de visualiser les variations de l'énergie liées à celles de paramètres du modèle et montre bien que trouver le point minimisant globalement l'énergie n'est pas évident, du fait de la présence de minima locaux assez proches, ou lorsque le profil d'énergie est plus propice à l'émergence de maxima que de minima (profil "plat" avec des bosses).

Un autre problème vient de l'implémentation de la méthode, où l'on construit un polygone dont les points sont en fait des pixels à coordonnées entières. Il y a donc une certaine dose d'approximation dans l'arrondi des coordonnées des points, et donc dans le tracé des côtés des polygones, eux aussi discrétisés. Cela a pour conséquence d'avoir une fonction d'énergie discrète, avec surtout la possibilité de sauts un peu brutaux lorsque la variation d'un paramètre fait bouger un point.

Pour contourner ces obstacles souvent rencontrés dans ce genre de méthode, nous utilisons une approche inspirée du recuit simulé, de façon à éviter les minima locaux d'énergie. Elle repose sur une variable de température \mathcal{T} qui va décroître progressivement et être brutalement augmentée à intervalle régulier. Cette température a un intérêt double. Elle force tout d'abord le déplacement, même



(a)



(b)

FIGURE 4.16 – Exemples de profils d'énergie en fonction de deux paramètres du modèle polygonal simple relativement à un point donné : largeur l et position p (a), angles α_B et α_A (b)

lorsque l'on se trouve en un minimum d'énergie, en augmentant la valeur des dérivées partielles. Elle va également venir conditionner l'amplitude Δ des déplacements, qui seront plus brutaux à haute température et plus fins lorsqu'elle diminue. On notera dès lors $\frac{\partial E}{\partial p}(\mathcal{T})$, et pour les déplacements $\Delta x(\mathcal{T})$ par exemple.

La température part d'une valeur initiale \mathcal{T}_{\max} et on la multiplie à chaque étape par un facteur λ inférieur à 1, jusqu'à atteindre une valeur \mathcal{T}_{\min} , à laquelle elle est remontée jusqu'à \mathcal{T}_{\max} . A chacune de ces hausses de températures, on fait subir au modèle une perturbation aléatoire sur les valeurs de ses paramètres de façon à repartir d'un point différent de l'espace des paramètres. Ces cycles sont répétés jusqu'à ce que la descente de gradient aboutisse à des valeurs des paramètres suffisamment proches entre deux cycles consécutifs. L'Algorithme 2 reprend le déroulement de cette méthode d'optimisation, et les résultats de l'évolution sur des feuilles simples sont donnés dans la Figure 4.17.

Algorithme 2 Optimisation du modèle de feuille simple par descente de gradient d'énergie

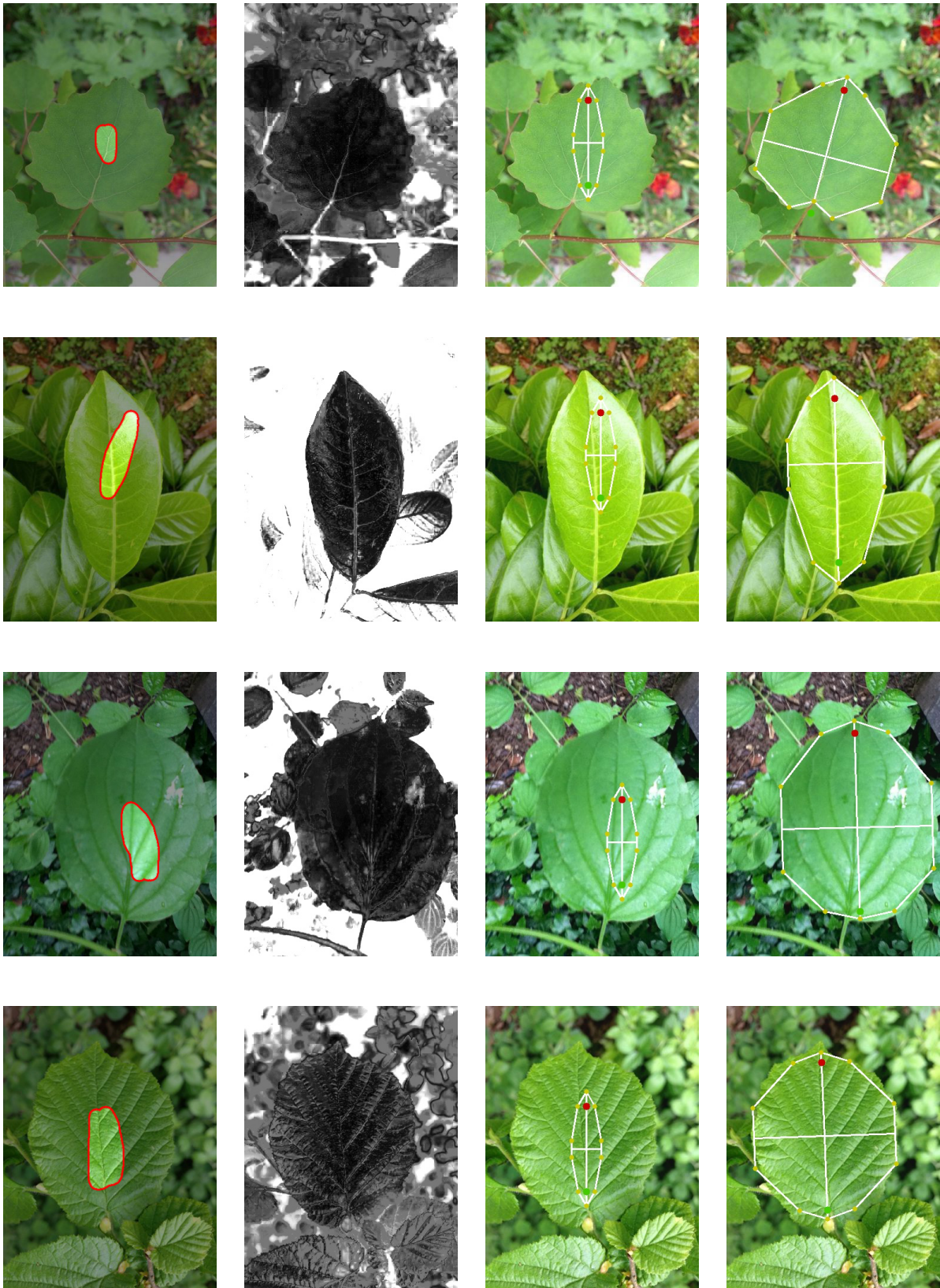
Entrée : $d_{Croyance}$ la carte de dissimilarité; $A^{(0)}, B^{(0)}, \mathbf{p}_{Polygone}^{(0)}$ point de départ de l'algorithme

Sortie : Π^* le modèle polygonal après évolution

```

 $\Pi^{(0)} \leftarrow \Pi(A^{(0)}, B^{(0)}, \mathbf{p}_{Polygone}^{(0)})$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{\max}$ 
   $\Pi^{(cycle)} \leftarrow$  perturbe( $\Pi^{(cycle-1)}$ )
  while  $\mathcal{T} > \mathcal{T}_{\min}$  do
    for all  $\mathbf{p} \in \{x_B, y_B, x_A, y_A, l, p, L, \alpha_B, \alpha_A\}$  do
       $\mathbf{p}^{(cycle)} \leftarrow \mathbf{p}^{(cycle)} + \frac{1}{\partial_{\max}} \left( \frac{\partial E}{\partial p}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  OK  $\leftarrow \left( \|\Pi^{(cycle)} - \Pi^{(cycle-1)}\| < \Delta \Pi_{\max} \right) \wedge (cycle < cycle_{\max})$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Pi^* \leftarrow \Pi^{(cycle)}$ 
return  $\Pi^*$ 

```



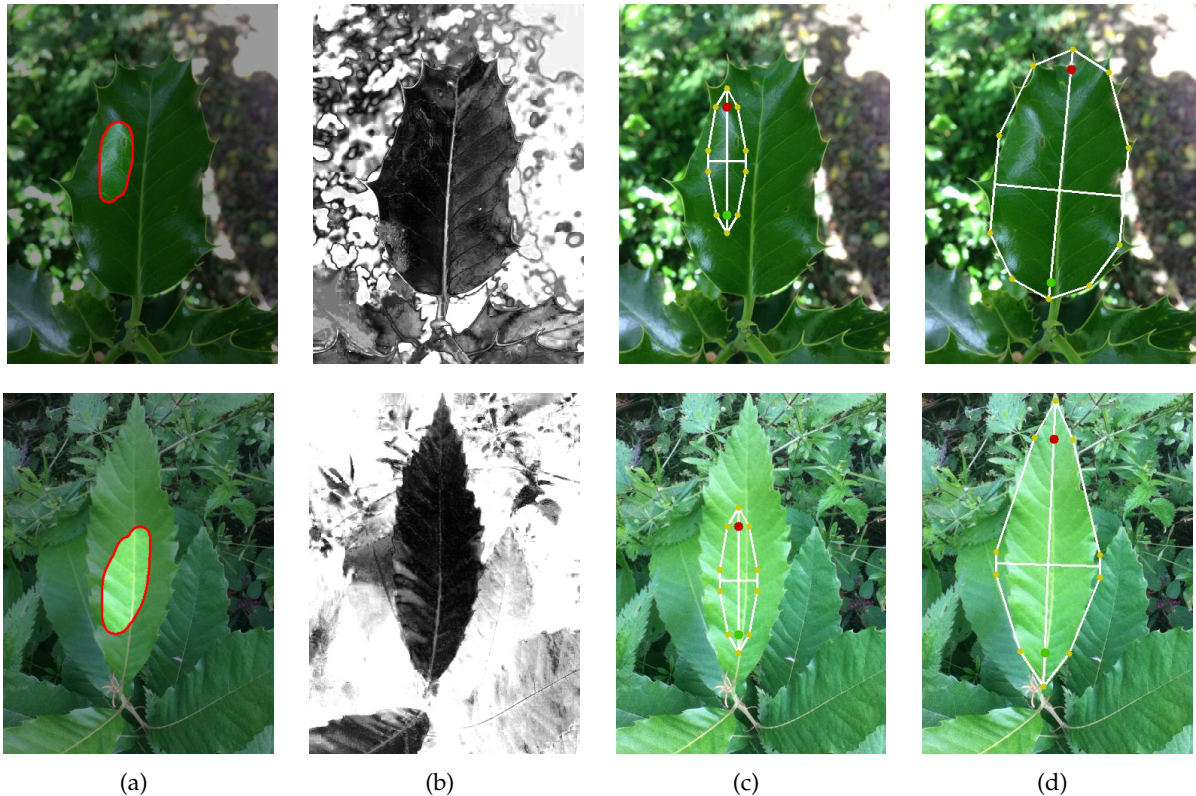


FIGURE 4.17 – Evolution du modèle polygonal paramétrique déformable de feuille simple : initialisation manuelle (a), carte de dissimilarité calculée à partir de la région initiale (b), polygone initial (c) et polygone après optimisation des paramètres (d)

Dans le cas des feuilles palmatilobées, le nombre de paramètres à prendre en compte pour l'optimisation est un peu plus conséquent. Cependant, si considérer les éléments de \mathbf{p}_{Lobes} de la même façon que les paramètres du polygone pour estimer les dérivées partielles ne pose pas de problème, il en va tout autrement du nombre de lobes. Un changement dans la valeur de ce paramètre se traduit par une déformation bien trop importante pour penser que l'on mesure une dérivée locale. Ce paramètre joue un rôle particulier, et doit par conséquent être optimisé de manière un peu différente.

L'idée est simple, et elle part du principe que l'on maîtrise l'effet d'une suppression de lobes alors que l'ajout d'une paire supplémentaire de polygones laisse trop d'incertitudes. Nous partons donc d'un modèle polygonal de feuille palmatilobée avec un nombre clairement excessif de lobes, par exemple $n_L = 6$ (il ne se trouve aucune espèce atteignant les 13 lobes parmi les essences natives européennes). Pour simplifier la suite, ces lobes vont être contraints à s'échelonner, c'est-à-dire que les paramètres \mathbf{p}_{Lobes} vont devoir vérifier : $\forall l \in \{1..n_L\}, (h_l < h_{l-1}) \wedge (\theta_l > \theta_{l-1})$

De cette façon, les lobes initiaux ne peuvent pas se recouvrir ni se croiser. On est ainsi assuré qu'ils seront effectivement disposés dans le modèle selon l'ordre de leurs indices, et que les lobes secondaires les plus grands seront les plus proches du lobe principal (ce qui est une constante chez les feuilles palmatilobées).

Une fois initialisé le modèle $\Pi^{(0)}$ avec ces lobes en excès, il est possible d'appliquer exactement le même algorithme de minimisation d'énergie qu'introduit ci-dessus, en considérant les variations de \mathbf{p}_{Lobes} en plus de celles de B , A et $\mathbf{p}_{Polygone}$ dans l'estimation des dérivées partielles de l'énergie et donc dans les déformations applicables au modèle. Ce que l'on observe alors, c'est qu'en évoluant les polygones simples Π_l vont avoir tendance à se regrouper à l'intérieur des véritables lobes de la feuille, à se superposer jusqu'à se recouvrir en grande partie comme le montre la Figure 4.18.



FIGURE 4.18 – Recouvrement des polygones en excès après évolution sur une feuille palmatilobée

C'est dans cet état, une fois le modèle stabilisé à la fin d'un cycle de température, que l'on va pouvoir observer la répartition des polygones représentant les lobes de façon à supprimer ceux qui sont superflus. Pour faire cette sélection, nous utilisons sur un facteur de recouvrement $\rho_{l-1 \rightarrow l}$. La formulation de ce facteur se base sur une approximation des polygones par des rectangles, et essaie d'estimer un rapport pertinent :

$$\rho_{l-1 \rightarrow l} = \frac{1}{h_l} \frac{\frac{l}{2} \times h_{l-1}}{\sin \theta} \cos(\theta) \quad \theta = \theta_l - \theta_{l-1} \quad (4.21)$$

Si l'on assimile les polygones de deux lobes successifs à deux rectangles faisant un angle θ entre eux, de longueur respective h_{l-1} et h_l et de largeur respective $l \times h_{l-1}$ et $l \times h_l$, l'axe BA_l du second intersecte le bord du premier en un point I_l tel que :

$$BI_l = \frac{\frac{l}{2} \times h_{l-1}}{\sin \theta}$$

La portion du second lobe recouverte par le premier peut alors être approximée par le rapport BI_l / BA_l . Cette mesure n'est pas tout à fait exacte, notamment lorsque l'angle θ augmente. Pour l'affiner, et également pour pénaliser plus les lobes très proches en angle, on pondère ce rapport par $\cos(\theta)$, ce qui ramène à l'expression de $\rho_{l-1 \rightarrow l}$.

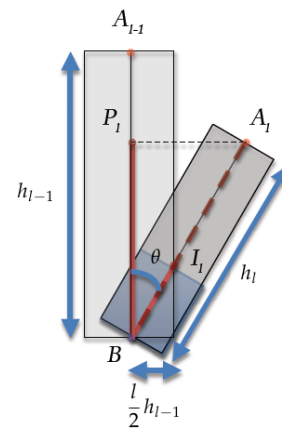


FIGURE 4.19 – Calcul du recouvrement entre deux lobes successifs

La méthode est alors simple : en partant du premier lobe après le lobe principal, on calcule ce facteur de recouvrement, de façon à supprimer les paires de polygones pour lesquelles il excède une certaine valeur, généralement fixée à 0.7. Lorsqu'une paire de lobes est supprimée, le recouvrement du lobe suivant n'est pas calculé avec celle-ci mais avec la première paire précédente restante, de façon à ne pas supprimer trop de lobes d'un coup. L'algorithme d'optimisation du modèle est alors celui décrit dans l'Algorithme 3, et il est aussi valable pour les feuilles simples, puisque par suppressions successives, on peut retomber à un modèle à un seul polygone.

Algorithme 3 Optimisation du modèle de feuille palmatilobée par descente de gradient d'énergie

Entrée : $d_{Croyance}$ la carte de dissimilarité ; $A^{(0)}, B^{(0)}, \mathbf{p}_{Polygone}^{(0)}, \mathbf{p}_{Lobes}^{(0)}$ point de départ de l'algorithme

Sortie : Π^* le modèle polygonal après évolution

```

 $\Pi^{(0)} \leftarrow \Pi(A^{(0)}, B^{(0)}, \mathbf{p}_{Polygone}^{(0)}, \mathbf{p}_{Lobes}^{(0)})$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{max}$ 
   $\Pi^{(cycle)} \leftarrow$  perturbe( $\Pi^{(cycle-1)}$ )
  while  $\mathcal{T} > \mathcal{T}_{min}$  do
    for all  $\mathbf{p} \in \{x_B, y_B, x_A, y_A, l, p, L, \alpha_B, \alpha_A, h_1, \theta_1, \dots, h_{n_L}, \theta_{n_L}\}$  do
       $\mathbf{p}^{(cycle)} \leftarrow \mathbf{p}^{(cycle)} + \frac{1}{\partial_{max}} \left( \frac{\partial E}{\partial \mathbf{p}}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  SUPPR  $\leftarrow$  false
  for all  $l \in \{1, \dots, n_L\}$  do
    if  $\rho_{l-1 \rightarrow l} > 0.7$  then
      suppress( $\Pi^{(cycle)}, l$ )
      SUPPR  $\leftarrow$  true
    end if
  end for
  OK  $\leftarrow$   $\left( \|\Pi^{(cycle)} - \Pi^{(cycle-1)}\| < \Delta \Pi_{max} \right) \wedge (\neg \text{SUPPR}) \wedge (cycle < cycle_{max})$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Pi^* \leftarrow \Pi^{(cycle)}$ 
return  $\Pi^*$ 

```

Cette démarche d'optimisation *a posteriori* permet d'estimer de façon relativement robuste un paramètre dont l'influence sur la forme est bien trop importante pour être pris en compte au même niveau que les autres. Il s'agit de plus d'un moyen élégant de le faire puisque cela fait partie du même processus d'optimisation, et ne nécessite pas un pré-traitement hasardeux qui chercherait à estimer le nombre de lobes avant-même de mener à bien la segmentation. Des exemples de ce mécanisme de suppression sont donnés dans la Figure 4.20.

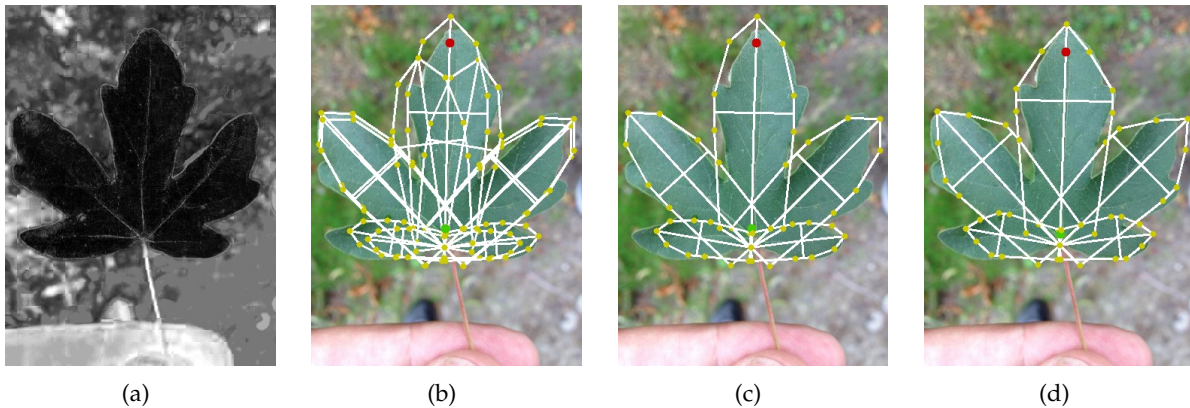


FIGURE 4.20 – Evolution du modèle polygonal paramétrique déformable de feuille palmatilobée : carte de dissimilarité (a), lobes après une première évolution (b), suppression des lobes à l’aide du recouvrement (c) et polygone final (d)

Ce que l’on obtient en fin de compte est une modélisation de la forme globale des feuilles simples, palmatilobées ou non, effectuée uniquement à l’aide de la similarité de couleur à une région initiale et d’une connaissance *a priori* de la forme. Cette contrainte de forme permet de contourner quelques difficultés, même quand la carte de dissimilarité n’est pas idéale et que des pixels se trouvant à l’arrière-plan sont considérés comme proches de la feuille. La rigidité relative du modèle permet en effet de ne pas déborder dans les régions voisines mais de toujours maintenir une forme de feuille plausible.

Les modèles optimisés automatiquement sur les images fournissent déjà une interprétation à assez haut niveau du contenu de l’image, avec des informations directement interprétables. Ils constituent surtout une première étape en vue de la détermination du contour exact de la feuille qui va rendre plus simple l’application de méthodes de segmentation.

4.3.3 Modèles déformables de feuilles composées

Les feuilles composées constituent des objets bien différents des feuilles simples, qui ne peuvent pas être modélisées de la même façon. Leur forme apparente peut être assez complexe, et leur compréhension nécessite à la fois la connaissance de leur structure et de l’agencement des folioles, et la forme globale de chacune de ces folioles, lorsque l’on considère le limbe simple. Pour représenter ces feuilles, nous utilisons donc des outils radicalement différents qui se résument à deux modèles paramétriques : l’un pour l’agencement des folioles et l’autre pour leur forme globale.

Il est donc pour ainsi dire impossible d’inclure la modélisation de leur forme dans le même cadre d’optimisation automatique de forme que pour les feuilles simples, ce que nous avons réussi à faire pour les feuilles palmatilobées du fait de leur plus grande proximité. La différence de nature est ici trop grande, et cela signifie par conséquent que le processus de traitement va diverger pour ces deux types de feuilles, et ce dès la phase de segmentation.

La première tâche va donc être de déterminer si on est en présence d’une feuille simple ou

composée. Garder une approche purement automatique pour ce faire poserait alors un véritable problème : il faudrait en effet avoir une idée de la forme de la feuille pour reconnaître son type, alors que c'est justement cette première idée de forme que l'on cherche à obtenir de façon robuste par le biais de modèles ! Tout cela, en plus de se mordre la queue, présente trop de risques étant donnée l'importance de cette première décision sur la suite de la reconnaissance. Dans la suite, nous supposons donc que la distinction entre feuille simple et feuille composée nous est donnée, avec l'image et la région initiale, comme entrée de l'algorithme.

L'optimisation de modèles déformables sur les images de feuilles composées va donc se faire en deux temps. La première étape consiste à estimer le nombre et la disposition des folioles de la feuille, en ajustant un premier modèle où les folioles sont représentées par des cercles. Il est important de noter que le recouvrement potentiel entre les folioles de la feuille empêche l'emploi de méthodes plus basiques, notamment en considérant des composantes connexes. Il est donc préférable de s'appuyer sur le modèle défini au Chapitre 4 en cherchant à faire en sorte que le modèle Φ défini par les points P et T , le nombre de paires de folioles n_F et les paramètres $\mathbf{p}_{Folioles} = (p_1; \dots; p_{n_F}; k; \delta; r)$ minimise l'énergie liée à la couleur.

La région initiale sur laquelle va être bâti le modèle de couleur doit être quelque peu différente pour correspondre aux spécificités des feuilles composées. Pour une initialisation manuelle, nous demandons à l'utilisateur de marquer 3 folioles au minimum : la foliole terminale (que l'on considère double pour les feuilles paripennées) et ses deux plus proches voisines. La marque de la foliole terminale va permettre de placer le point T qui restera fixe. L'ensemble de ces régions disjointes va ensuite être utilisé pour calculer le modèle de couleur.

La nature discontinue de la région formée par la feuille composée dans l'image n'est alors pas sans poser de gros problèmes. Notamment, la propagation de la moyenne locale échoue à s'adapter correctement aux variations de couleurs lorsque celles-ci se font d'une foliole à l'autre, et qu'une bande d'arrière-plan les sépare. De plus, la région initiale étant souvent plus réduite, la plage de luminosité qu'elle couvre risque d'être trop réduite pour que la régression linéaire évalue correctement l'axe du cylindre de couleurs. Pour ces raisons, et pour limiter les risques de cartes de dissimilarité complètement erronées, nous avons choisi de conserver une mesure plus simple, la distance de Mahalanobis aux couleurs $L^*a^*b^*$ de la région initiale $d_{M,Lab}(x, y, \Omega_0)$, et de s'appuyer donc sur l'énergie de couleur définie à l'équation 4.15, pour obtenir l'énergie d'attache suivante :

$$E(\Phi) = E_{attache}(\Phi) = \int_{\Phi} (d_{M,Lab}(x, y, \Omega_0) - d_{max}) dx dy \quad (4.22)$$

Nous n'avons pas jugé opportun d'introduire une énergie interne additionnelle, puisque les contraintes de formes sont induites par la construction du modèle, et par quelques limites rigides pour les valeurs des paramètres. En particulier nous forçons les folioles à être échelonnées (à ne pas se

dépasser les unes les autres) et les autres paramètres à prendre des valeurs acceptables :

$$\begin{aligned} \forall f \in \{2 \dots n_F\}, p_f > p_{f-1} \\ |k| < 0.5; \quad 0 < \delta < 0.5; \quad 0.05 < r < 0.3 \end{aligned} \quad (4.23)$$

La méthode d'optimisation est exactement la même que pour le précédent modèle, avec estimation des dérivées partielles et cycles de température pour estimer de façon aussi robuste que possible les valeurs de $\mathbf{p}_{Foliolles}$ et les coordonnées de P (le point T étant fixe) qui permettent d'ajuster au mieux le modèle à l'image.

De la même manière que le nombre de lobes du modèle de feuille simple, le nombre de folioles n_F doit être optimisé séparément, étant donné que les changements qu'il engendre dans la forme du modèle sont bien trop importants pour le considérer au même niveau que les autres paramètres dans une descente de gradient. C'est la même méthode qui est donc employée, à savoir : initialiser le modèle avec un nombre excessif de folioles (en l'occurrence 33 au total), laisser les disques s'agglomérer dans les folioles réelles de la feuille, et éliminer les paires de folioles superflues du modèle. Le critère de recouvrement $\rho_{f-1 \rightarrow f}$ utilisé pour déterminer les folioles à supprimer se base alors sur les distances entre les centres des cercles représentant les folioles :

$$\rho_{f-1 \rightarrow f} = \max \left(\frac{r \times h}{F_f^g F_{f-1}^g}, \frac{r \times h}{F_f^d F_{f-1}^d} \right) \quad (4.24)$$

Les folioles seront alors supprimées dans tous les cas si $\rho_{f-1 \rightarrow f}$ dépasse 3, ou s'il dépasse 2 en étant supérieur au recouvrement moyen $\bar{\rho}$ sur l'ensemble des folioles. Une fois encore, lorsqu'une paire de folioles est supprimée, on ne la prend plus en compte pour le calcul du recouvrement, les distances étant alors calculées avec la dernière paire de folioles conservée. Cette phase de suppression est répétée à chaque fin de cycle de température, et la convergence du modèle permet idéalement de se retrouver avec un cercle par foliole de la feuille. L'Algorithme 4 résume ce processus, et la Figure 4.21 donne quelques exemples d'évolution du modèle d'agencement de feuille composée.

Algorithme 4 Optimisation du modèle de feuille composée par descente de gradient d'énergie

Entrée : d_{Croyance} la carte de dissimilarité ; $P^{(0)}, T^{(0)}, \mathbf{p}_{\text{Folioles}}^{(0)}$ point de départ de l'algorithme

Sortie : Φ^* le modèle d'organisation après évolution

```

 $\Phi^{(0)} \leftarrow \Phi(P^{(0)}, T^{(0)}, 16, \mathbf{p}_{\text{Folioles}}^{(0)})$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{\text{max}}$ 
   $\Phi^{(\text{cycle})} \leftarrow \text{perturbe}(\Phi^{(\text{cycle}-1)})$ 
  while  $\mathcal{T} > \mathcal{T}_{\text{min}}$  do
    for all  $\mathbf{p} \in \{x_p, y_p, k, \delta, r, p_1, \dots, p_{n_F}\}$  do
       $\mathbf{p}^{(\text{cycle})} \leftarrow \mathbf{p}^{(\text{cycle})} + \frac{1}{\sigma_{\text{max}}} \left( \frac{\partial E}{\partial \mathbf{p}}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  SUPPR  $\leftarrow$  false
  for all  $f \in \{2, \dots, n_F\}$  do
    if  $(\rho_{f-1 \rightarrow f} > 3) \vee ((\rho_{f-1 \rightarrow f} > 2) \wedge (\rho_{f-1 \rightarrow f} > \bar{\rho}))$  then
      suppress( $\Phi^{(\text{cycle})}, f$ )
      SUPPR  $\leftarrow$  true
    end if
  end for
  OK  $\leftarrow$   $\left( \|\Phi^{(\text{cycle})} - \Phi^{(\text{cycle}-1)}\| < \Delta \Phi_{\text{max}} \right) \wedge (\neg \text{SUPPR}) \wedge (\text{cycle} < \text{cycle}_{\text{max}})$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Phi^* \leftarrow \Phi^{(\text{cycle})}$ 

```

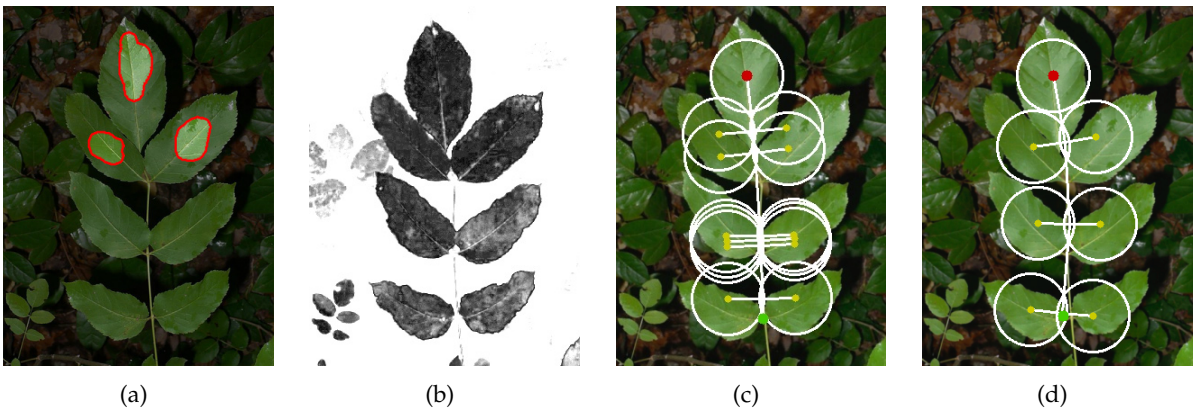


FIGURE 4.21 – Evolution du modèle d'agencement de feuille composée : coloriage de la région initiale (a), carte de dissimilarité au modèle gaussien de couleur (b), folioles superposées après une première évolution (c) et modèle obtenu après suppression et évolution (d)

En vue d'une segmentation exacte, et dans un souci de description, nous voulons ensuite modéliser la forme de chacune des folioles de la feuille par un modèle polygone de feuille simple non-lobée, en estimant la modélisation paramétrique jointe de folioles qui s'adaptera le mieux à l'image. Après l'évaluation du premier modèle de feuille composée, nous disposons d'une bonne idée du nombre et de la localisation des folioles, ce qui va permettre d'initialiser les modèles de folioles dans une position assez proche de celle qu'il prendra au final.

Le nombre de modèles étant fixé à $2n_F + 1$ il ne fait plus partie des quantités à optimiser. Il faut en revanche trouver les positions des points B_f et A_f ($f \in \{-n_F \dots n_F\}$) et à ajuster la forme du polygone en jouant sur les valeurs de $\mathbf{p}_{Polygone}$ de façon à ce que le modèle Π minimise encore une fois une énergie.

$$\begin{aligned} \Pi &= \Pi(n_F, B_{-n_F}, A_{-n_F}, \dots, B_{n_F}, A_{n_F}, \mathbf{p}_{Polygone}) \\ &= \bigsqcup_{f=-n_F}^{n_F} \Pi_f(B_f, A_f, \mathbf{p}_{Polygone}) \end{aligned} \quad (4.25)$$

Le terme d'attache de l'énergie est le même que précédemment, mais une énergie interne est cette fois-ci rajoutée, de façon à la fois à s'assurer que les folioles conservent une forme plausible, mais également pour pousser les paires de folioles à demeurer régulières. En particulier, on veut éviter une trop grande disparité de taille, et essayer de conserver des paires de folioles deux à deux symétriques. Un terme basé sur la variance des longueurs $B_f A_f$ des modèles pousse à avoir des tailles identiques, et un second terme basé sur le déterminant pousse les paires de folioles à être symétriques (le déterminant étant nul lorsque $(B_{-f} B_f) \parallel (A_{-f} A_f)$). La définition de l'énergie interne du modèle est alors la suivante :

$$\begin{aligned} E_{interne}(\Pi) &= D_{Feuille}(\mathbf{p}_{Polygone}) + \omega_{symetrie} D_{symetrie}(B_{-n_F}, A_{-n_F}, \dots, B_{n_F}, A_{n_F}) \\ D_{symetrie}(B_{-n_F}, \dots, A_{n_F}) &= \left(\frac{1}{h^2} \text{var} \left(\{A_f B_f\}_{f \in \{-n_F \dots n_F\}} \right) + \frac{1}{n_F} \sum_{f=1}^{n_F} \left(\frac{1}{h^2} \det(\overrightarrow{B_{-f} B_f}, \overrightarrow{A_{-f} A_f})^2 \right) \right) \end{aligned} \quad (4.26)$$

L'énergie que le modèle doit minimiser au cours de son évolution est alors comme en 4.20 une somme pondérée de ses énergies interne et d'attache. Ici, tous les paramètres peuvent être pris au même niveau, et inclus dans le même processus d'optimisation. La seule différence étant qu'à tout moment, il faut laisser la possibilité à un modèle se trouvant là où il n'y a pas de foliole (foliole effectivement manquante dans la feuille, nombre impair de modèles pour une feuille paripennée, ou tout simplement nombre de folioles surévalué par le modèle précédent) de disparaître définitivement, et de ne plus rentrer en ligne de compte lors de l'évolution. On procède de fait à cette suppression dès lors que $B_f A_f / h < h_{\min}$, mais il ne s'agit pas d'une déformation du modèle, seulement d'une simplification des calculs lorsque l'optimisation des paramètres tend à faire disparaître certains des polygones.

Le fait que les folioles évoluent ainsi de manière simultanée, et qu'elles ne doivent pas s'intersecter

(union disjointe), assure une sorte d'auto-contrainte dans l'évolution. C'est ce qui limite notamment les risques de recouvrement, même dans les cas où les folioles de la feuille se recouvrent elles-mêmes largement. Les paramètres de forme sont de plus considérés comme étant les mêmes pour toutes les folioles, ce qui rajoute une forte contrainte, et garantit au final que les modèles optimaux sont obtenus de manière bien plus robuste que s'ils étaient calculés indépendamment pour chaque foliole. En plus du gain en temps de calcul, cela permet un moyennage implicite des erreurs, qui accorde une confiance plus grande dans la valeur des paramètres $\mathbf{p}_{Polygone}$ obtenus en vue de la description de forme. Le détail du processus d'optimisation est donné dans l'Algorithme 5, et quelques exemples de modèles obtenus dans la Figure 4.22.

Algorithme 5 Optimisation du modèle joint de folioles par descente de gradient d'énergie

Entrée : $d_{Croyance}$ la carte de dissimilarité ; Φ^* le modèle d'organisation après évolution

Sortie : Π^* le modèle polygonal joint après évolution

```

 $\Pi^{(0)} \leftarrow \Pi(n_F, B_{-n_F}^{(0)}(\Phi^*), A_{-n_F}^{(0)}(\Phi^*), \dots, B_{n_F}^{(0)}(\Phi^*), A_{n_F}^{(0)}(\Phi^*), \mathbf{p}_{Polygone}^{(0)}$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{max}$ 
   $\Pi^{(cycle)} \leftarrow$  perturbe( $\Pi^{(cycle-1)}$ )
  while  $\mathcal{T} > \mathcal{T}_{min}$  do
    for all  $\mathbf{p} \in \{x_{B_{-n_F}}, y_{B_{-n_F}}, x_{A_{-n_F}}, y_{A_{-n_F}}, \dots, x_{B_{n_F}}, y_{B_{n_F}}, x_{A_{n_F}}, y_{A_{n_F}}, l, p, L, \alpha_B, \alpha_A\}$  do
       $\mathbf{p}^{(cycle)} \leftarrow \mathbf{p}^{(cycle)} + \frac{1}{\partial_{max}} \left( \frac{\partial E}{\partial \mathbf{p}}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
      for all  $f \in \{-n_F, \dots, n_F\}$  do
        if  $(B_f A_f / h < h_{min})$  then
          suppress( $\Pi^{(cycle)}, f$ )
        end if
      end for
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  OK  $\leftarrow$   $(\|\Pi^{(cycle)} - \Pi^{(cycle-1)}\| < \Delta \Pi_{max}) \wedge (cycle < cycle_{max})$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Pi^* \leftarrow \Pi^{(cycle)}$ 
return  $\Pi^*$ 

```

L'information obtenue après cette phase de pré-segmentation des feuilles composées est très riche. Elle donne tout d'abord le nombre de folioles de la feuille (à 1 près, étant donné que le modèle est créé pour des feuilles avec un nombre impair de folioles) une information qu'il est par ailleurs assez compliqué d'obtenir *a posteriori*. Surtout, une telle modélisation permet une individualisation des

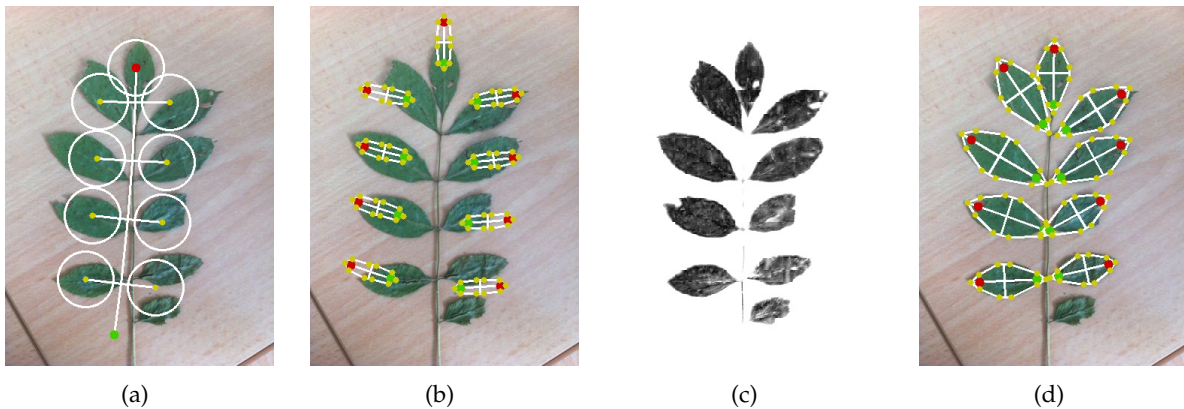


FIGURE 4.22 – Evolution du modèle joint de folioles : résultat du modèle de feuille composée (a), initialisation des modèles polygonaux de folioles (b), carte de dissimilarité (c) et modèle final (d)

folioles, qui vont dès lors pouvoir être segmentées, analysées, décrites séparément. Les modèles polygonaux fournissent des données sur la taille, ou les axes des folioles qui s'avèrent très utile pour la compréhension de la feuille. Enfin, les régions que délimitent le modèle constituent une bonne base en vue de la segmentation exacte et de la récupération du contour de la feuille.

4.4 Contours actifs guidés par les modèles de forme

Au terme de la première phase de segmentation, l'optimisation de modèles paramétriques de formes foliaires permet d'obtenir une région polygonale Π^* , éventuellement composée de plusieurs polygones individualisés. Si elle ne correspond pas rigoureusement à la feuille, cette région en constitue une approximation généralement peu distante, et la déformation de son contour qui permettrait d'arriver à la forme exacte du bord de la feuille peut être considérée comme minime.

La méthode que nous avons choisie pour mener à bien la segmentation finale va alors être une simple transposition de cette idée, en cherchant à déformer un contour Γ en s'appuyant sur le guide que constitue la région Π^* . Le choix d'un modèle de contour actif semble alors naturel, en se servant de Π^* à la fois comme point de départ du contour à déformer et comme contrainte de forme dont le contour doit à tout moment rester proche.

4.4.1 Energie de guidage par la forme

Pour prendre en compte la forme dans l'évolution d'un contour actif, quelle que soit la représentation paramétrique ou implicite qui est choisie, il faut définir un terme dans l'énergie dont la minimisation conduira à une forme parfaitement en accord avec la contrainte posée. Autrement dit, il est nécessaire d'avoir un terme d'énergie, que l'on va dire externe puisque ne provenant que d'informations extérieures à l'image donnée en entrée mais variables, qui renferme dans sa définition la contrainte sur la forme.

Dans notre cas, la contrainte qui a été formulée est que le contour ne devait pas trop s'éloigner de

la forme définie par le polygone Π^* , et il semble donc naturel de baser cette énergie sur la distance du contour aux côtés du polygone. On peut cependant laisser une marge de manœuvre au contour, une zone autour du polygone où il sera libre d'évoluer, que l'on peut voir comme la taille des déformations nécessaires pour s'ajuster aux sinuosités ou aux dents assez marquées du bord d'une feuille.

Un problème vient toutefois du fait que, le modèle polygonal étant une approximation de la forme globale, il risque parfois de se trouver localement assez loin du bord de la feuille, notamment au niveau de la base et de l'apex qui ne sont généralement modélisés que par une valeur d'angle. Il va donc falloir donner un petit peu plus de latitude au contour dans ces zones, et ne pas le contraindre à rester trop proche du polygone, cela mettrait en danger l'évaluation correcte de la forme à la base ou l'apex, caractéristiques essentielles dans le processus d'identification des espèces.

On va donc considérer comme référence une version légèrement modifiée du polygone Π^* que l'on va allonger suivant son axe principal au niveau de sa base et de son apex. On définit ainsi deux points A_{\uparrow}^A et B_{\downarrow}^B à partir de l'axe (BA) du polygone de façon à ce qu'ils se trouvent à l'extérieur du polygone :

$$\begin{aligned}\overrightarrow{BA_{\uparrow}^A} &= \overrightarrow{BA^A} + 0.1 \cdot \overrightarrow{BA} \\ \overrightarrow{BB_{\downarrow}^B} &= \overrightarrow{BB^B} - 0.05 \cdot \overrightarrow{BA}\end{aligned}$$

La région que l'on prend comme référence pour l'évolution du contour actif sera alors formée du polygone auquel on rajoute une bande de pixels autour de l'axe formé par le segment $[B_{\downarrow}^B A_{\uparrow}^A]$, en considérant simplement les pixels à une distance suffisamment faible. On obtient ainsi une nouvelle région de polygone étendu Π_{\downarrow}^* :

$$\Pi_{\downarrow}^* = \Pi^* \cup \left\{ (x; y) \mid \|(x; y) - [B_{\downarrow}^B A_{\uparrow}^A]\| < 0.05 \cdot h \right\} \quad (4.27)$$

L'idée est alors d'utiliser la distance d'un point au contour de cette région $C(\Pi_{\downarrow}^*)$ comme mesure de base permettant de juger de l'éloignement à la forme de référence. Cette distance d'un point $(x; y)$ au contour, que l'on notera $\|(x; y) - C(\Pi_{\downarrow}^*)\|$ est définie très simplement de la manière suivante :

$$\begin{aligned}\|(x; y) - C(\Pi_{\downarrow}^*)\| &= \min_{(x'; y') \in C(\Pi_{\downarrow}^*)} \|(x; y) - (x'; y')\| \\ &= \min_{(x'; y') \in C(\Pi_{\downarrow}^*)} \left(((x; y) - (x'; y'))^T ((x; y) - (x'; y')) \right)^{\frac{1}{2}}\end{aligned} \quad (4.28)$$

La fonction que l'on utilise alors pour pousser un point du contour à rester à une distance acceptable de la forme de référence est définie à partir de cette distance mais de façon asymétrique. En effet son comportement ne doit pas être le même selon que l'on se trouve à l'intérieur ou à l'extérieur de la région Π_{\downarrow}^* . Cela est principalement dû au fait que, de la même manière que pour l'énergie des modèles paramétriques déformables, on va utiliser une énergie ballon qui va pousser constamment

le contour à grandir. Pour ramener le contour vers la forme référence, l'énergie de forme va alors pousser dans la même direction que le ballon lorsque l'on est à l'intérieur, et au contraire s'opposer à lui dès que l'on sort du polygone étendu. Il faut donc que la force soit plus forte à l'extérieur pour s'opposer à l'effet du ballon.

On définit alors la fonction de guidage γ_{forme} qui peut être baptisée fonction pochoir, et qui détermine en tout point de l'image l'intensité avec laquelle le contour devra être repoussé vers la forme Π_{\downarrow}^* . Elle utilise un paramètre de marge m qui détermine dans quelle mesure on peut s'éloigner du contour de la région référence sans que cela soit pénalisé, ce qui définit une zone uniformément nulle, d'où l'appellation de pochoir. Pour contrecarrer l'effet du ballon, cette fonction est alors quadratique à l'extérieur du polygone étendu, alors qu'elle est linéaire à l'intérieur, ce qui mène à la définition suivante :

$$\gamma_{forme}(x, y, \Pi_{\downarrow}^*) = \begin{cases} -\max\left(\|(x; y) - C(\Pi_{\downarrow}^*)\| - m, 0\right) & \text{si } (x; y) \in \Pi_{\downarrow}^* \\ \left(\max\left(\|(x; y) - C(\Pi_{\downarrow}^*)\| - m, 0\right)\right)^2 & \text{si } (x; y) \notin \Pi_{\downarrow}^* \end{cases} \quad (4.29)$$

D'un point de vue pratique, cette fonction est calculée en tout point de l'image une seule et unique fois avant l'évolution du contour actif, de façon à disposer d'une carte de distance au polygone, de la même façon qu'on disposait d'une carte de dissimilarité aux couleurs d'une région initiale. Des exemples de la constitution de telles cartes de distance sont données dans la Figure 4.23.

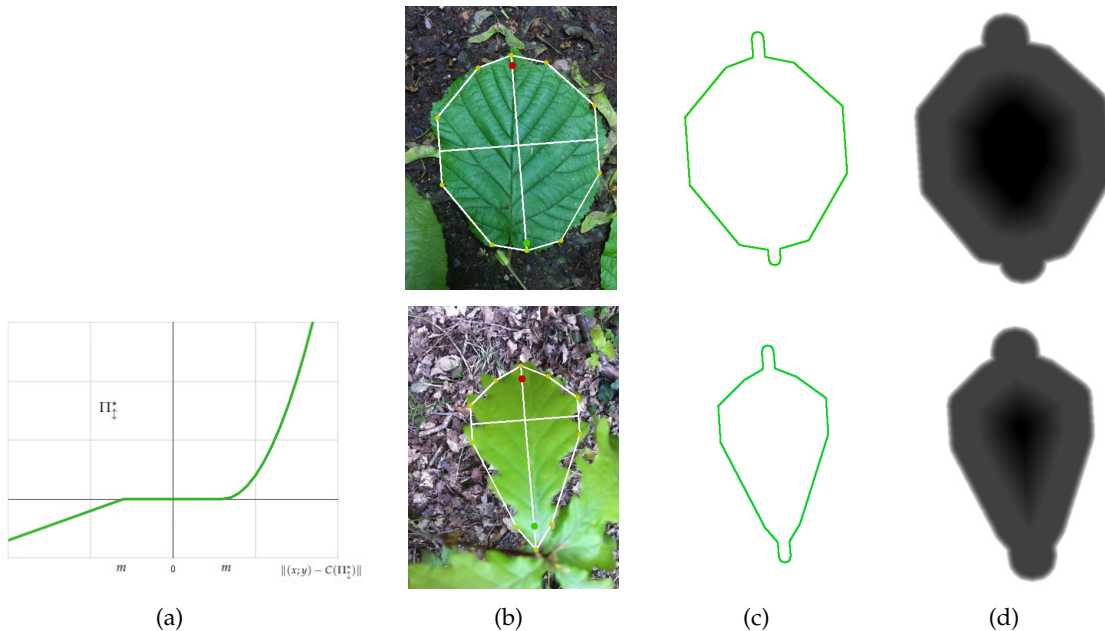


FIGURE 4.23 – Calcul de la fonction de guidage par contrainte de forme (a) : modèle polygonal final Π_{\downarrow}^* (b), extension du polygone à la base et à l'apex Π_{\downarrow}^* (c) et carte de distance (d)

Le terme d'énergie externe qui va être associé à cette fonction a simplement pour but de ramener le contour en évolution Γ vers la zone de déformation acceptable. L'objectif étant que les points

appartenant à la région $\Omega(\Gamma)$ délimitée par le contour aient une fonction de guidage la plus faible possible, le fait qu'elle soit positive à l'extérieur fait qu'il sera coûteux de déborder, et le fait qu'elle soit négative à l'intérieur qu'il sera coûteux de perdre des pixels situés à l'intérieur donc de "rentrez" trop loin de Π_{\downarrow}^* . On définit ainsi une énergie de guidage de forme de la manière suivante :

$$E_{\text{guidage}}(\Gamma) = \omega_{\text{guidage}} \int_{\Omega(\Gamma)} \left(\gamma_{\text{forme}}(x, y, \Pi_{\downarrow}^*) \right) dx dy \quad (4.30)$$

Une telle énergie permet de garantir que le contour, au cours de son évolution ne pourra pas s'éloigner dans une trop grande mesure du polygone obtenu au cours de la première phase. Les contraintes incluses dans la construction de ce dernier agissent comme un *a priori* de forme qui l'empêche de déborder dans les feuilles voisines de couleur similaire, mais c'est un risque qui redevient bien présent dès lors qu'on se défait de la contrainte polygonale et que l'on retrouve un contour plus libre. Une énergie de forme garantit alors que ce genre de débordement, s'il a lieu, ne sera pas trop préjudiciable au résultat final.

4.4.2 Contours actifs guidés

L'énergie de guidage de la forme va alors pouvoir être incluse dans le terme d'énergie d'un modèle de contour actif. On considère que l'on déforme un contour fermé Γ que l'on parcourt en sens antihoraire suivant une abscisse curviligne u , et qui définit une région $\Omega(\Gamma)$. Au cours de l'évolution de la région, il va falloir prendre en compte différents aspects et l'énergie qui va devoir être minimisée peut donc être écrite comme une somme de plusieurs termes indépendants :

$$E(\Gamma) = E_{\text{couleur}}(\Gamma) + E_{\text{gradient}}(\Gamma) + E_{\text{lissage}}(\Gamma) + E_{\text{guidage}}(\Gamma) + E_{\text{ballon}}(\Gamma) \quad (4.31)$$

Le contour que l'on souhaite obtenir à la fin de l'évolution correspond au bord de la feuille, donc bel et bien à une transition entre deux régions, soit une zone de fort gradient. L'information de gradient calculée sur l'image I va donc ici avoir un rôle prépondérant, puisque c'est essentiellement elle qui va permettre au contour de coller avec précision au bord de la feuille. L'énergie de gradient est alors faite pour pénaliser les points du contour qui se trouveraient sur des pixels de faible intensité de gradient, et s'écrit simplement :

$$E_{\text{gradient}}(\Gamma) = \omega_{\text{gradient}} \int_{\Gamma} \left(- \left\| \vec{\nabla} I(\Gamma(u)) \right\| \right) du \quad (4.32)$$

En l'absence d'un gradient fort, il faut tout de même faire la distinction entre la feuille d'intérêt et l'arrière-plan. Pour accomplir cette séparation, nous avons décidé de réutiliser la carte de dissimilarité de couleur de l'étape précédente, en considérant qu'il s'agit déjà d'une bonne mesure de l'appartenance des pixels à la feuille. Le terme d'énergie lié à la couleur est alors le même que précédemment, mais

cette fois-ci calculé sur l'ensemble de la région $\Omega(\Gamma)$ délimitée par le contour, et devient alors :

$$E_{couleur}(\Gamma) = \omega_{couleur} \int_{\Omega(\Gamma)} (d_{croyance}(x, y, \Omega_0)) dx dy \quad (4.33)$$

Bien que le contour final doive être suffisamment précis pour capturer les pointes assez fines et les petites dents du bord des feuilles, il est important d'ajouter une certaine dose de lissage, pour éviter de se retrouver avec un contour trop bruité. Nous rajoutons donc un terme de lissage des plus classiques :

$$E_{lissage}(\Gamma) = \omega_{lissage} \int_{\Gamma} \left\| \frac{d\Gamma}{du} \right\| du \quad (4.34)$$

La contrainte de forme est appliquée via le terme d'énergie de guidage introduit ci-dessus. Ce terme d'énergie externe va agir en compétition avec les autres énergies d'attache notamment. Enfin l'énergie de ballon vient contrebalancer et stabiliser les autres énergies en ajoutant une force constante dirigée vers l'extérieur du contour. Il s'agit de la même énergie que précédemment, minimale lorsque l'aire de la région est maximale :

$$E_{ballon}(\Gamma) = -\omega_{ballon} \int_{\Omega(\Gamma)} dx dy \quad (4.35)$$

La définition de ces différents termes d'énergie permet de calculer en tout point du contour $\Gamma(u) = (x(u); y(u))$ la dérivée fonctionnelle de l'énergie totale à minimiser $\vec{\partial E} / \partial \Gamma$. Il s'agit d'une grandeur vectorielle qui peut être calculée en utilisant le calcul des variations. Pour les termes d'énergie basés sur la région $\Omega(\Gamma)$ on peut montrer que le vecteur résultant suit la direction du vecteur normal unitaire intérieur du contour Γ au point considéré que l'on notera $\vec{n}_{\Gamma}(u)$. Par calcul des variations, on obtient donc les dérivées fonctionnelles suivantes (la quantité $\kappa_{\Gamma}(u)$ désigne la courbure du contour au point $\Gamma(u)$) :

$$\begin{aligned} \frac{\vec{\partial E}}{\partial \Gamma}(u) &= \frac{\vec{\partial E}_{gradient}}{\partial \Gamma}(u) + \frac{\vec{\partial E}_{couleur}}{\partial \Gamma}(u) + \frac{\vec{\partial E}_{lissage}}{\partial \Gamma}(u) + \frac{\vec{\partial E}_{guidage}}{\partial \Gamma}(u) + \frac{\vec{\partial E}_{ballon}}{\partial \Gamma}(u) \\ \frac{\vec{\partial E}_{gradient}}{\partial \Gamma}(u) &= -\omega_{gradient} \vec{\nabla} \left\| \vec{\nabla} I(\Gamma(u)) \right\| \\ \frac{\vec{\partial E}_{couleur}}{\partial \Gamma}(u) &= -\omega_{couleur} d_{croyance}(\Gamma(u), \Omega_0) \vec{n}_{\Gamma}(u) \\ \frac{\vec{\partial E}_{lissage}}{\partial \Gamma}(u) &= \omega_{lissage} \kappa_{\Gamma}(u) \vec{n}_{\Gamma}(u) \\ \frac{\vec{\partial E}_{guidage}}{\partial \Gamma}(u) &= -\omega_{guidage} \gamma_{forme}(\Gamma(u), \Pi_{\downarrow}^*) \vec{n}_{\Gamma}(u) \\ \frac{\vec{\partial E}_{ballon}}{\partial \Gamma}(u) &= \omega_{ballon} \vec{n}_{\Gamma}(u) \end{aligned} \quad (4.36)$$

C'est la donnée de $\vec{\partial E}/\partial \Gamma$ qui va guider l'évolution du contour, qui pour minimiser l'énergie E va se déplacer localement dans la direction opposée au gradient d'énergie. On peut considérer que l'évolution du contour ne se fait que dans la direction de sa normale, puisque les déformations tangentes s'effectuent le long du contour et n'ont aucune influence sur la géométrie de la forme résultante. On peut donc en tout point $\Gamma(u)$ du contour définir une vitesse de déplacement $\vec{v}_\Gamma(u)$ orientée le long de la normale, en projetant la dérivée fonctionnelle de l'énergie sur le vecteur normal :

$$\vec{v}_\Gamma(u) = - \left(\frac{\vec{\partial E}}{\partial \Gamma}(u) \cdot \vec{n}_\Gamma(u) \right) \vec{n}_\Gamma(u) \quad (4.37)$$

Cette vitesse peut ainsi être décomposée en plusieurs vecteurs représentant l'influence de chacun des termes d'énergie sur le déplacement local du contour, en reprenant les expressions issues du calcul des variations :

$$\begin{aligned} \vec{v}_\Gamma(u) = & \omega_{gradient} \left(\vec{\nabla} \left\| \vec{\nabla} I(\Gamma(u)) \right\| \cdot \vec{n}_\Gamma(u) \right) \vec{n}_\Gamma(u) \\ & + \omega_{couleur} d_{croyance}(\Gamma(u), \Omega_0) \vec{n}_\Gamma(u) \\ & - \omega_{lissage} \kappa_\Gamma(u) \vec{n}_\Gamma(u) \\ & + \omega_{guidage} \gamma_{forme}(\Gamma(u), \Pi_\downarrow^*) \vec{n}_\Gamma(u) \\ & - \omega_{ballon} \vec{n}_\Gamma(u) \end{aligned} \quad (4.38)$$

Plutôt que de représenter le contour Γ de façon explicite et de faire se déplacer les points qui le composent, nous avons choisi de le représenter de façon implicite en utilisant une méthode de *level-set* où le contour suit une formulation par courbe de niveau [MSV95]. Ce que l'on fait alors évoluer est une fonction implicite $\phi(x, y)$ qui est à valeurs négatives à l'intérieur de la région $\Omega(\Gamma)$ et positives à l'extérieur, le contour correspondant alors aux passages à zéro de la fonction ϕ . On ne va donc pas considérer le déplacement effectif des points $\Gamma(u)$ mais la mise à jour de la fonction ϕ de façon à inclure ou supprimer des pixels $(x; y)$ de la région $\Omega(\Gamma)$. Cela passe par le calcul d'un coefficient de vitesse de déplacement $v_\phi(x, y)$ dérivée du calcul des variations, qui n'est plus définie que pour les points du contour mais en tout point $(x; y)$ de l'image. La vitesse calculée remplace simplement le lissage par courbure par un lissage Laplacien utilisant $\Delta\phi$, qui a un effet similaire, pour arriver finalement à la formulation suivante :

$$\begin{aligned} v_\phi(x, y) = & \omega_{gradient} \left(\vec{\nabla} \left\| \vec{\nabla} I(x, y) \right\| \cdot \vec{\nabla} \phi(x, y) \right) \\ & + \omega_{couleur} d_{croyance}(x, y, \Omega_0) \left\| \vec{\nabla} \phi \right\| \\ & - \omega_{lissage} \Delta\phi(x, y) \left\| \vec{\nabla} \phi \right\| \\ & + \omega_{guidage} \gamma_{forme}(x, y, \Pi_\downarrow^*) \left\| \vec{\nabla} \phi \right\| \\ & - \omega_{ballon} \left\| \vec{\nabla} \phi \right\| \end{aligned} \quad (4.39)$$

D'un point de vue de l'implémentation, on se place dans le cadre d'un *level-set* binaire [SK08] où la fonction ϕ est à valeurs dans $\{-1; 1\}$. On peut alors se contenter de calculer les vitesses de déplacement pour tous les pixels situés dans une bande étroite de largeur 1 pixel autour du contour Γ , en s'intéressant à leur signe pour déterminer lesquels doivent être rajoutés ou supprimés de la région. Les pixels de cette bande autour de Γ pour lesquels la vitesse est négative sont ajoutés à la région (ou conservés s'ils en faisaient déjà partie) et ceux pour lesquels la vitesse est positive sont retirés de la région s'ils en faisaient partie. L'évolution se poursuit alors, en rajoutant ou en supprimant des pixels dans la région jusqu'à qu'il n'y ait plus de changements possibles, c'est-à-dire que tous les pixels appartenant à $\Omega(\Gamma)$ aient une vitesse négative, et que tous ceux extérieurs à la région aient une vitesse positive.

Les diverses énergies qui interviennent dans l'évolution du contour, bien qu'ayant chacune leur rôle à jouer, forment un ensemble difficile à harmoniser en un tout cohérent, qui soit efficace quelle que soit l'image. Certaines d'entre elles peuvent même présenter des interactions conflictuelles, le terme de dissimilarité de couleur empêchant par exemple le contour d'atteindre des zones de transitions vers lesquelles le terme de gradient l'aurait fait se diriger. La façon dont ces énergies interagissent dépend évidemment de la définition des coefficients ω , et plutôt que de chercher à avoir un jeu de valeurs fixes qui fasse coopérer sans heurt ces forces disparates, nous avons décidé de les faire varier avec le temps.

Le contour n'est pas initialisé rigoureusement sur le polygone Π^* mais sur une version légèrement contractée de cette forme Π_{\bowtie}^* , obtenue par érosion et lissage. Le but est de partir d'un contour $\Gamma^{(0)} = C(\Pi_{\bowtie}^*)$ qui se trouve intégralement à l'intérieur de la feuille. Au début de l'évolution, le coefficient prédominant est celui du gradient, alors que celui de couleur est relativement faible. Cela permet au contour de s'étendre, poussé par le ballon, sans être freiné par des zones de couleurs distantes, pour aller s'arrêter sur un fort gradient que l'on espère être le bord de la feuille. Si un tel gradient ne se présente pas, le terme de forme assure que le débordement ne sera pas trop important. Ensuite, au cours de l'évolution, le terme de couleur est progressivement augmenté jusqu'à atteindre une influence comparable à celle du gradient. Cela ne suffira pas pour décrocher le contour des zones de transition marquées, même si la couleur n'est localement pas celle de la feuille, mais permet de ramener les éventuels débordements dans des zones où la couleur est plus plausible.

En pratique, les valeurs des poids ont été sélectionnées empiriquement, de façon à donner les résultats les plus visiblement fiables. Les valeurs données correspondent aux images en environnement naturel ; en effet le cas beaucoup plus simple des feuilles sur fond uni ne nécessite pas autant d'énergies (pas de risque de débordement, influence de la couleur limitée) et n'est pas aussi difficile à équilibrer. Les valeurs choisies sont les suivantes :

$$\omega_{\text{gradient}} = 60.0 ; \omega_{\text{couleur}} = 0.2 \rightarrow 0.66 ; \omega_{\text{lissage}} = 0.36 ; \omega_{\text{guidage}} = 0.02 ; \omega_{\text{ballon}} = 0.24$$

L'évolution du contour, lorsqu'elle se fait à partir de l'intérieur de la feuille, va alors ressembler à

une extension poussée par le ballon jusqu'à s'arrêter sur un fort gradient ou à atteindre une couleur vraiment trop distante. L'énergie de lissage assure que cette évolution se fasse de façon régulière et en particulier empêche la création de chemins très étroits qui permettraient de passer d'une feuille à sa voisine. Ensuite, au fur et à mesure que le coefficient $\omega_{couleur}$ augmente, le contour va avoir tendance à se rétracter vers les zones dont les couleurs sont celles de la feuille, tout en restant accroché sur les forts gradients lorsqu'il s'y est arrêté. Des exemples de cette évolution dans différents contextes sont données dans la Figure 4.24.

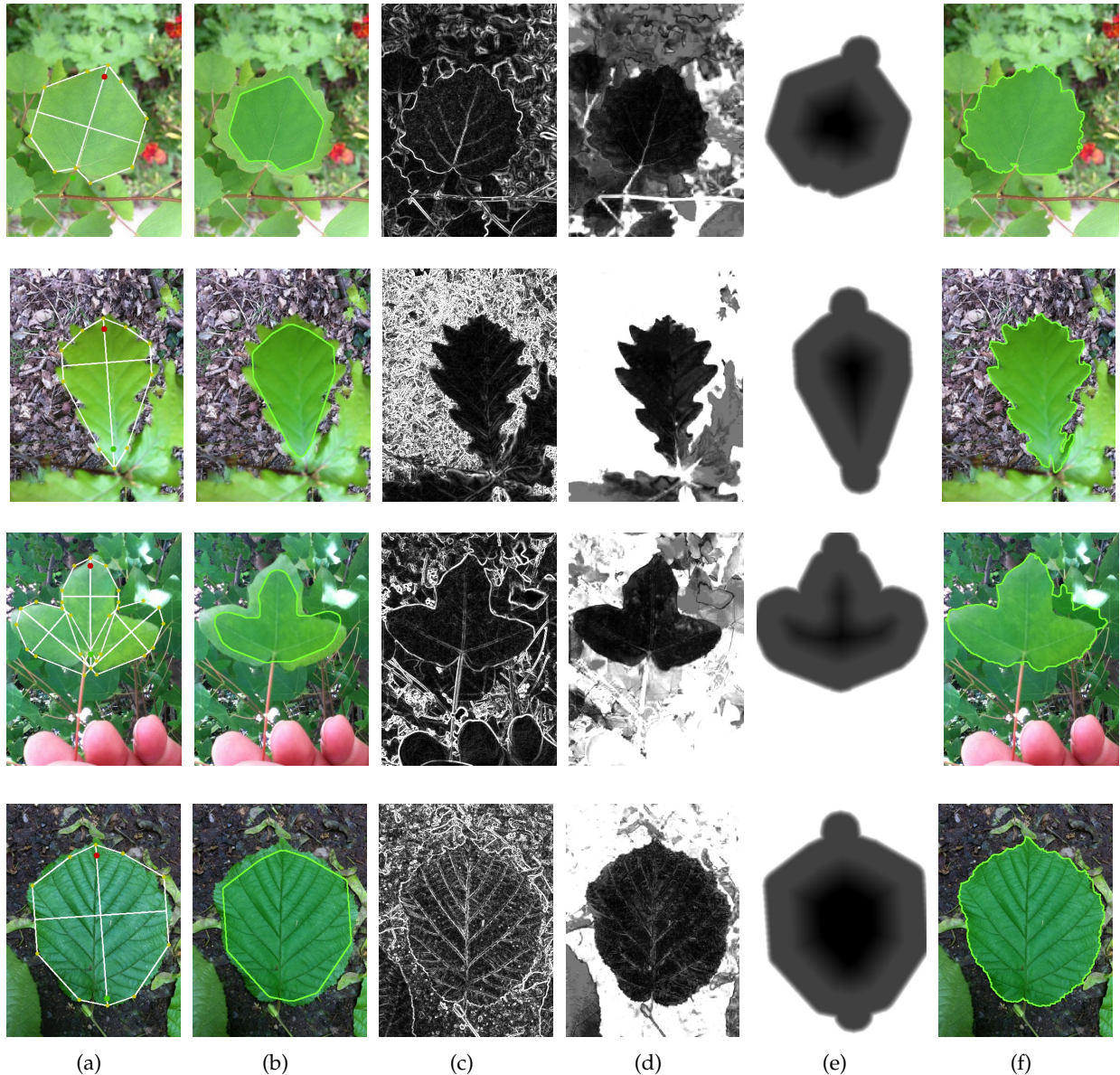


FIGURE 4.24 – Segmentation par contour actif guidé : modèle polygonal obtenu à la fin de la première phase (a), initialisation du contour (b), gradient de l'image (c), carte de dissimilarité de couleur (d), fonction de guidage du contour (e) et région segmentée finale (f)

4.4.3 Contours guidés multiples de folioles

Dans le cas des feuilles composées, il serait également souhaitable d'arriver à obtenir le contour exact de la feuille, de façon à analyser les propriétés morphologiques des folioles. Mais plutôt que d'extraire un seul contour, autant essayer de se servir de l'individualisation des folioles permise par la modélisation de façon à obtenir un contour par foliole modélisée, ce qui permettra une description plus fine. L'idée est donc encore de déformer le contour polygonal Π_f^* obtenu pour chacune des n_F^* folioles tout en l'utilisant comme contrainte pour un contour actif guidé.

Chaque foliole f non supprimée au cours de l'évolution du modèle joint va ainsi être représentée par un contour Γ_f , ou plus exactement par la région induite $\Omega(\Gamma_f)$, toujours dans une représentation de type *level-set* binaire étendue au cas de régions multiples [BW06], où l'on fait évoluer une fonction binaire implicite ϕ_f par région. Les contours vont, de la même manière que pour une feuille simple, être initialisés sur une version réduite du polygone $\Pi_{\triangleright f}^*$, et rester guidés au cours de leur évolution par la fonction γ_{forme} calculée à partir d'une version étendue de leur modèle polygonal $\Pi_{\downarrow f}^*$, en minimisant la même énergie que précédemment.

En particulier, les vitesses sont calculées de la même façon pour chacune des régions, en utilisant le gradient local $\vec{\nabla} \phi_f$ ou le laplacien de la fonction associée à la région $\Delta \phi_f$. Pour chacun des n_F^* contours, on peut donc calculer une vitesse différente :

$$\begin{aligned}
v_{\phi_f}(x, y) = & \omega_{gradient} \left(\vec{\nabla} \left\| \vec{\nabla} I(x, y) \right\| \cdot \vec{\nabla} \phi_f(x, y) \right) \\
& + \omega_{couleur} d_{M, Lab}(x, y, \Omega_0) \left\| \vec{\nabla} \phi_f \right\| \\
& - \omega_{lissage} \Delta \phi_f(x, y) \left\| \vec{\nabla} \phi_f \right\| \\
& + \omega_{guidage} \gamma_{forme}(x, y, \Pi_{\downarrow f}^*) \left\| \vec{\nabla} \phi_f \right\| \\
& - \omega_{ballon} \left\| \vec{\nabla} \phi_f \right\|
\end{aligned} \tag{4.40}$$

D'un point de vue de l'implémentation, on préfère manipuler, plutôt que n_F^* fonctions implicites binaires, une seule fonction $\bar{\phi}_{1..n_F^*}$ à valeurs dans $\{-1; 1; \dots; n_F^*\}$, tout en gardant la possibilité d'évaluer localement les grandeurs dépendant des ϕ_f . Cela présente en plus l'avantage de s'assurer qu'un pixel de l'image ne pourra appartenir qu'à une seule et unique région, du fait même de la définition de la fonction $\bar{\phi}_{1..n_F^*}$ qui ne peut associer à tout point $(x; y)$ qu'une valeur. Cela permet donc, par construction, de pousser les contours à s'auto-contraindre, et donc de s'assurer que les débordements entre folioles seront maîtrisés.

Au final on obtient un ensemble de régions disjointes, dont l'évolution est rendue robuste à la fois par le guidage via la forme polygonale et par l'interaction implicite avec les autres régions. La feuille n'est alors pas vue comme une unique région mais comme plusieurs régions indépendantes, ce qui permet une plus grande interprétabilité, et une meilleure description par le choix des régions les plus pertinentes. Il s'agit donc d'une segmentation avec un contenu sémantique ajouté, dont quelques résultats peuvent être vus sur la Figure 4.25.

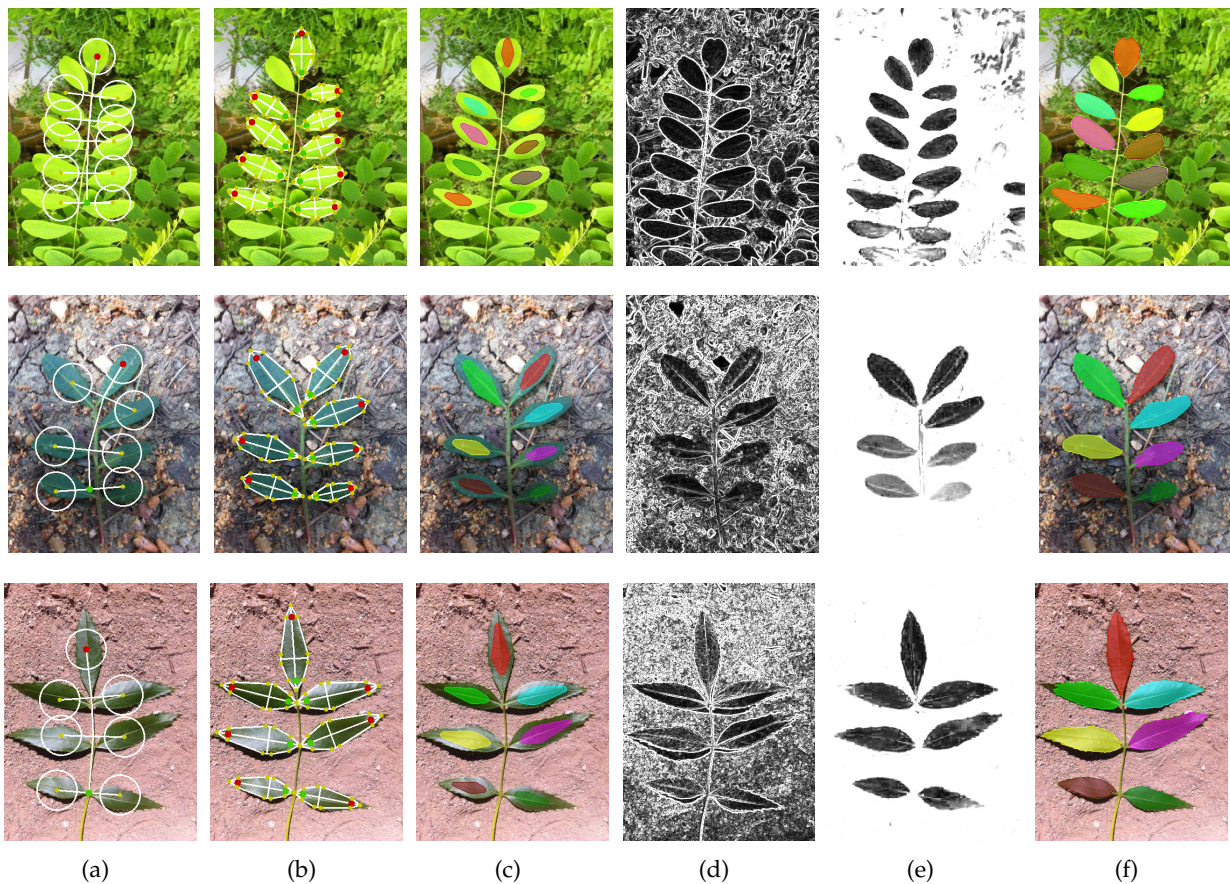


FIGURE 4.25 – Segmentation par contours actifs guidés multiples : modèle de feuille composée (a) modèles polygonaux de folioles obtenus à la fin de la première phase (b), initialisation du contour (c), gradient de l’image (d), carte de dissimilarité de couleur (e) et régions segmentées finales (f)

A la fin de cette phase de segmentation, on obtient de manière générale une région Ω^* , qui peut éventuellement être subdivisée en plusieurs régions indépendantes $\Omega^* = \bigsqcup_f \Omega_f^*$. C’est sur ces régions que l’on va chercher à effectuer des mesures pour décrire la forme de la feuille, et tout est donc mis en place pour que les segmentations que l’on obtient soient les plus proches possible de la région Ω_{feuille} délimitée par la feuille.

La mise en place de contraintes comme *l’a priori* de forme globale de feuille inclus dans la modélisation, puis la contrainte de guidage des contours actifs, a pour but de limiter les risques de résultats aberrants, parfaitement envisageables étant donné la complexité des images en entrée. Là où certaines méthodes moins contraintes échoueraient, nous nous retrouvons généralement avec une région qui, si elle a parfois du mal à s’attacher au contour exact de la feuille à cause d’ombres, de reflets, ou de feuilles voisines, est globalement similaire à quelques artefacts près à la forme de la feuille. Dans certains cas, l’influence du guidage peut être négative, lorsqu’elle empêche par exemple le contour d’atteindre l’extrémité de lobes ou de dents trop prononcés, mais c’est à ce prix que l’on parvient à une robustesse certaine de la segmentation.

4.5 Résultats et performances de segmentation

Pour évaluer de manière plus précise la performance de notre méthode de segmentation par contours actifs guidés, nous avons essayé d'en mesurer les résultats en comparant les segmentations obtenues avec des vérités terrains créées en segmentant manuellement des images de feuilles sur fond naturel. La base utilisée pour cette évaluation est un sous-ensemble de 232 photographies de la base **Pl@ntLeaves** [GBJ⁺11], constituée de manière participative par les contributeurs du réseau botanique francophone TelaBotanica. Les images retenues de la Figure 4.26 sont des photographies de feuilles simples éventuellement palmatilobées, où une feuille d'intérêt peut être clairement désignée, et sur un fond naturel assez complexe.

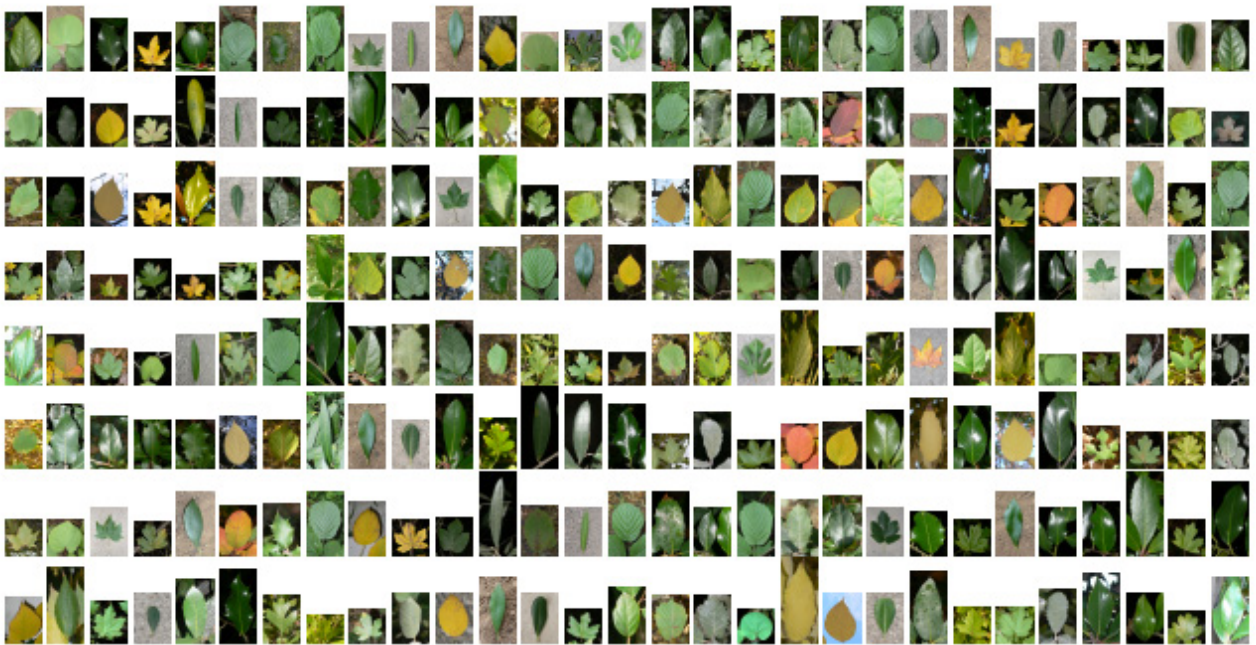


FIGURE 4.26 – Base de photographies utilisées pour évaluer les performances de segmentation

De façon à valider la pertinence de notre approche, nous avons non seulement évalué différents critères de qualité de segmentation sur les résultats de notre méthode, mais également implémenté ou utilisé des versions existantes d'autres méthodes de segmentation s'appliquant au même type d'images. En calculant aussi les mêmes critères pour ces approches différentes, on constitue une base solide de comparaison permettant de tirer des enseignements quant au bien-fondé des parti-pris de la méthode des contours actifs guidés.

4.5.1 Mesures de qualité

L'évaluation de la qualité d'une segmentation peut s'effectuer de manière non-supervisée par la mesure de grandeurs liées à l'image sur la région Ω^* , notamment du point de vue du contraste ou de l'homogénéité, des critères objectifs permettant de définir une "bonne" segmentation dans un cadre assez général. Cela semble néanmoins nettement insuffisant dans notre cadre particulier, et on a vu que l'homogénéité de la feuille était un critère assez difficile à définir. De telles mesures de qualité ne semblent donc pas appropriées pour l'objet de notre étude.

Par conséquent, l'examen de la qualité de la segmentation doit être réalisé de manière supervisée, c'est-à-dire en associant aux images traitées une segmentation de référence fournie par un expert. Autrement dit, la mesure de qualité se base pour chaque image sur la comparaison entre la région segmentée Ω^* et la vérité terrain Ω_{ref} représentant la région réellement délimitée par la feuille dans l'image.

De nombreuses méthodes existent alors pour comparer deux régions binaires et évaluer leur ressemblance ou leur proximité. En voyant la segmentation comme la recherche des pixels appartenant effectivement à Ω_{ref} , et la région Ω^* comme le résultat de cette recherche, certaines mesures statistiques très courantes tirées de la recherche d'information vont alors pouvoir être utilisées. En notant $|\Omega|$ le cardinal d'une région Ω , on parlera alors de :

- ▶ **Vrais Positifs** : pour les pixels de Ω_{ref} bien présents dans Ω^* : $VP = |\Omega^* \cap \Omega_{\text{ref}}|$
- ▶ **Faux Positifs** : pour les pixels absents de Ω_{ref} mais présents dans Ω^* : $FP = |\Omega^* \cap \overline{\Omega_{\text{ref}}}|$
- ▶ **Faux Négatifs** : pour les pixels de Ω_{ref} qui n'ont pas été retrouvés dans Ω^* : $FN = |\overline{\Omega^*} \cap \Omega_{\text{ref}}|$
- ▶ **Vrais Négatifs** : pour les pixels absents de Ω_{ref} et également absents de Ω^* : $VN = |\overline{\Omega^*} \cap \overline{\Omega_{\text{ref}}}|$

La Figure 4.27 donne une idée de l'interprétation de ces indices pour évaluer une segmentation.

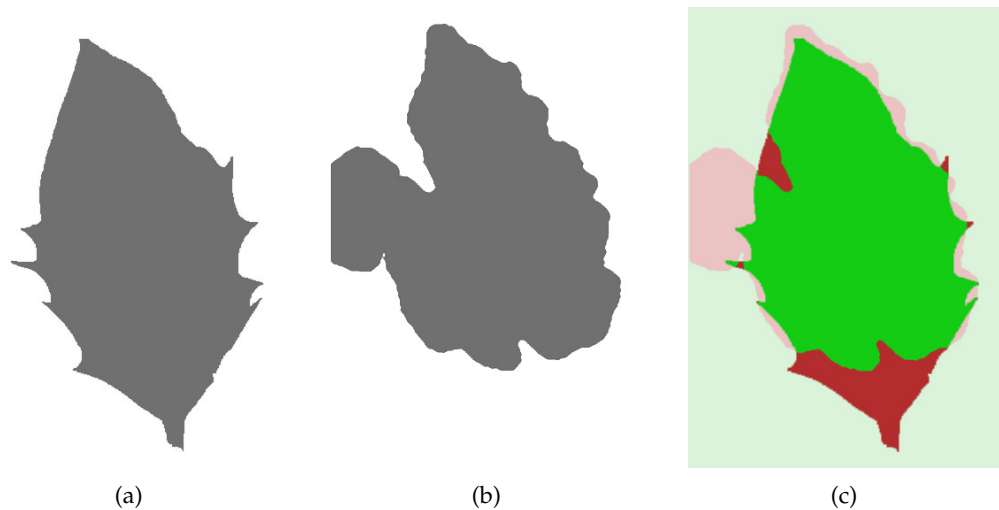


FIGURE 4.27 – Indices de base utilisés pour comparer une vérité terrain (a) et une région segmentée (b) : ■ vrais positifs, ■ vrais négatifs, ■ faux positifs, ■ faux négatifs (c)

A partir de ces mesures de base, on peut définir un certain nombre d'indices permettant de mesurer la qualité d'une segmentation, du point de vue de la simple correspondance avec le résultat attendu. Aucun de ces indices n'est exactement représentatif en lui-même, et la qualité est plutôt agréée par l'ensemble de leurs valeurs que par une valeur individuelle. Ils partagent en revanche tous la caractéristique d'être maximaux lorsque la segmentation est optimale c'est à dire lorsque $\Omega^* = \Omega_{\emptyset}$:

- **Précision** : rapport entre le nombre de vrais positifs et le nombre de positifs (maximal dès lors que $\Omega^* \subseteq \Omega_{\emptyset}$)

$$\text{precision}(\Omega^*, \Omega_{\emptyset}) = \frac{VP}{VP + FP} = \frac{|\Omega^* \cap \Omega_{\emptyset}|}{|\Omega^*|}$$

- **Rappel** ou **Sensibilité** : rapport entre le nombre de vrais positifs et le nombre de pixels à retrouver (maximal dès lors que $\Omega_{\emptyset} \subseteq \Omega^*$)

$$\text{rappel}(\Omega^*, \Omega_{\emptyset}) = \frac{VP}{VP + FN} = \frac{|\Omega^* \cap \Omega_{\emptyset}|}{|\Omega_{\emptyset}|}$$

- **Spécificité** : rapport entre le nombre de vrais négatifs et le nombre de pixels extérieurs (maximal dès lors que $\overline{\Omega_{\emptyset}} \subseteq \overline{\Omega^*}$ donc $\Omega^* \subseteq \Omega_{\emptyset}$)

$$\text{specificite}(\Omega^*, \Omega_{\emptyset}) = \frac{VN}{VN + FP} = \frac{|\overline{\Omega^*} \cap \overline{\Omega_{\emptyset}}|}{|\overline{\Omega_{\emptyset}}|}$$

- **Recouvrement** ou **Indice de Jaccard** : rapport entre le nombre de vrais positifs et la réunion de pixels positifs et à retrouver (maximal uniquement lorsque $\Omega^* = \Omega_{\emptyset}$)

$$\text{jaccard}(\Omega^*, \Omega_{\emptyset}) = \frac{VP}{VP + FP + FN} = \frac{|\Omega^* \cap \Omega_{\emptyset}|}{|\Omega^* \cup \Omega_{\emptyset}|}$$

- **F-mesure** ou **Indice de Dice** : rapport entre deux fois le nombre de vrais positifs, et la somme du nombre de positifs et de pixels à trouver (maximal uniquement lorsque $\Omega^* = \Omega_{\emptyset}$)

$$\text{dice}(\Omega^*, \Omega_{\emptyset}) = \frac{2VP}{2VP + FP + FN} = \frac{2|\Omega^* \cap \Omega_{\emptyset}|}{|\Omega^*| + |\Omega_{\emptyset}|} = 2 \times \frac{\text{precision}(\Omega^*, \Omega_{\emptyset}) \times \text{rappel}(\Omega^*, \Omega_{\emptyset})}{\text{precision}(\Omega^*, \Omega_{\emptyset}) + \text{rappel}(\Omega^*, \Omega_{\emptyset})}$$

- **Indice de Manhattan** : rapport entre le nombre pixels corrects, et le nombre total de pixels (maximal uniquement lorsque $\Omega^* = \Omega_{\emptyset}$)

$$\text{manhattan}(\Omega^*, \Omega_{\emptyset}) = \frac{VP + VN}{VP + FP + FN + VN} = \frac{|\Omega^* \cap \Omega_{\emptyset}| + |\overline{\Omega^*} \cap \overline{\Omega_{\emptyset}}|}{|I|}$$

Cette quantité est à mettre en lien avec la **distance de Hamming** qui comptabilise les différences

entre deux ensembles binaires :

$$d_{\text{Hamming}}(\Omega^*, \Omega_{\emptyset}) = |\Omega^* \cap \overline{\Omega_{\emptyset}}| + |\overline{\Omega^*} \cap \Omega_{\emptyset}| = |\Omega^* \sqcup \Omega_{\emptyset}| = FP + FN$$

$$\text{manhattan}(\Omega^*, \Omega_{\emptyset}) = 1 - \frac{1}{|I|} d_{\text{Hamming}}(\Omega^*, \Omega_{\emptyset})$$

Il est important de prendre en compte plusieurs de ces indices pour évaluer une segmentation, étant donné que plusieurs éléments interviennent dans la qualité. Les 4 grandeurs VP , FP , FN et VN sont importantes pour caractériser une méthode, et sa propension à sur-segmenter ou sous-segmenter, à se montrer robuste ou risquée. Etant donnée que leur somme est constante, la donnée de 3 de ces grandeurs suffit et on pourrait donc considérer que la qualité d'une segmentation est le mieux évaluée par une donnée tri-dimensionnelle par exemple (precision; rappel; spécifité). On peut donc voir tous les indices scalaires donnés ici comme des moyens de mesurer à qu'elle point le point 3D caractérisant la qualité de la segmentation se rapproche du point idéal correspondant à $\Omega^* = \Omega_{\emptyset}$, mais sans dire comment elle s'en approche. Il est donc pertinent d'en garder un certain nombre pour mieux interpréter les valeurs obtenues.

Toutefois, ces mesures prennent en compte la région dans sa globalité, et au final un poids très important est donné à l'intérieur de la région, qui n'apporte que peu d'information en lui-même. Il pourrait être tout aussi intéressant de considérer la similarité entre les contours $C(\Omega^*)$ et $C(\Omega_{\emptyset})$ de façon à mesurer les écarts entre ce qui permet véritablement de définir les régions. Dans notre cadre spécifique, le bord des feuilles présentant des formes très caractéristiques des espèces, il est d'autant plus important que ce soit le contour qui corresponde au bord réel de la feuille, bien plus que le fait que la région recouvre majoritairement celle de la feuille.

Nous utilisons donc deux mesures supplémentaires qui permettent de caractériser la dissimilarité des contours, et pénalisent nettement plus les écarts entre les pixels de $C(\Omega^*)$ et $C(\Omega_{\emptyset})$ que ne le font les mesures basées sur la taille des intersections de régions. Elles utilisent notamment la distance d'un pixel $(x; y)$ à une région Ω que l'on note $\|(x; y) - \Omega\| = \min_{(x'; y') \in \Omega} \|(x; y) - (x'; y')\|$.

- **Distance de Hausdorff** : écart maximal entre les deux régions ; bien que définie sur les régions, elle se ramène à une distance entre contours, la distance maximale étant trivialement atteinte pour les points extrêmes des régions soit leurs contours (minimale si et seulement si $\Omega^* = \Omega_{\emptyset}$)

$$\begin{aligned} d_{\text{Hausdorff}}(\Omega^*, \Omega_{\emptyset}) &= \max \left(\max_{(x; y) \in \Omega^*} \|(x; y) - \Omega_{\emptyset}\|, \max_{(x; y) \in \Omega_{\emptyset}} \|(x; y) - \Omega^*\| \right) \\ &= \max \left(\max_{(x; y) \in C(\Omega^*)} \|(x; y) - C(\Omega_{\emptyset})\|, \max_{(x; y) \in C(\Omega_{\emptyset})} \|(x; y) - C(\Omega^*)\| \right) \end{aligned}$$

- **Distance absolue moyenne** (*mean absolute distance*) : écart moyen entre les deux contours (minimale si et seulement si $\Omega^* = \Omega_{\emptyset}$)

$$d_{\text{MAD}}(\Omega^*, \Omega_{\emptyset}) = \max \left(\frac{1}{|C(\Omega^*)|} \sum_{(x;y) \in C(\Omega^*)} \|(x;y) - C(\Omega_{\emptyset})\|, \frac{1}{|C(\Omega_{\emptyset})|} \sum_{(x;y) \in C(\Omega_{\emptyset})} \|(x;y) - C(\Omega^*)\| \right)$$

Ces mesures permettent d’avoir une idée de la différence entre le contour de la région segmentée et celui de la région attendue, la première pénalisant très fortement tout écart un peu trop prononcé, et la seconde permettant de quantifier cette différence sur toute la longueur du contour.

L’ensemble de ces indices de qualité permet de fournir une description assez complète de l’adéquation d’une segmentation avec une vérité terrain, prenant à la fois en compte la région dans sa globalité et des différences plus fines au niveau du contour. Prises dans leur ensemble, il s’agit surtout d’un excellent moyen de comparaison entre différentes méthodes de segmentation qui donne une bonne idée des avantages et des points faibles des approches comparées.

4.5.2 Méthodes implémentées pour comparaison

Dans le cadre de l’évaluation de notre méthode de segmentation, nous avons implémenté ou utilisé l’implémentation de divers algorithmes destinés à produire des résultats comparables, mais généralement sans prise en compte d’une connaissance antérieure sur la forme. Certains sont dédiés aux images complexes et même utilisés dans le cadre de la segmentation d’images végétales, d’autres moins appropriés. Pour rendre les choses équitables, nous avons quand cela était possible utilisé la même modélisation de la couleur de l’image qu’introduite dans la Section 5.2, de façon à ce que toutes les méthodes aient le même point de départ.

Les algorithmes que nous avons retenus sont ceux qui nécessitaient le moins d’adaptation fine de leurs paramètres, et par conséquent qui étaient les plus à même de fonctionner correctement sur la variété relativement importante d’images contenues dans la base Pl@ntLeaves. Dans le cas où certains paramètres devaient tout de même être optimisés, nous avons fait en sorte de les ajuster de façon à obtenir des performances visuellement les plus satisfaisantes possibles sur des images difficiles. Les approches choisies couvrent un large spectre, allant d’approches globales à des algorithmes basés contour ou région, exploitant l’information de gradient ou de couleur. Au total, nous avons produit des résultats de segmentation à l’aide de 9 méthodes :

- ▶ **Seuillage d’Otsu** [Ots79] : méthode globale affectant uniformément les pixels à la région Ω^* ou non en une seule passe, en se basant sur la définition d’un seuil optimal sur l’histogramme des niveaux de gris. Cette méthode a été adaptée pour traiter comme source la carte de dissimilarité produite par la modélisation de la couleur.
- ▶ **Watershed** [BM92] : méthode basée région où l’image est vue comme un relief, et la région Ω^* comme un bassin versant calculé en simulant une montée des eaux, utilisant donc notamment l’information de gradient. Cette méthode peut utiliser la modélisation de la couleur en s’appliquant sur le gradient de la carte de dissimilarité.
- ▶ **Mean Shift** [CM99] : méthode s’apparentant à un *clustering*, basée sur un filtrage de l’image

pour déterminer des modes locaux de couleur ensuite utilisés pour regrouper les pixels dans le domaine joint de la couleur et de l'espace. Cette méthode peut également être appliquée aux cartes de dissimilarité de façon à inclure la modélisation de la couleur.

- ▶ **Mean Shift Pyramidal** [LXGM10] : algorithme dérivé du précédent, mais où le filtrage est appliqué à une pyramide d'images de résolutions différentes, de façon à lisser l'influence de certains détails gênants. De la même façon, cette méthode peut s'appuyer sur les cartes de dissimilarité de couleur.
- ▶ **Graph-Cuts** [BJ01] : méthode où l'image est représentée par un graphe dont les pixels sont les sommets et les arêtes pondérées par la distance entre leurs couleurs, et où la segmentation s'apparente à la recherche d'une coupe optimale dans le graphe. La carte de dissimilarité de couleur peut également remplacer l'image comme entrée de la méthode.
- ▶ **GrabCut** [RKB04] : méthode également basée sur les coupes de graphes, mais reposant plus sur l'information de couleur (par conséquent la dissimilarité comme base n'est pas suffisante) et permettant une interactivité avec l'utilisateur. La méthode pourra bénéficier d'une initialisation manuelle.
- ▶ **Segmentation basée Graphe** [Fel04] : méthode basée sur la même représentation de l'image par un graphe mais où les composantes segmentées croissent de manière itérative en utilisant un critère de décision basé sur la différence minimale entre les régions. Cette méthode peut bénéficier de notre modélisation de la couleur et s'appliquer sur la carte de dissimilarité.
- ▶ **Contour Actifs** [CV01] : méthode de contours évoluant de façon à minimiser une énergie basée sur l'homogénéité en termes de couleurs des régions délimitées et sur un terme de lissage. En se basant sur la carte dissimilarité, l'énergie minimisée est alors sensiblement la même que notre modèle. Cette méthode est sensible à la position initiale et le contour pourra donc être initialisé manuellement.
- ▶ **B-Spline Snakes** [BHU00] : méthode fonctionnant de la même manière que la précédente, mais où le contour est formé de morceaux de courbes paramétriques, et donc plus régulier et plus rigide. L'évolution peut également se baser sur la dissimilarité pour arriver à une formulation de l'énergie équivalente à la nôtre. De la même manière une initialisation manuelle du contour permettra une plus grande robustesse.

Pour les méthodes qui le permettent, ainsi que pour notre méthode, nous avons donné la possibilité de démarrer à partir d'une région initiale Ω_0 sous forme d'une marque continue laissée par un utilisateur. Cette initialisation correspond au genre d'interaction qu'il est possible de demander à l'utilisateur d'une application pour smartphone disposant d'une interface tactile. Dans tous les cas, une initialisation automatique est utilisée, sous forme d'un polygone initial Π_0 placé au centre de l'image, et construit selon le modèle polygonal de feuille simple. Cette initialisation fait des suppositions sur la position de feuille dans l'image qui ne sont pas toujours rigoureusement vérifiées et comporte donc une part de risque. Il faut donc s'attendre à ce que les performances soient toujours meilleures avec

une initialisation manuelle qui élimine le risque d’avoir une région initiale qui déborde de la feuille, comme cela peut se produire avec une initialisation automatique (Figure 4.28).



FIGURE 4.28 – Différentes initialisations possibles des méthodes de segmentation, automatique par un polygone centré, ou manuelle par une trace d’utilisateur

Les résultats fournis par l’ensemble de ces méthodes, avec ou sans initialisation manuelle, constituent alors une masse de données importante qui devrait permettre de situer la performance de notre algorithme de segmentation sur un ensemble d’images pour lesquelles il est dédié, par rapport à un vaste panel de l’état de l’art.

4.5.3 Résultats et comparaisons

Mesurer la qualité des segmentations permet, avant même de comparer les méthodes entre elles, de valider notre approche de façon numérique, en particulier le choix de la modélisation de la couleur. On peut ainsi comparer les valeurs moyennes prises par les différents indices de qualité sur l’ensemble des 232 images de la base, de façon à observer quelle fonction de dissimilarité de couleur donne globalement les meilleurs résultats. Nous comparons donc les segmentations obtenues par évolution du modèle polygonal et contour actif guidé de différentes manières :

- ▶ d_{LinReg} : les fonctions d’énergies utilisent uniquement la dissimilarité au modèle global de couleur
- ▶ d_{Locale} : les fonctions d’énergies utilisent uniquement la dissimilarité à la moyenne locale attendue
- ▶ $d_{Croyance}$: les fonctions d’énergies utilisent la combinaison de ces deux dissimilarités obtenue par fonctions de croyance.

De plus on compare une initialisation automatique par un polygone centré Π_0 avec celle faite manuellement par un utilisateur laissant une trace Ω_0 . Les résultats de cette première analyse sont donnés dans la Table 4.1 pour les indices globaux, à valeur constamment entre 0 et 1, et dont la moyenne sur la base peut être plus facilement interprétée.

De façon à mieux visualiser ces résultats, il est également possible de tracer ces mesures sous formes de courbes, représentant l’histogramme des fréquences cumulées décroissantes de chaque mesure. Autrement dit, pour chaque valeur d’un indice à valeur entre 0 et 1 on peut donner la proportion d’images de la base atteignant au moins cette valeur pour l’indice considéré. La courbe idéale étant constante à 1 (la valeur maximale de l’indice atteinte pour 100% des exemples), plus la

Dissimilarité	Précision	Rappel	Spécificité	Dice	Jaccard	Manhattan
$d_{LinReg} + \Pi_0$	0.951	0.915	0.974	0.928	0.874	0.950
$d_{Locale} + \Pi_0$	0.948	0.916	0.974	0.927	0.872	0.952
$d_{Croyance} + \Pi_0$	0.963	0.927	0.980	0.942	0.895	0.960
$d_{Croyance} + \Omega_0$	0.967	0.947	0.982	0.955	0.917	0.968

TABLE 4.1 – Comparaisons des performances de segmentation des contours actifs guidés obtenus avec différentes mesures de dissimilarité de couleur

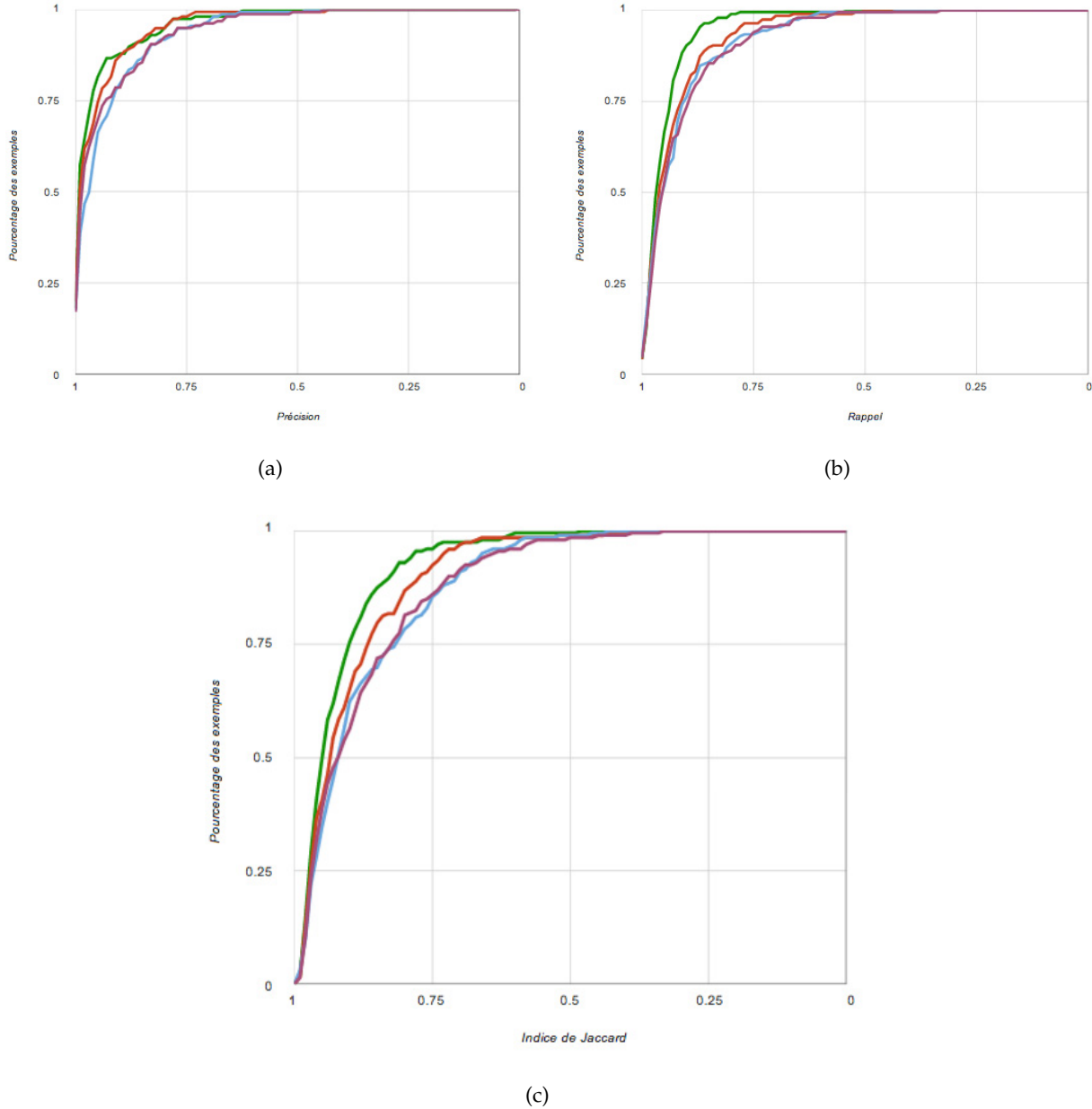


FIGURE 4.29 – Courbes de performance comparées des contours actifs guidés pour la précision (a), le rappel (b) et l'indice de Jaccard (c) : \blacksquare $d_{Croyance} + \Omega_0$, \blacksquare $d_{Croyance} + \Pi_0$, \blacksquare $d_{Locale} + \Pi_0$, \blacksquare $d_{LinReg} + \Pi_0$

courbe se rapproche du point $(1;1)$, mieux se comporte la méthode. La Figure 4.29 donne par exemple ces courbes pour les différentes options de notre méthode de segmentation sur la précision, le rappel, et l'indice de Jaccard.

Le bénéfice de la combinaison de deux modélisations de la couleur, par ailleurs satisfaisantes en elles-mêmes, apparaît ici clairement. Ce que tendent à montrer ces résultats est que l'amélioration produite par l'utilisation de $d_{croyance}$, visible sur tous les indices, est principalement due à un gain en précision (nettement apparent sur l'écart entre la courbe de $d_{croyance}$ et les autres courbes de précision de la Figure 4.29 (a)). Cela signifie que combiner les deux dissimilarités permet de mieux éliminer des pixels de l'arrière-plan qu'en les prenant séparément. Ceci est logique, puisque chacune agit comme garde-fou de la propension de l'autre à considérer certains pixels extérieurs comme proches.

L'apport de l'initialisation manuelle se traduit quant à lui bien plus nettement en termes de rappel (Figure 4.29 (b)). En effet, la définition de la région initiale étant généralement plus large que celle de Π_0 , plus de couleurs vont entrer dans la modélisation de la couleur de la feuille, qui sera capable de mieux couvrir l'ensemble de la surface de la feuille. Même si l'amélioration en précision est peu visible (l'arrière-plan est toujours aussi distant) l'écart en rappel (la baisse du nombre de faux négatifs) permet une amélioration nette des indices globaux.

La performance en termes de segmentation valide donc les choix faits du point de vue de la modélisation de la couleur, et notamment de la combinaison qui semble ainsi bénéficier des avantages de chacune des deux visions, locale et globale, des couleurs de la feuille. De manière assez évidente, mais qu'il est toujours bon de confirmer par les chiffres, l'utilisateur apporte une véritable aide pour la segmentation (que l'on qualifiera plus de supervisée que d'interactive) et plus généralement pour l'ensemble processus.

Pour comparer alors les méthodes de segmentation à proprement parler, nous évaluons les différents indices de qualité sur les segmentations des 232 images obtenues en utilisant les divers algorithmes implémentés, et en prenant en compte autant que possible la même modélisation de la couleur, de façon à évaluer uniquement la méthode. Les résultats de ces évaluations sont présentées dans la Table 4.2 sous forme de valeurs moyenne sur l'ensemble de la base.

Méthode	Précision	Rappel	Spécificité	Dice	Jaccard	Manhattan	$d_{\text{Hausdorff}}$	d_{MAD}
Otsu [Ots79]	0.848	0.951	0.888	0.887	0.815	0.909	48.72	6.69
Watershed [BM92]	0.778	0.963	0.811	0.840	0.752	0.858	62.66	6.13
MeanShift [CM99]	0.794	0.957	0.834	0.850	0.767	0.874	61.42	6.31
MS Pyramidal [LXGM10]	0.836	0.957	0.867	0.878	0.803	0.901	51.08	5.77
Graph-Cuts [BJ01]	0.744	0.934	0.798	0.812	0.707	0.845	63.13	6.81
GrabCut [RKB04]	0.835	0.979	0.868	0.884	0.821	0.902	41.56	7.16
Felzenszwalb [Fel04]	0.882	0.826	0.911	0.821	0.733	0.877	34.57	4.37
Snakes [CV01]	0.788	0.968	0.838	0.854	0.768	0.881	63.57	5.98
BSpline Snakes [BHU00]	0.812	0.970	0.877	0.877	0.789	0.910	30.9	4.20
Contours Actifs Guidés	0.963	0.927	0.980	0.942	0.895	0.960	11.56	1.56

TABLE 4.2 – Comparaisons des performances des différentes méthodes de segmentation initialisées automatiquement

Le principal enseignement est que, sur l'ensemble des mesures de qualité complètes (maximales uniquement pour une segmentation parfaite), la méthode des contours actifs guidés par la modélisation polygonale s'avère supérieure à l'ensemble des méthodes classiques tirées de l'état de l'art. Elle montre notamment une amélioration nette d'au minimum 5% sur ses meilleures concurrentes (seuillage d'Otsu utilisant la carte de dissimilarité, GrabCut et B-Spline snakes) sur les indices de Dice, de Jaccard et de Manhattan. Ce qui apparaît ici est la plus grande robustesse de notre méthode, qui s'appuie sur de fortes contraintes de forme, et se montre donc moins encline à produire des segmentations aberrantes. Des exemples de segmentations produites par les méthodes les plus compétitives sont présentés dans la Figure 4.30.

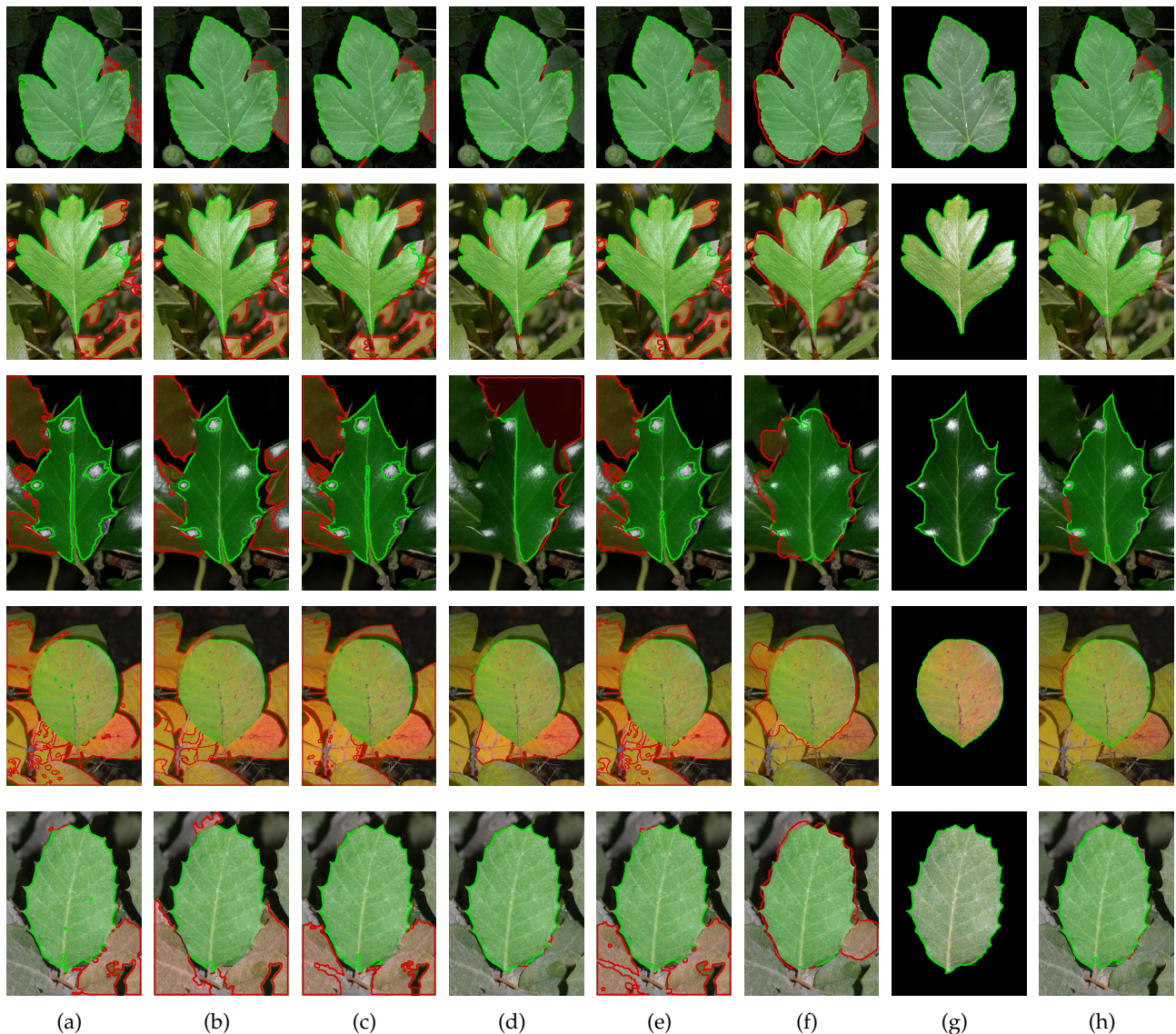


FIGURE 4.30 – Exemples de résultats de segmentation, ■ vrais positifs et ■ faux positifs : seuillage d'Otsu (a), watershed (b), mean shift pyramidal (c), GrabCut (d), contours actifs (e), contours actifs B-spline (f) à comparer avec la vérité terrain (g) et le résultat des contours actifs guidés (h)

Pour affiner cette analyse, on peut observer les performances en termes de précision et de rappel. Les valeurs moyennes permettent de donner une première idée, en voyant que bien que le point où notre méthode se détache le plus est la précision, donc la moindre prise de risque. Elle serait au contraire plutôt en retrait en ce qui concerne le rappel, donc la couverture de l'ensemble de la feuille. Pour mieux visualiser cet écart, nous avons tracé les mêmes courbes de fréquences cumulées que précédemment, en sélectionnant les méthodes les plus compétitives. Les résultats pour la précision et le rappel sont donnés dans la Figure 4.31.

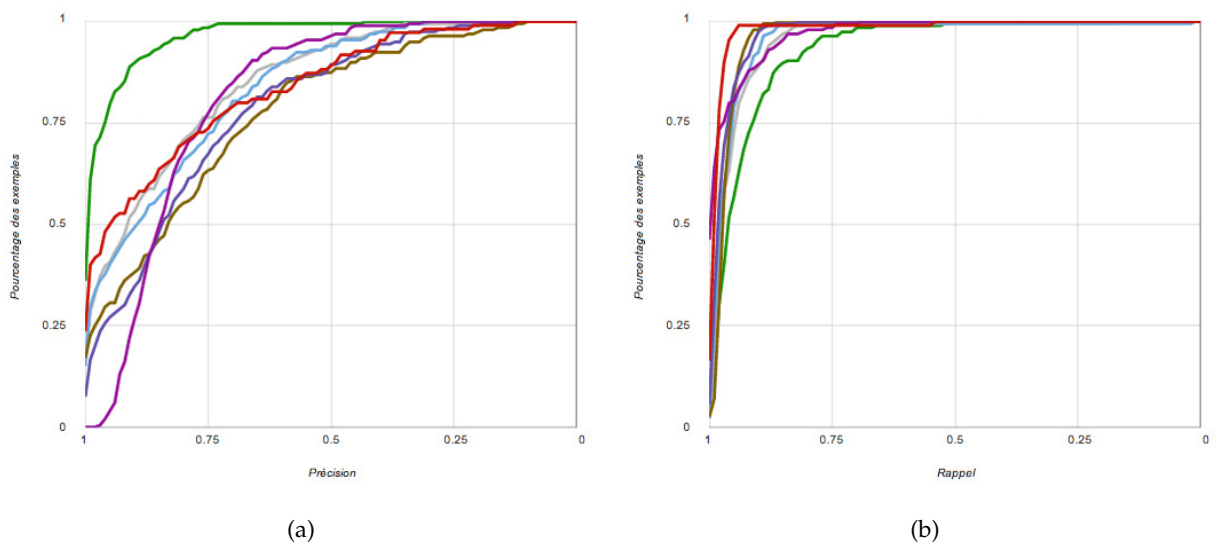


FIGURE 4.31 – Courbes de performance comparées de différentes méthodes de segmentation pour la précision (a) et le rappel (b) : ■ seuillage d'Otsu, ■ watershed, ■ mean shift pyramidal, ■ GrabCut, ■ contours actifs, ■ contours actifs B-spline, ■ contours actifs guidés

La différence en précision est ici criante, et elle souligne la précaution de notre méthode à ne pas déborder de la feuille. Le fait d'évaluer dans un premier temps un modèle polygonal aux bords rigides et de l'utiliser comme limite à l'évolution du contour garantit de limiter les débordements, un problème dont souffrent de nombreuses méthodes moins contraintes. L'influence des contraintes de forme sur le débordement est mise en évidence dans la Figure 4.32. De ce point de vue, la méthode qui s'en sort le mieux est la plus contrainte à savoir celle des B-spline snakes, mais paradoxalement sa faible flexibilité la fait aussi systématiquement déborder de quelques pixels. Ces résultats sont confirmés par les valeurs de la distance de Hausdorff qui pénalise lourdement les débordements et où les méthodes les plus prudentes (B-spline snakes et Felzenswalb) s'en sortent le mieux.

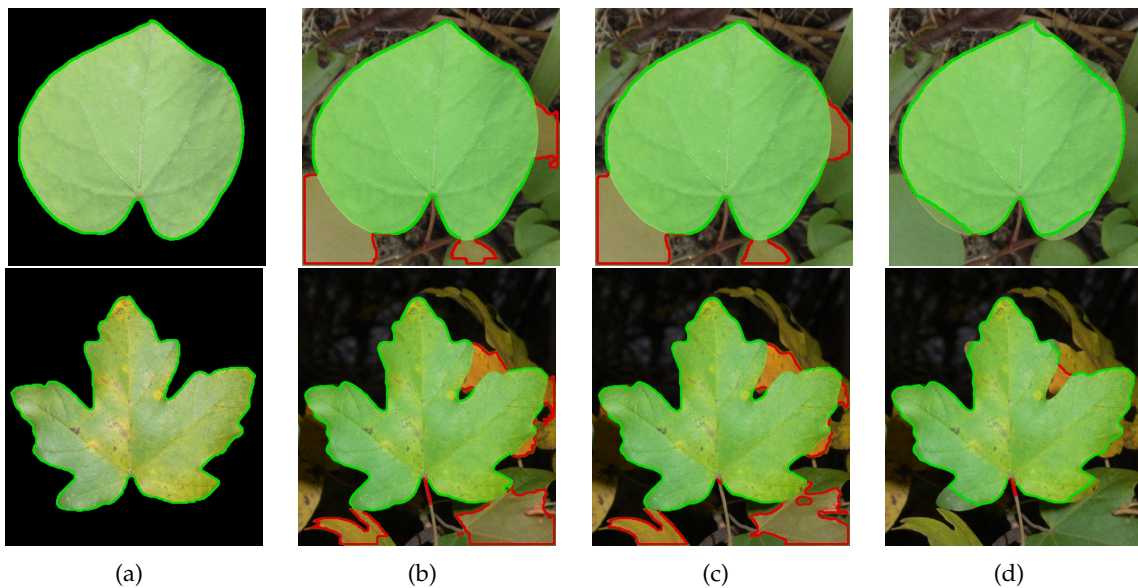


FIGURE 4.32 – Exemples de problèmes de débordement avec les méthodes classiques (mean shift pyramidal (b) et contours actifs (c)) et influence de la contrainte de forme (d)

En ce qui concerne le rappel, notre méthode se classe bonne dernière parmi les méthodes retenues, ce qui n'est pas forcément étonnant. En effet la contrainte de forme peut parfois avoir un effet limitant en empêchant le contour d'aller jusqu'au bord de la feuille, notamment dans le cas de formes très arrondies (Figure 4.32, première ligne) ou de lobes peu symétriques. C'est là le prix à payer pour garantir une plus grande robustesse, et l'écart avec le bord est généralement assez faible comme le montre la performance en distance moyenne (MAD) ; c'est en tout cas un point tout à fait cohérent avec le compromis choisi. Certaines méthodes comme GrabCut privilégient manifestement le rappel en cherchant à retrouver la feuille dans sa totalité mais la contrepartie en est le risque plus élevé de débordement, caractérisé par un nombre important d'images à faible précision.

Si l'on trace maintenant ces courbes pour un indice de qualité comme l'indice de Jaccard (le plus pénalisant en cas de mauvaise segmentation) la méthode des contours actifs guidés apparaît clairement comme la plus robuste (Figure 4.33). Si une méthode comme GrabCut semble produire de meilleures segmentations sur les feuilles faciles (la courbe étant au-dessus pour des valeurs de l'indice proches de 1) notre approche prouve sa supériorité par sa capacité à produire des segmentations satisfaisantes dans la majorité des cas (plus de 80 % des images avec un indice de Jaccard supérieur à 0.8). C'est par exemple ce que ne fait pas la méthode des B-spline snakes où la contrainte forte dans la construction du contour lui permet d'être meilleure que les contours actifs non-contraints, mais l'empêche quasiment de produire des segmentations avec un indice de Jaccard supérieur à 0.9. Il est aussi intéressant de noter la bonne position de la méthode assez basique de seuillage d'Otsu lorsqu'elle est appliquée à une représentation adaptée de la couleur.

Puisque nous avons démontré l'importance de l'initialisation manuelle pour améliorer les performances, il est intéressant de mesurer comment les autres méthodes se comportent lorsqu'on leur

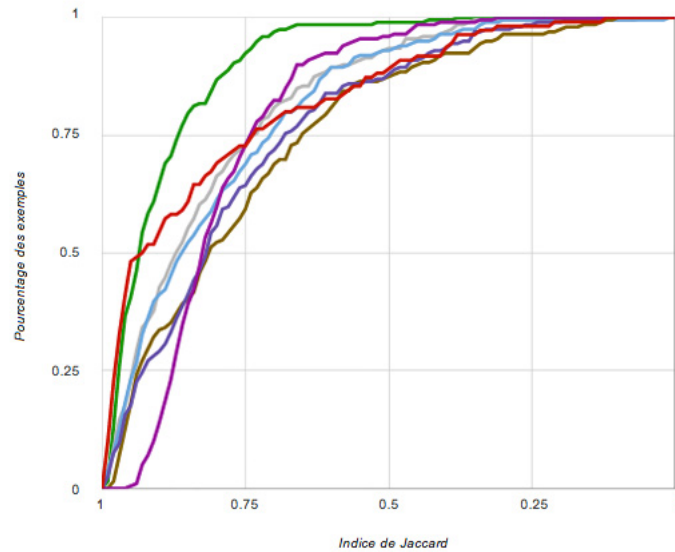


FIGURE 4.33 – Courbes de performance comparées de différentes méthodes de segmentation pour l'indice de Jaccard : ■ seuillage d'Otsu, ■ watershed, ■ mean shift pyramidal, ■ GrabCut, ■ contours actifs, ■ contours actifs B-spline, ■ contours actifs guidés

donne un meilleur point de départ. Les résultats alors obtenus en moyenne pour les différents indices sont donnés dans la Table 4.3, où il apparaît que notre méthode demeure meilleure même lorsque les méthodes concurrentes sont mieux initialisées.

Méthode	Précision	Rappel	Spécificité	Dice	Jaccard	Manhattan	$d_{\text{Hausdorff}}$	d_{MAD}
GrabCut [RKB04]	0.910	0.975	0.930	0.933	0.888	0.945	33.55	6.44
Snakes [CV01]	0.896	0.937	0.930	0.908	0.844	0.932	40.92	5.61
BSpline Snakes [BHU00]	0.803	0.976	0.869	0.874	0.786	0.907	29.13	3.79
Contours Actifs Guidés	0.967	0.949	0.982	0.955	0.917	0.968	10.82	1.06

TABLE 4.3 – Comparaisons des performances des différentes méthodes de segmentation initialisées manuellement

De manière assez surprenante, pour les méthodes moins contraintes comme les contours actifs ou GrabCut, c'est surtout au niveau de la précision que l'amélioration se fait, comme le montre la Figure 4.34. Alors que l'on avait vu que l'initialisation manuelle apportait une baisse du nombre de faux négatifs dans notre méthode, elle semble ici contribuer à limiter les faux positifs. Ce n'est pas le cas pour les B-spline snakes, qui du fait de leur rigidité intrinsèque ne profitent pas tellement d'une initialisation manuelle.

Dans tous les cas, l'initialisation manuelle permet une correction intéressante de la segmentation comme l'illustre la Figure 4.35. Dans le cadre d'une application mobile, sa mise en place via une interaction intuitive semble indispensable pour minimiser les risques de segmentation aberrante. Une méthode comme GrabCut devient dès lors très intéressante, et fournit d'ailleurs plus d'excellentes segmentations que la nôtre, mais elle pêche très nettement dans certains cas par son incapacité à séparer la feuille d'intérêt de celles de l'arrière plan. L'avantage de la méthode des contours actifs

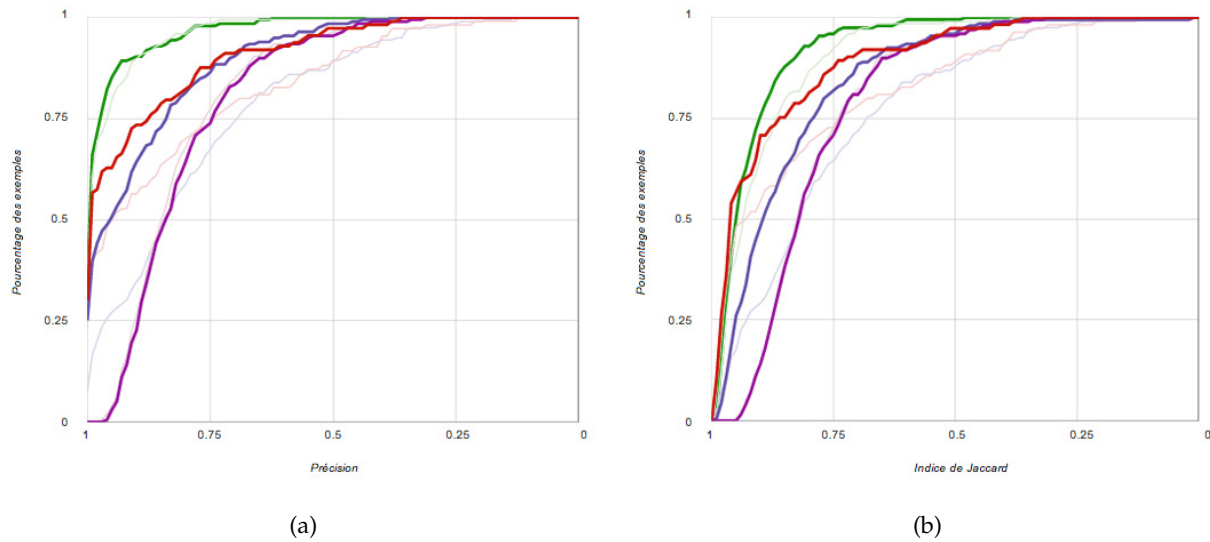


FIGURE 4.34 – Courbes de performance comparées de différentes méthodes de segmentation initialisées manuellement pour la précision (a) l'indice de Jaccard (b), les courbes plus claires représentant la performance de la méthode sans initialisation : ■ GrabCut, ■ contours actifs, ■ contours actifs B-spline, ■ contours actifs guidés

guidés, qui transparaît dans les valeurs des indices et dans les courbes, provient donc pour l'essentiel de sa fiabilité dans la plupart des cas de figures, encore amélioré par une aide de l'utilisateur. La contrainte de forme s'avère être une garantie indispensable de la qualité des segmentations produites.

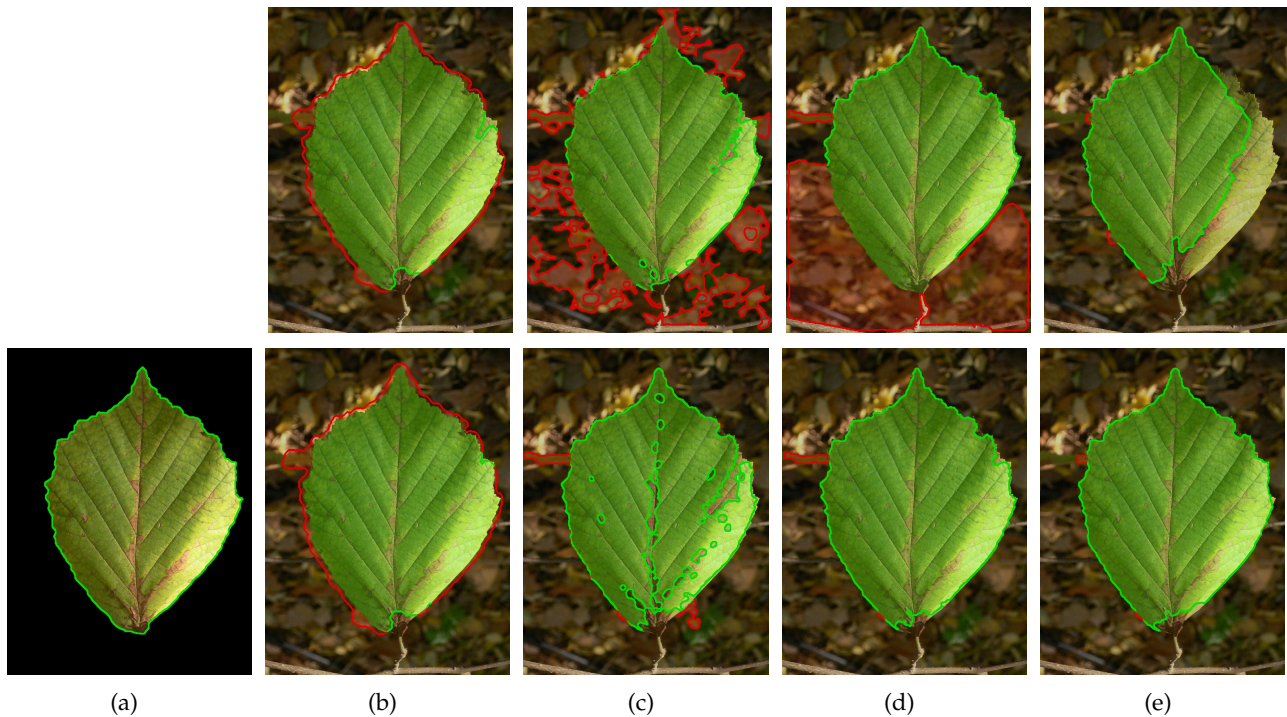


FIGURE 4.35 – Influence de l'initialisation manuelle sur la segmentation : B-splines snakes (b), contours actifs (c), GrabCut (d) et contours actifs guidés (e)

La méthode de segmentation présentée dans ce chapitre s'impose comme une façon fiable, robuste et répétable, de segmenter les feuilles dans des images en environnement naturel. Elle repose sur une modélisation de la couleur qui permet une analyse fine du contenu de l'image, et qui limite notamment le risque de confusion de la feuille d'intérêt avec des objets de couleur proches de l'arrière plan. La première phase d'évaluation d'un modèle paramétrique de feuille est alors cruciale puisque c'est elle, via les fortes contraintes de formes tirées de la connaissance *a priori*, qui va permettre de limiter la zone d'intérêt dans l'image à un objet ayant la forme d'une feuille.

En guidant les contours actifs par cette contrainte forte, la segmentation qu'on obtient peut être qualifiée de prudente, mais elle évite surtout les potentiels artéfacts dus aux débordements si fréquents dans le contexte d'images aussi complexes. Eviter de sacrifier la reconnaissance + précision. Le résultat est que les objets ainsi extraits possèdent une régularité et une constance qui offre la possibilité d'une description plus poussée des formes foliaires locales. Cette première brique si déterminante dans le processus de reconnaissance se devait de bénéficier d'efforts particuliers pour ne pas compromettre la suite de l'analyse et permettre d'envisager une compréhension complète des feuilles dans les images naturelles.

Bibliographie

- [Bez81] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers, Norwell, MA, USA, 1981.
- [BHU00] Patrick Brigger, Jeff Hoeg, and Michael Unser. B-spline snakes : a flexible tool for parametric contour detection. *IEEE Transactions on Image Processing*, pages 1484–1496, 2000.
- [BJ01] Y.Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In *IEEE International Conference on Computer Vision*, volume 1, pages 105–112, 2001.
- [BL79] S. Beucher and C. Lantuejoul. Use of Watersheds in Contour Detection. In *International Workshop on Image Processing : Real-time Edge and Motion Detection/Estimation*, 1979.
- [BM92] Serge Beucher and Fernand Meyer. The morphological approach to segmentation : the watershed transformation. *Mathematical Morphology in Image Processing*, 1992.
- [BW06] T. Brox and J. Weickert. Level set segmentation with multiple regions. *IEEE Transactions on Image Processing*, 15(10) :3213–3218, 2006.
- [Can86] John Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6) :679–698, 1986.
- [Che95] Yizong Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8) :790–799, 1995.
- [CM99] Dorin Comaniciu and Peter Meer. Mean shift analysis and applications. In *IEEE International Conference on Computer Vision*, volume 2, pages 1197–1203, 1999.
- [CTCG95] Timothy F. Cootes, Christopher J. Taylor, David H. Cooper, and Jim Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1) :38–59, 1995.
- [CV01] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2) :266–277, 2001.

- [Dem68] Arthur P. Dempster. A generalization of bayesian inference. *Journal of the Royal Statistical Society*, 30 :205–247, 1968.
- [Fel04] P. Felzenszwalb. Representation and detection of deformable shapes. *PAMI*, 27 :208–220, 2004.
- [FH04] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59, 2004.
- [GBJ⁺11] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The CLEF 2011 plant images classification task. In *CLEF 2011 Working Notes*, 2011.
- [GPS89] D. Greig, B. Porteous, and A. Seheult. Exact Maximum A Posteriori Estimation for Binary Images. *Royal Journal on Statistical Society*, 51(2) :271–279, 1989.
- [HP74] S. L. Horowitz and T. Pavlidis. Picture Segmentation by a directed split-and-merge procedure. *International Joint Conference on Pattern Recognition*, pages 424–433, 1974.
- [KWT88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, 1(4) :321–331, 1988.
- [LXGM10] An Li, Xuemei Xu, Qiaoyun Guo, and Qin Mo. Robust segmentation using marked regular pyramid and mean shift. In *International Symposium on Intelligence Information Processing and Trusted Computing*, pages 341–344, 2010.
- [MH80] David Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London Series B*, 207 :187–217, 1980.
- [MSV95] Ravikanth Malladi, James A. Sethian, and Baba C. Vemuri. Shape modeling with front propagation : a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2) :158–175, 1995.
- [OS88] Stanley Osher and James A. Sethian. Fronts propagating with curvature dependent speed : Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1) :12–49, 1988.
- [Ots79] Nobuyuki Otsu. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 :62–66, 1979.
- [RKB04] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut" : interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23(3) :309–314, 2004.

- [Ron94] Remi Ronfard. Region-based strategies for active contour models. *International Journal of Computer Vision*, 13(2) :229–251, 1994.
- [Sha76] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [SK08] Yonggang Shi and William Clement Karl. A real-time algorithm for the approximation of level-set-based curve evolution. *IEEE Transactions on Image Processing*, 17(5) :645–656, 2008.
- [YHC92] A. Yuille, P. Hallinan, and D. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 8(2) :99–111, 1992.
- [ZLMY⁺02] C. Zimmer, E. Labruyere, V. Meas-Yedid, N. Guillen, and J-C Olivo-Marin. Improving active contours for segmentation and tracking of motile cells in videomicroscopy. In *International Conference on Pattern Recognition*, volume 2, pages 286–289, 2002.

Chapitre 5

Description structurelle du contour

Compréhension et représentation locale des formes foliaires



Ilex aquifolium L.

Résumé du Chapitre

Ce chapitre propose une méthode pour la description des contours de feuilles obtenus après la phase de segmentation de l'image. Elle repose sur une interprétation sémantique à haut niveau des contours s'appuyant sur la modélisation de la feuille. L'analyse qui s'ensuit permet la détection précise de points caractéristiques sur le contour, ce qui va rendre possible une description des formes locales du bord, de la base et du sommet de la feuille cohérente avec les notions botaniques. Nous introduisons une représentation spatialisée du bord basée sur une structure de chaîne, ainsi que les outils permettant sa manipulation. Enfin la modélisation des formes basales et apicales est adaptée pour caractériser la forme locale de la feuille en utilisant une approche de type modèle déformable.

UNNE fois extraite la forme de la feuille dans l'image, l'objet que l'on obtient doit être décrit le plus précisément possible, en prenant en compte toutes ses spécificités, pour disposer du maximum d'informations pertinentes pour la reconnaissance. Le fait est que la botanique apporte une connaissance forte des caractéristiques qui vont se montrer discriminantes dans le cadre de l'identification des espèces. Pour utiliser cette connaissance, pour en tirer parti de façon à optimiser la pertinence des informations extraites de la forme, il est nécessaire de considérer l'objet extrait comme une feuille, que l'on souhaite voir et comprendre de la même façon que le ferait un observateur humain.

Cela passe dès lors par l'interprétation de la forme obtenue avec un regard suffisamment haut, de façon à appliquer sur un ensemble de pixels peu porteur de sens des notions sémantiques qui semblent assez évidentes à l'œil, comme l'apex de la feuille, la nervure principale, le pétiole et son point d'insertion, ou encore les lobes et les dents éventuels. C'est une analyse qui ne peut se faire qu'en émettant certaines hypothèses qui vont permettre d'introduire une forme de connaissance experte dans l'interprétation de l'image. Cette interprétation a pour but d'enrichir la forme obtenue, de ne plus la voir comme une forme quelconque mais comme une forme de feuille, que l'on sait décrire de façon spécifique, en observant des localisations précises, pour caractériser ses formes locales, à la base, à l'apex ou sur les bords.

Un contour interprété permet alors d'envisager une description de haut niveau des propriétés de la feuille, s'attachant à capturer les propriétés étudiées par les botanistes. L'idée est en effet d'arriver à caractériser par une description numérique les informations discriminantes que la botanique décrit par un vocabulaire spécifique. Si les modélisations de la forme des feuilles accomplissent cette tâche au niveau de la forme globale, il faut, pour y parvenir au niveau des formes plus locales, analyser le contour en se penchant sur des endroits bien précis.

5.1 Interprétation du contour des feuilles

L'interprétation du contour obtenu après segmentation de l'image est grandement facilitée par le fait que nous disposons d'une modélisation polygonale de la forme globale, qui constitue elle-même une première interprétation de la feuille dans l'image. En particulier, ce modèle fournit la définition d'un ou plusieurs axes principaux de la feuille ou de la foliole. On peut ainsi estimer connaître l'axe de la nervure principale, et des nervures secondaires donnant naissance aux lobes éventuels. Une telle information permet de parvenir à une compréhension très complète du contour qui rend possible l'évaluation de descripteurs de haut niveau.

5.1.1 Partition sémantique du contour

Ce que nous cherchons à produire à partir du contour de la forme segmentée et de la donnée du modèle polygonal de la forme globale de la feuille est en fait un découpage en plusieurs parties

bien différentes, correspondant chacune à une zone d'intérêt où on s'attend à observer une forme particulière. Ces zones, dont la définition est guidée par les notions utilisées en botanique, sont les suivantes :

- ▶ **Apex** de la feuille, partie située à l'extrémité de la nervure principale, souvent en pointe
- ▶ **Apex des lobes** de la feuille, extrémités des nervures secondaires pour une feuille palmatilobée
- ▶ **Pétiole** de la feuille, fin et plus ou moins allongé, reliant le limbe de la feuille à la tige
- ▶ **Base** de la feuille, partie située de part et d'autre du point d'insertion du pétiole

En pratique, il est difficile de garantir que le pétiole de la feuille fera bien partie de la forme extraite de l'image. En effet, il n'est pas rare que sa couleur soit sensiblement différente de celle de la feuille, ce qui le rendra difficile à inclure par notre méthode de segmentation basée sur la similarité de couleur. De plus, la contrainte de forme imposée au contour fait que même lorsque le pétiole pourrait être atteint, il le sera rarement jusqu'au bout, étant donné qu'il n'est pas pris en compte par le modèle polygonal pris comme forme de référence. On se retrouve donc souvent avec un pétiole absent ou incomplet, mais il est tout de même intéressant de le détecter lorsqu'il est présent, même si peu d'informations fiables pourront en être tirées.

Pour arriver à partitionner le contour, l'idée est de considérer les intersections du contour avec des secteurs angulaires définis à partir des points du modèle polygonal, de façon à isoler des portions particulières de contour comme sur la Figure 5.1. De manière plus précise, nous allons construire des sous-ensembles Γ_B , Γ_A et Γ_{A_l} du contour Γ^* , correspondant respectivement à la base, à l'apex et aux apex des lobes. La définition de chacun de ces ensembles s'appuie sur un secteur angulaire partant du point extrême opposé du modèle polygonal Π^* , soit A^A pour la base, et B^B pour les différents apex. Le secteur est tronqué de façon à ne garder que des points suffisamment éloignés, ce qui donne par exemple pour Γ_B :

$$\Gamma_B = \left\{ \Gamma(u) \mid \left(\cos \left(\overrightarrow{A^A \Gamma(u)}, \overrightarrow{A^A B^B} \right) > \cos(8^\circ) \right) \wedge \left(\frac{A^A \Gamma(u)}{A^A B^B} > 0.75 \right) \right\}$$

Pour les extrémités des lobes, le rapport de longueur est fixé à 0.9, de façon à ne contenir que la pointe du lobe dans Γ_{A_l} . Cette méthode très simple donne une partition fiable du contour, du moment que le modèle polygonal obtenu correspond bien à l'alignement et à la forme de la feuille. Elle permet également de dégager deux portions supplémentaires du contour, à savoir le bord gauche Γ_g et le bord droit Γ_d comme les points non-affectés du contour se situant de part et d'autre des zones Γ_B et Γ_A .

Avec une telle définition, le pétiole est nécessairement inclus dans la zone de la base Γ_B , et on procède à une subdivision de cet ensemble pour former la zone Γ_p du pétiole. Celui-ci étant une structure fine, il disparaît très rapidement lorsqu'on procède à l'érosion de l'image binaire représentant $\Omega(\Gamma)$. Utiliser les opérations de morphologie mathématique pour supprimer le pétiole est une pratique courante [WHD⁺08, KBB⁺12] et nous faisons de même pour obtenir un contour érodé Γ_\ominus par un

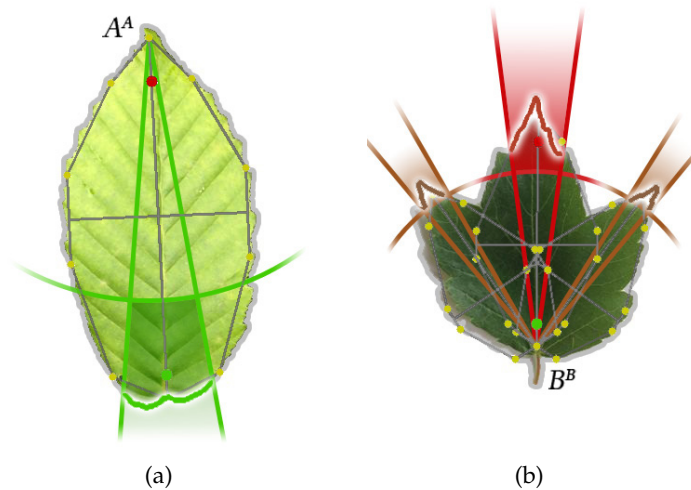
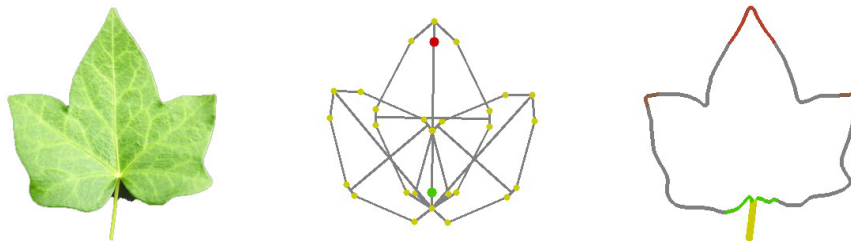


FIGURE 5.1 – Partition du contour de la feuille utilisant les axes du modèle polygonal pour délimiter la zone de la base (a) la zone de l'apex et les apex des lobes (b)

élément structurant de taille r , où le pétiole a donc disparu. On considère alors comme appartenant au pétiole les points de la zone de base qui sont trop lointains de ce contour érodé :

$$\Gamma_P = \Gamma_B \cap \left\{ \Gamma(u) \mid \|\Gamma(u) - \Gamma_{\ominus}^*\| > r\sqrt{2} \right\}$$

Le résultat de cette affectation des points du contour à des ensembles significatifs est une interprétation quasi sémantique du contour qui permet de savoir où regarder pour caractériser les formes locales de la feuille, que ce soit la forme de l'apex ou de la base, ou bien la répartition des dents sur le reste du bord. La description peut alors se faire sans être perturbé par des éléments comme la pointe des lobes ou le pétiole qui ne rentrent pas en ligne de compte dans la définition de ces formes. Une illustration de l'interprétation produite est donnée dans la Figure 5.2.



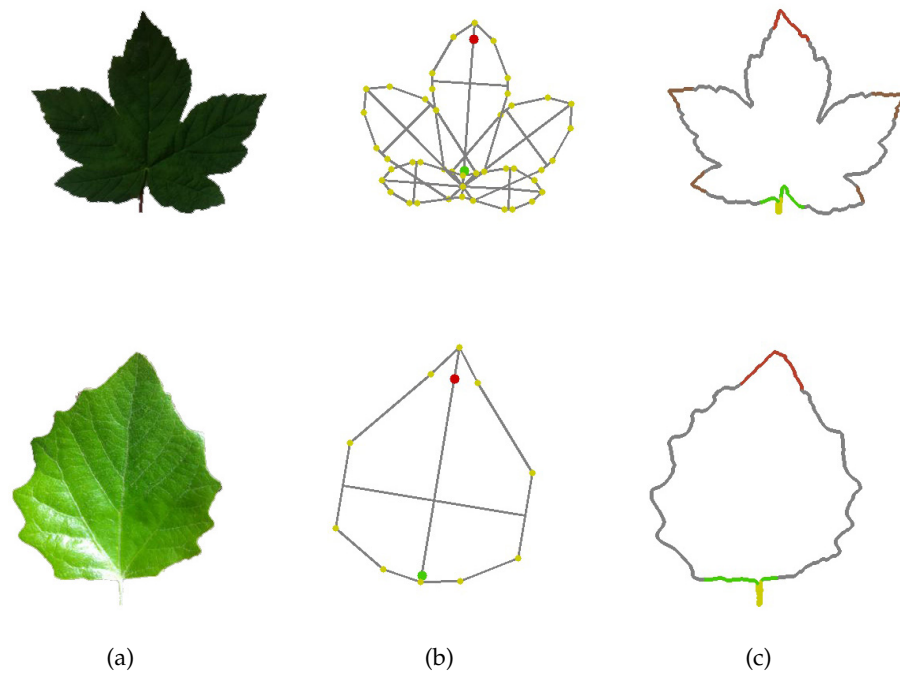


FIGURE 5.2 – Interprétation de la forme segmentée (a) à l’aide de la modélisation polygonale (b) par partition du contour de la feuille (c) : ■ zone de l’apex Γ_A , ■ zone de l’apex des lobes Γ_L , ■ zone de la base Γ_B , dont ■ pétiole Γ_P

5.1.2 Analyse du *Curvature-Scale Space*

Les éléments que l’on souhaite localiser pour examiner les formes foliaires locales sont généralement faciles à repérer du point de vue de la géométrie, se détachant assez clairement lorsqu’on jette un œil à la courbure du contour. L’un des principaux problèmes de cette vision intuitive est que les niveaux de détail auquel se trouvent les petites dents, les ondulations ou l’apex de la feuille ne sont absolument pas les mêmes, et considérer le contour avec un point de vue uniforme ne peut par conséquent pas être suffisant pour détecter de façon adéquate les points déterminants du contour.

Une représentation plus riche du contour qui évite cet obstacle de l’échelle optimale à laquelle considérer le contour est la transformation dans l’espace courbure-échelle ou *Curvature-Scale Space* (CSS) [MM92], qui a d’ailleurs déjà été utilisé dans le cadre de la reconnaissance de feuilles basée sur la forme [AMK97, MA04, CA10]. Elle consiste à empiler des mesures de courbures calculées en chaque point du contour à des échelles croissantes, de telle sorte que toute structure concave ou convexe sur le contour apparaîtra dans le CSS à son échelle optimale.

Dans la méthode originale, le contour subit un lissage par des noyaux gaussiens d’écart-type croissant de façon à calculer la courbure sur différentes échelles. Il y a cependant de nombreux estimateurs de la courbure (inverse du rayon du cercle osculateur, mesures d’angle, dérivée seconde) qui nécessitent bien souvent la donnée d’une portion de contour comme support pour être calculés.

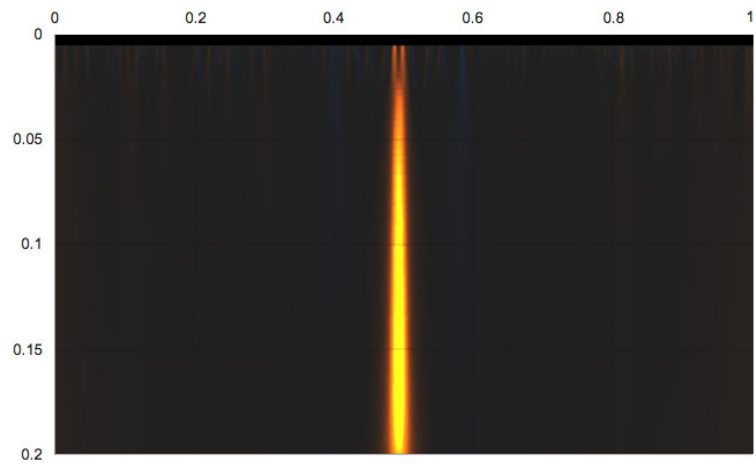
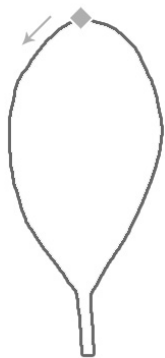
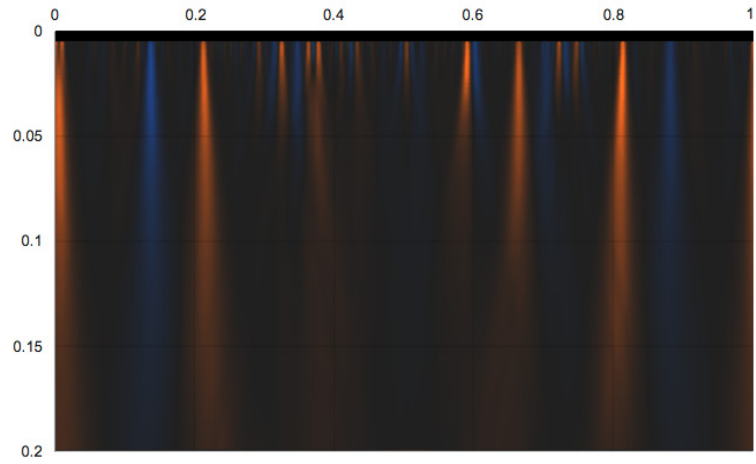
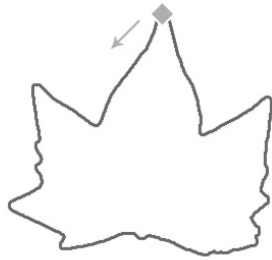
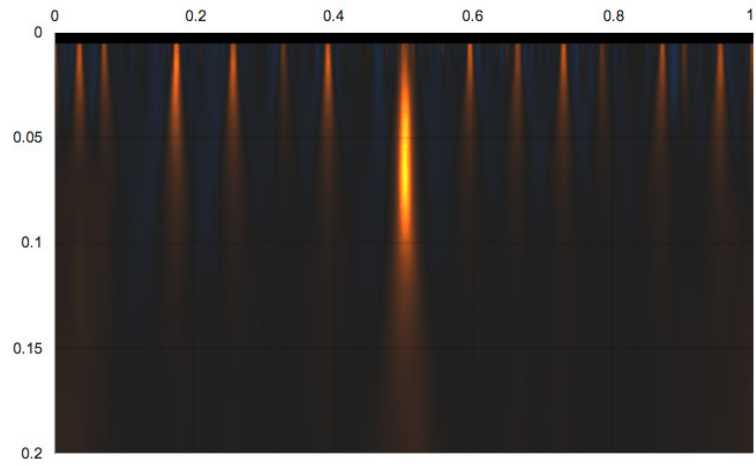
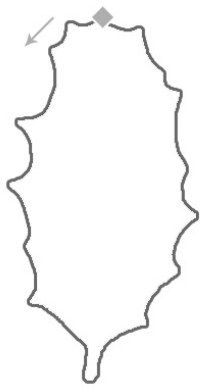
Si la taille de ce support est fixe, on risque de manquer les structures dont l'échelle ne correspond pas à cette taille. Bien qu'il soit possible d'évaluer pour tout point du contour le support optimal pour l'évaluation de la courbure, nous avons préféré utiliser des supports de taille croissante pour la calculer aux différentes échelles. Par conséquent, le fait de lisser la courbe ferait double emploi et nous calculons alors toujours une mesure de courbure sur le contour original, en considérant un voisinage d'échelle croissante autour du point.

La mesure de courbure que nous utilisons s'appuie sur la définition mathématique de la courbure κ comme la norme $\|\partial^2\Gamma/\partial u^2\|$ de la dérivée seconde de la courbe par rapport à son abscisse curviligne. Dans le plan, où $\Gamma(u) = (x(u); y(u))$, cette formule se simplifie, et fait intervenir les dérivées premières et secondes en x et en y . Ce sont ces dérivées que nous estimons sur des voisinages de taille s croissante. La courbure au point d'abscisse curviligne u à valeur entre 0 et 1, pour une échelle donnée s est alors calculée comme suit :

$$\begin{aligned}\kappa(u, s) &= \frac{x'_s(u) \cdot y''_s(u) - x''_s(u) \cdot y'_s(u)}{(x'_s(u)^2 + y'_s(u)^2)^{3/2}} \\ x'_s(u) &= \frac{1}{s} \left(x\left(u + \frac{s}{2}\right) - x\left(u - \frac{s}{2}\right) \right) \\ y'_s(u) &= \frac{1}{s} \left(y\left(u + \frac{s}{2}\right) - y\left(u - \frac{s}{2}\right) \right) \\ x''_s(u) &= \frac{1}{s^2} \left(x\left(u + \frac{s}{2}\right) - 2x(u) + x\left(u - \frac{s}{2}\right) \right) \\ y''_s(u) &= \frac{1}{s^2} \left(y\left(u + \frac{s}{2}\right) - 2y(u) + y\left(u - \frac{s}{2}\right) \right)\end{aligned}\tag{5.1}$$

Le CSS peut alors être vu comme une matrice où chaque ligne correspond à une valeur de s , chaque colonne à une abscisse curviligne u , et contenant les valeurs de $\kappa(u, s)$. Pour le calcul, le pas de discrétisation est de 0.001, aussi bien pour l'échelle que pour l'abscisse curviligne. Pour accélérer le processus, on peut se contenter de prendre un pas plus large pour l'échelle, et d'extrapoler les valeurs de courbure intermédiaires de façon linéaire, les variations d'une échelle à l'autre étant continues. On obtient dans tous les cas une représentation telle que sur la Figure 5.3, qui pourrait s'apparenter à un relief, où toutes les convexités et les concavités du contour apparaissent clairement ainsi que l'échelle jusqu'à laquelle elles persistent.

Pour des applications de reconnaissance de formes, le CSS a souvent été utilisé en ne considérant que les CSS-contours, c'est-à-dire les courbes formées par les passages à zéro de la courbure. Si c'est un objet qui permet effectivement de localiser les structures remarquables du contour (notamment les concavités, qui sont généralement entourées par des zéros de courbure, celle-ci tendant à être positive lorsque l'échelle augmente) cela constitue une perte considérable d'information puisque les valeurs concrètes de la courbure sont occultées. Il semble donc plus judicieux de conserver l'information dans son ensemble, pour garder une représentation parfaitement fidèle du contour.



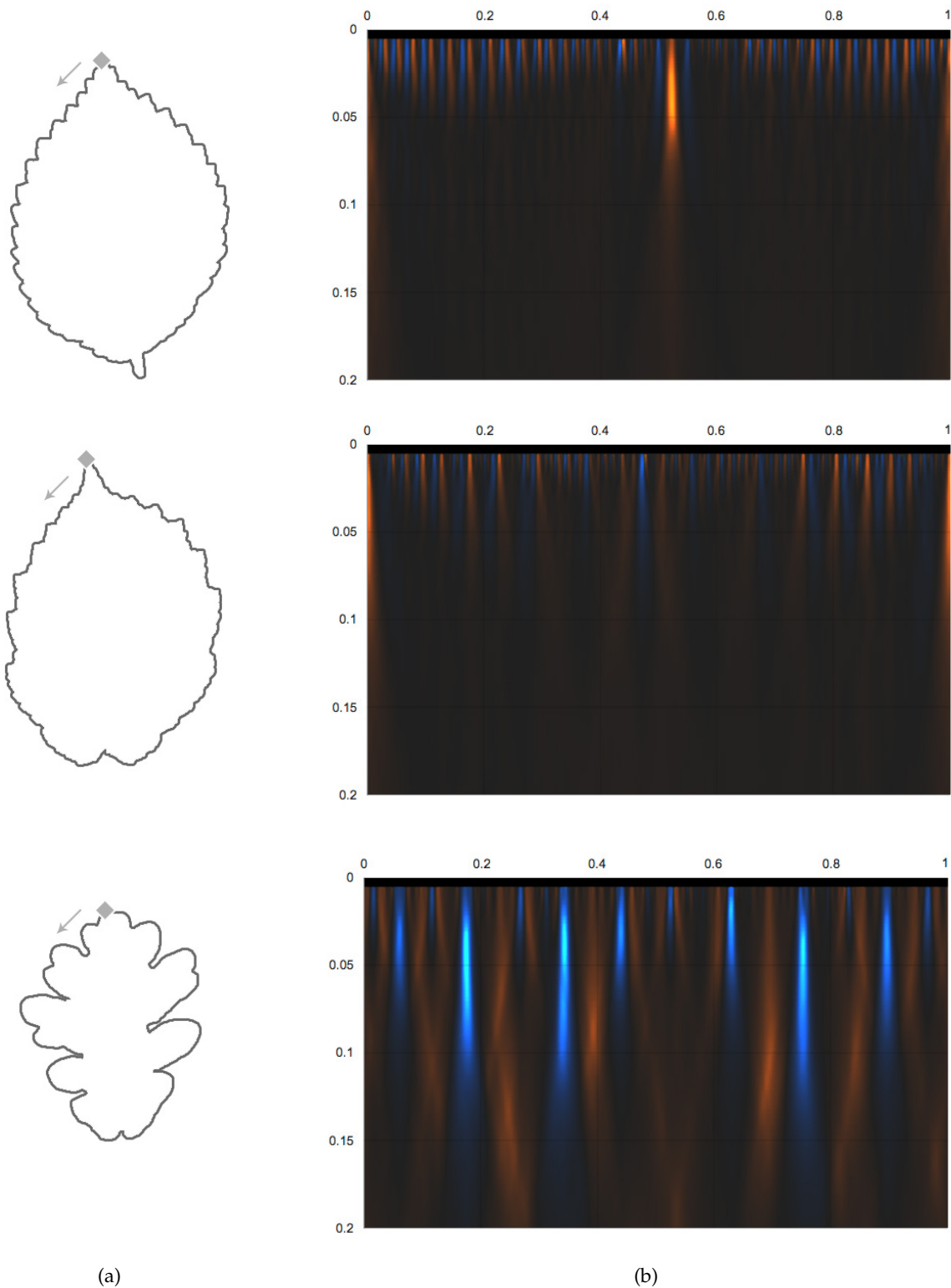




FIGURE 5.3 – Transformation de contours de feuille (a) en *Curvature-Scale Space* (b) : abscisse curviligne du contour en abscisse, échelle en ordonnée, la couleur représentant la courbure (courbure positive  et courbure négative )

Sur les images représentant le CSS, chaque dent, lobe ou creux sur le bord de la feuille peut en effet être associé à une strie verticale, d'autant plus longue que la taille de l'élément concerné est importante, et d'autant plus intense qu'il est marqué. Le pétiole et l'apex forment généralement les raies les plus importantes, tout comme les apex des lobes éventuels. Un bord lisse ne présentera ainsi pas d'autre strie importante sur les portions correspondant au bord de la feuille, alors qu'une feuille dentée montrera une succession plus ou moins régulière de stries concaves et convexes dont la hauteur représente la taille des dents.

La transformation du contour de la feuille en *Curvature-Scale Space* offre donc une représentation visuellement très riche du contour, qui peut être utilisée pour détecter et pour caractériser des formes locales du bord de la feuille. Telle quelle, elle est cependant trop volumineuse et informative pour être considérée comme un descripteur, mais son analyse et son interprétation doivent amener une quantité importante d'informations.

En particulier, l'application la plus évidente serait de localiser ces stries visuellement si évidentes, et de reporter leur position, et donc celle de l'élément signalé, sur le contour de la feuille. Ce serait ainsi un moyen de détecter toutes les structures localement saillantes sur lesquelles s'appuyer pour la description du contour, et l'information contenue dans le CSS doit permettre de les caractériser complètement.

5.1.3 Détection et description de points caractéristiques

Les points que l'on recherche dans le *Curvature-Scale Space* correspondent à des structures définies précisément sur le bord de la feuille par leur courbure (dents, pointes, sinus) et il est par conséquent plus approprié de regarder, non pas les points où la courbure s'annule, mais au contraire les endroits où la courbure est minimale ou maximale dans le CSS. Ce sont en effet bien ces points qui pourront être rattachés directement à une concavité ou à une convexité du contour.

Chercher ces points particuliers revient en pratique à détecter les points dominants du contour, un sujet passablement connu [TC89], et même résolu en passant par le domaine courbure-échelle [PL92]. Le principe de base est de détecter les points extrêmes de la courbure, ceux-ci étant suffisamment saillants pour qu'une reconstruction polygonale basée uniquement sur ces points soit encore assez proche de la forme du contour. La méthode que nous utilisons est donc largement inspirée de ces travaux.

La détection de ces points va donc s'effectuer à chaque niveau d'échelle s , en commençant par le plus bas. En pratique, la valeur de départ pour s est 0.005, considérant que la courbure pour des valeurs inférieures ne permet pas de différencier les véritables dentelures de la feuille du bruit lié à la discrétisation du contour. L'idée est alors simplement de localiser les points correspondant aux minima et aux maxima locaux de la courbure, cette notion de localité étant définie par un voisinage de taille s . On évalue donc pour des valeurs croissantes de s un ensemble de points \mathcal{D}_s contenant initialement l'ensemble des points du contour et affiné par le processus suivant :

1. Ne conserver que les points pour lesquels la valeur de la courbure est un optimum local :

$$\mathcal{D}_s \leftarrow \mathcal{D}_s \cap \left\{ \Gamma(u) \mid \forall u', \left(|u' - u| < \frac{s}{2} \right) \implies |\kappa(u, s)| \geq |\kappa(u', s)| \right\}$$

2. Supprimer les points pour lesquels la valeur de la courbure n'atteint pas un seuil minimal κ_{\min} :

$$\mathcal{D}_s \leftarrow \mathcal{D}_s \cap \{ \Gamma(u) \mid |\kappa(u, s)| > \kappa_{\min} \}$$

3. Supprimer les points qui ne peuvent être associés à un point à l'échelle inférieure (si $s > 0.005$) :

$$\mathcal{D}_s \leftarrow \mathcal{D}_s \cap \left\{ \Gamma(u) \mid \exists u', \left(|u' - u| < \frac{s}{2} \right) \wedge (u' \in \mathcal{D}_{s-0.001}) \right\}$$

4. Ne conserver que le point d'abscisse médiane des groupes potentiels de taille $\leq s$:

$$\forall u \in \mathcal{D}_s, g_s(u) = \mathcal{D}_s \cap \left\{ \Gamma(u') \mid \left(|u' - u| < \frac{s}{2} \right) \right\}, \mathcal{D}_s \leftarrow (\mathcal{D}_s \setminus g_s(u)) \cup \{ \text{median}(g_s(u)) \}$$

Le résultat de cette détection est un ensemble de points isolés à chaque niveau d'échelle (puisqu'il ne peut pas subsister de groupe de points) mais qui forment en fait des chaînes de points à travers les différentes échelles (étant donnée l'étape 3) suivant idéalement les lignes de crête et les fonds de vallées du CSS. Autrement dit, en traçant ces points dominants comme sur la Figure 5.4, on résume les raies visibles dans le CSS à des chemins dans l'espace courbure-échelle, qui correspondent donc chacun à une partie convexe ou concave du contour.

Etant donné les spécificités de notre méthode et notamment du calcul de la courbure, les points que l'on détecte ne sont pas toujours stables à travers les échelles, ce qui amène une certaine imprécision quant à la localisation du point sur le contour. Cependant, ce décalage renseigne également sur l'orientation de la convexité ou concavité représentée, puisqu'il sera d'autant plus important que celle-ci est dissymétrique, ce qui s'explique facilement par la prise en compte de supports symétriques pour le calcul de la courbure.

Dans le cadre de l'interprétation de ces points dominants, on va chercher à déterminer la taille et la courbure de chaque élément détecté sur le contour. On voit très clairement que toutes les chaînes ne vont pas aussi loin dans les échelles, et que l'on peut établir une correspondance directe entre le niveau d'échelle jusqu'auquel elles persistent et la taille de la structure qu'elles représentent. Autrement dit, le point qui va nous intéresser pour caractériser la taille d'une pointe ou d'un creux n'est autre que le point terminal de la chaîne.

Afin de ne conserver que ces points terminaux, il faut parcourir l'ensemble des niveaux d'échelles, en partant cette fois-ci du plus élevé, et lorsque l'on rencontre un point dominant à l'échelle s , supprimer tous ceux qui appartiennent à sa chaîne pour la suite de l'analyse, en éliminant les points de même signe de courbure des échelles inférieures dans un voisinage de taille s . Simplement, pour ne pas que cette suppression nuise à l'information extraite, nous choisissons ici d'effectuer les traitements

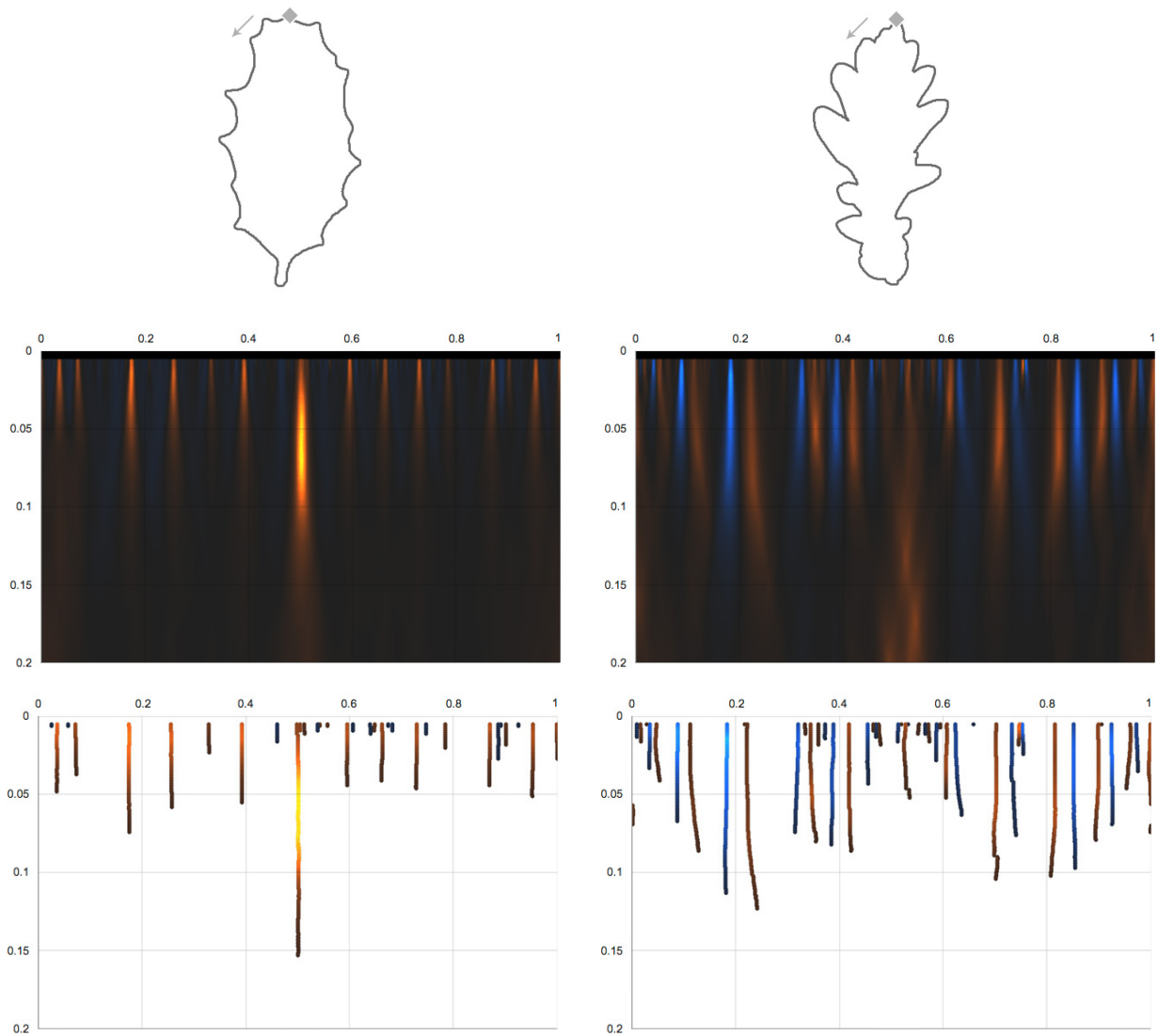


FIGURE 5.4 – Détection des points dominants du contour à différentes échelles à partir du CSS (courbure positive ■ et courbure négative ■)

séparément selon la zone du contour dans laquelle on se trouve. L'objectif est notamment que les convexités très importantes que sont l'apex et le pétiole n'aillent pas absorber les dents plus modestes des bords. Lorsqu'un point dominant $\Gamma(u)$ faisant partie de la zone Γ_P est rencontré à l'échelle s , on procède à la suppression partielle aux échelles inférieures :

$$\forall s' < s, \mathcal{D}_{s'} \leftarrow \mathcal{D}_{s'} \setminus \left\{ \Gamma(u') \in \Gamma_P \mid \left(|u' - u| < \frac{s}{2} \right) \wedge (\text{sign}(\kappa(u, s)) = \text{sign}(\kappa(u', s'))) \right\}$$

On s'assure ainsi de ne pas compter deux fois la même structure, et que les petites structures qui ne seraient que des sous-parties d'une plus grande (comme les sinuosités dessinant le contour d'un lobe) soient toutes reconnues comme une seule. Le fait de dissocier les zones dans le même temps

garantit que cette même fusion ne s'applique pas aux convexités présentes à la base et à l'apex de la feuille, de façon à ce que les structures plus fines puissent être prises en compte et analysées.

Au cours du même processus, on évalue une mesure de la courbure de la concavité ou de la convexité, en effectuant la moyenne $\bar{\kappa}(u, s)$ des courbures des points dominants considérés comme appartenant à la chaîne et supprimés, en gardant un seul point point de valeur absolue maximale à chaque échelle. La valeur de courbure $\kappa(u, s)$ du point terminal est généralement très faible, et bien peu représentative de celle de l'ensemble de chaîne. Cette estimation permet alors d'ajouter à la caractérisation de la taille de l'élément donnée par s une information relative à l'acuité de la pointe ou du creux considéré, et ce sans phase de traitement additionnelle. On peut ainsi écrire, si $\kappa(u, s) > 0$ (respectivement avec un min si $\kappa(u, s) < 0$) :

$$\bar{\kappa}(u, s) = \frac{1}{s} \sum_{s'=1}^s \max_{\substack{u' \in \mathcal{D}_{s'} \cap \Gamma_P \\ |u'-u| < \frac{s}{2}}} \kappa(u', s') \quad (5.2)$$

De la même manière, on peut calculer sur une chaîne de points dominants une mesure de l'orientation de la dent ou du creux concerné. Cette orientation o est apparente dans la déviation du point détecté à travers les échelles, et peut donc être mesurée par l'écart moyen en abscisse des points de la chaîne avec le point terminal. Ainsi, lorsqu'on rencontre le point d'abscisse u à l'échelle s , on peut calculer en même temps que l'on supprime les points des échelles inférieures :

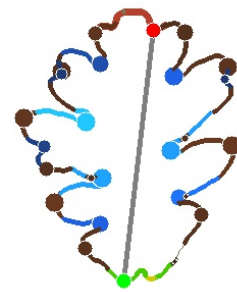
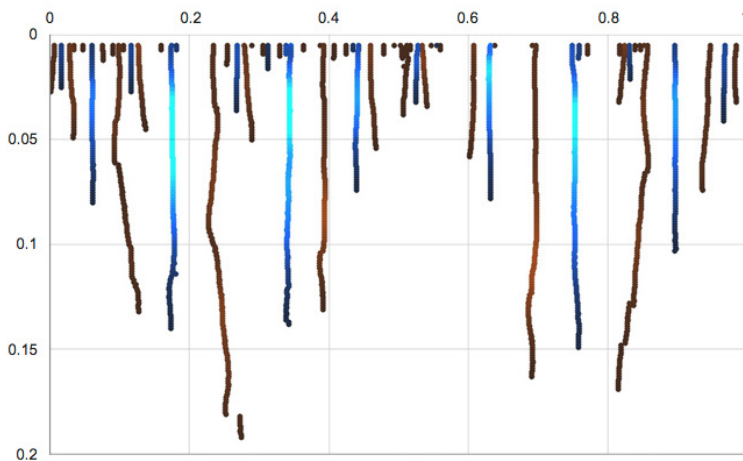
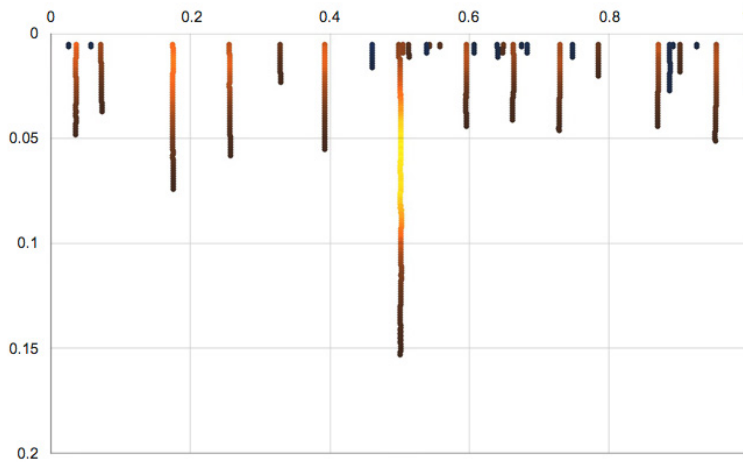
$$o(u, s) = \frac{1}{s} \sum_{s'=1}^s \max_{\substack{u' \in \mathcal{D}_{s'} \cap \Gamma_P \\ |u'-u| < \frac{s}{2}}} |u' - u| \quad (5.3)$$

On obtient alors plusieurs ensembles de points caractéristiques qui correspondent aux chaînes extraites du CSS, un pour chaque zone du contour, chacun des ces points étant associé à un ensemble de valeurs permettant de caractériser la portion de contour à laquelle il correspond.

La donnée de $\mathcal{D}_A = \bigcup_s \mathcal{D}_s \cap \Gamma_A$ et de $\mathcal{D}_B = \bigcup_s \mathcal{D}_s \cap \Gamma_B$ fournit des points qui sont des candidats potentiels pour le point d'apex de la feuille A^* et le point d'insertion du pétiole B^* . En sélectionnant parmi ces points celui dont le produit $s\sqrt{|\kappa|}$ est le plus élevé, on peut localiser précisément sur le contour ces points remarquables.

Les autres ensembles $\mathcal{D}_g = \bigcup_s \mathcal{D}_s \cap \Gamma_g$ et $\mathcal{D}_d = \bigcup_s \mathcal{D}_s \cap \Gamma_d$ correspondent aux dents présentes sur les bords gauche et droit de la feuille, qui vont pouvoir être analysées pour décrire la forme du bord de la feuille, caractéristique déterminante pour l'identification des espèces. Chacun de ces points est caractérisé par son abscisse curviligne u , sa taille s et sa courbure $\bar{\kappa}(u, s)$, et peut alors être représenté sur le contour de la feuille de façon à visualiser l'information extraite par analyse du CSS. Il en ressort alors une interprétation très riche du point de vue sémantique du contour de la feuille, où sa base et son apex sont localisés de manière fine, et où toutes les dents présentes sur son bord sont détectées et décrites indépendamment. Une telle interprétation du contour est illustrée dans la Figure 5.5.

L'un des avantages de placer d'aussi forts concepts sémantiques sur le contour est aussi d'atteindre une invariance à la rotation. En effet, les points de base et d'apex sont détectés par une mesure de courbure indépendante de la rotation de l'objet, et une fois localisés précisément, ils peuvent être utilisés comme points de référence. Tout traitement effectué à la suite pourra alors s'appuyer sur cette référence, rendant superflus les procédés coûteux de décalage et d'alignement utilisés pour faire des comparaisons. En particulier, des descripteurs non-agrégatifs (distance centre-contour, *shape contexts*, ...) bénéficieraient de l'introduction d'un point d'origine stable, qui permettraient de réduire grandement les temps de comparaison.



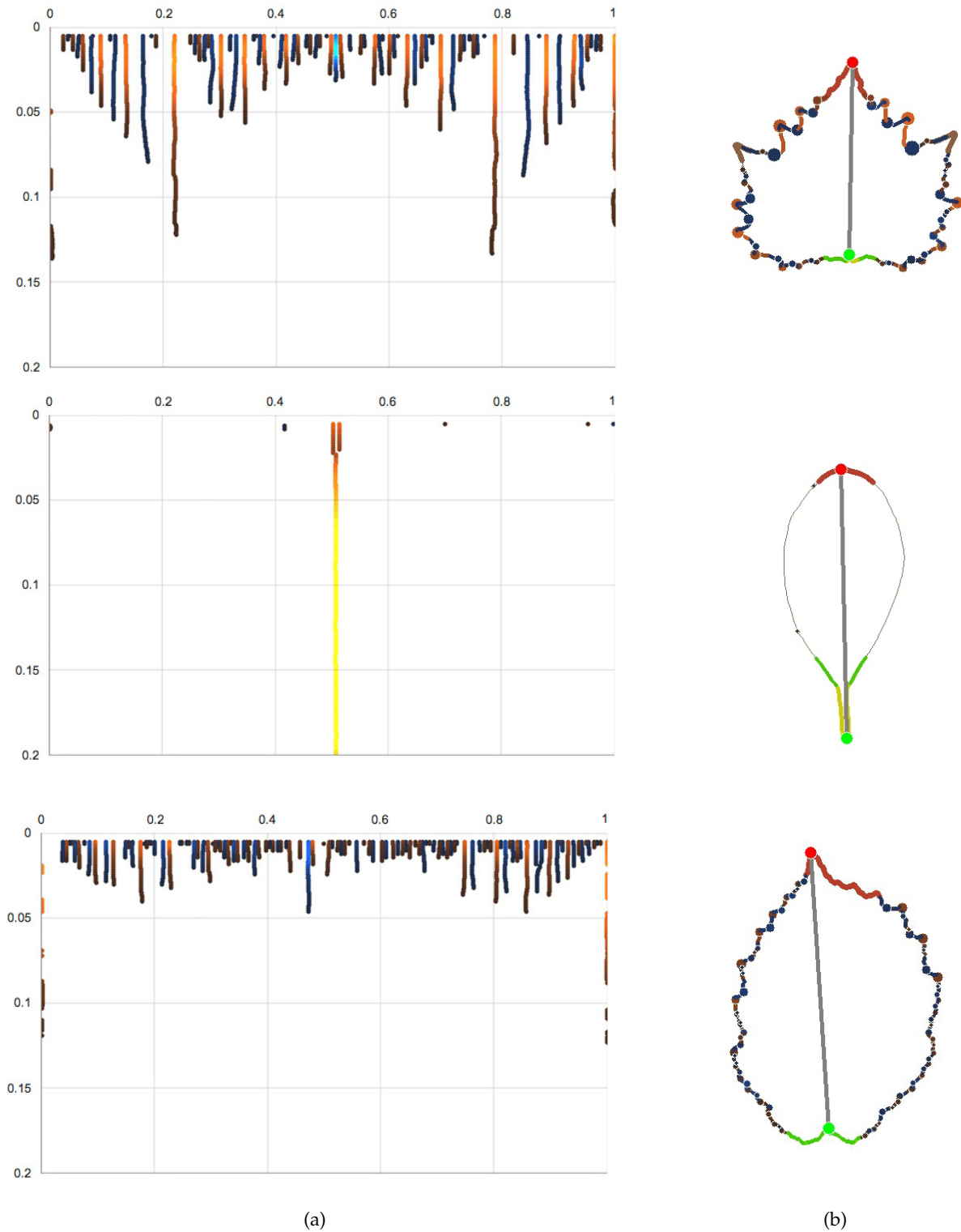


FIGURE 5.5 – Chaînes de points dominants (a) et contour interprété avec la base ● et l’apex ● détectés, et les dents et creux du bords caractérisés par l’échelle (représentée par la taille des points) et la courbure (courbure positive — et courbure négative —)

5.2 Mots de bord

La forme du bord de la feuille fait partie des caractéristiques les plus souvent discriminantes, et une de celles qui présentent le moins de variabilité au sein d'une espèce. Si l'on se réfère aux critères utilisés en botanique pour la décrire, plusieurs facteurs semblent entrer en ligne de compte. En premier lieu, la présence ou non de dents, ou du moins d'un quelconque relief sur le bord, est une distinction inévitable. Ensuite, si aspérités il y a, les termes utilisés renvoient aux caractéristiques des dents : leur taille, leur acuité, leur orientation. Mais les dents prises indépendamment ne suffisent pas à décrire le bord, et leur répartition, leur fréquence, leur régularité sont alors autant de paramètres supplémentaires à considérer.

Un descripteur de la forme du bord qui se contenterait d'agréger les propriétés des dents, sans prendre en compte la dimension spatiale dans sa formulation gommerait ainsi une bonne partie de l'information discriminante aux yeux des botanistes. Par conséquent, il y a un intérêt certain à concevoir une description du bord des feuilles qui garde une trace de l'information de spatialité, ce qui est possible notamment en utilisant des représentations structurelles.

5.2.1 Une représentation structurelle pour le bord des feuilles

Les approches structurelles (chaînes, arbres, graphes, grammaires) sont assez communes en reconnaissance de formes, même si leur nature les rend moins facilement manipulables que des données vectorielles. Dans notre cas, la description que nous avons choisie pour représenter les bords s'apparente à une structure de chaîne.

En effet, les ensembles de points \mathcal{D}_g et \mathcal{D}_d caractéristiques des dents du bord, peuvent être vus, indexés par leur abscisse curviligne u comme deux séquences de points. Pour une meilleure visualisation, on peut représenter leur répartition sur le bord entre les points de base et d'apex. Nous introduisons alors une nouvelle mesure d'abscisse curviligne $u^*(u, B^*, A^*)$ à valeur entre 0 et 1 qui va décrire les deux portions de courbes entre les points B^* et A^* , calculée à partir de leurs abscisses u_{B^*} et u_{A^*} .

En plus de la caractérisation par la taille s et la courbure κ , nous utilisons l'axe formé par les points B^* et A^* pour calculer pour chacun des points de \mathcal{D}_g et \mathcal{D}_d une valeur de position verticale $\pi(u, s)$. Il s'agit simplement du rapport entre la distance de B^* au point projeté sur la droite (B^*A^*) d'une part, et la longueur B^*A^* d'autre part (Figure 5.6 (a)). Cette mesure n'est pas tout à fait équivalente à l'abscisse u^* , l'ordre des points ne correspondant pas toujours à leur position verticale.

Référencées par l'abscisse u^* , \mathcal{D}_g et \mathcal{D}_d forment alors deux séquences parallèles de points \mathcal{P}_g et \mathcal{P}_d munis d'une caractérisation individuelle sous forme d'un vecteur comprenant leur abscisse relative, leur courbure, leur taille, et leur orientation. Ces séquences sont simplement la transposition des ensembles de points en ensembles de vecteurs descriptifs ordonnés selon u^* , et que l'on peut alors

représenter comme sur la Figure 5.6 :

$$\mathcal{P}_g = \left\{ \mathcal{P}_{g,i} = (u^*; \kappa; s; \pi; o)_{g,i} \right\}_{i=1..|\mathcal{D}_g|} = \bigcup_s \left\{ (u^*(u, B^*, A^*); \bar{\kappa}(u, s); s; \pi(u, s); o(u, s)), \Gamma(u) \in \mathcal{D}_s \cap \Gamma_g \right\}$$

$$\mathcal{P}_d = \left\{ \mathcal{P}_{d,i} = (u^*; \kappa; s; \pi; o)_{d,i} \right\}_{i=1..|\mathcal{D}_d|} = \bigcup_s \left\{ (u^*(u, B^*, A^*); \bar{\kappa}(u, s); s; \pi(u, s); o(u, s)), \Gamma(u) \in \mathcal{D}_s \cap \Gamma_d \right\}$$

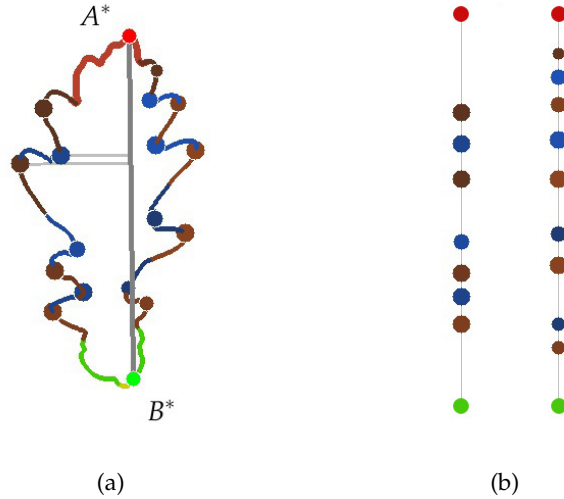


FIGURE 5.6 – Dents et creux détectés sur le contour de la feuille (a) et leur transposition en séquences parallèles représentatives des bords gauche et droit (b)

L'idée est alors de considérer ces séquences de vecteurs comme des chaînes de caractères, où les symboles utilisés ne seront pas les lettres d'un alphabet discret, mais plutôt les vecteurs $(\kappa; s; \pi)$ décrivant une dent ou un creux dans un espace tri-dimensionnel. L'orientation des dents est ici volontairement laissée de côté pour ne pas complexifier outre mesure l'espace dans lequel sont définis les symboles. Ces chaînes de symboles, peuvent être représentées comme sur la Figure 5.7.

Les objets ainsi formés décrivent fidèlement le profil du contour, en incluant une information forte quant à la répartition spatiale des structures sur le bord de la feuille. L'essentiel de l'information discriminante relative au bord se voit ainsi condensée dans un mot, une chaîne de symboles sur laquelle on va pouvoir travailler en s'appuyant sur des outils usuels de traitement des chaînes.

5.2.2 Outils de manipulation des chaînes

Parmi les représentations structurelles, les chaînes, initialement chaînes de caractères, ont la structure la plus simple et sont de ce fait les plus faciles à manipuler. Leur comparaison peut s'effectuer simplement par le calcul de distances d'édition, basées sur la distance de Levenshtein [Lev66], qui s'obtiennent de manière efficace, en même temps que les chemins optimaux d'édition par programmation dynamique [WF74]. Ces chemins sont en fait des suites d'opérations permettant la transformation d'une première chaîne en une seconde, qui utilisent des opérations unitaires de base pouvant être de

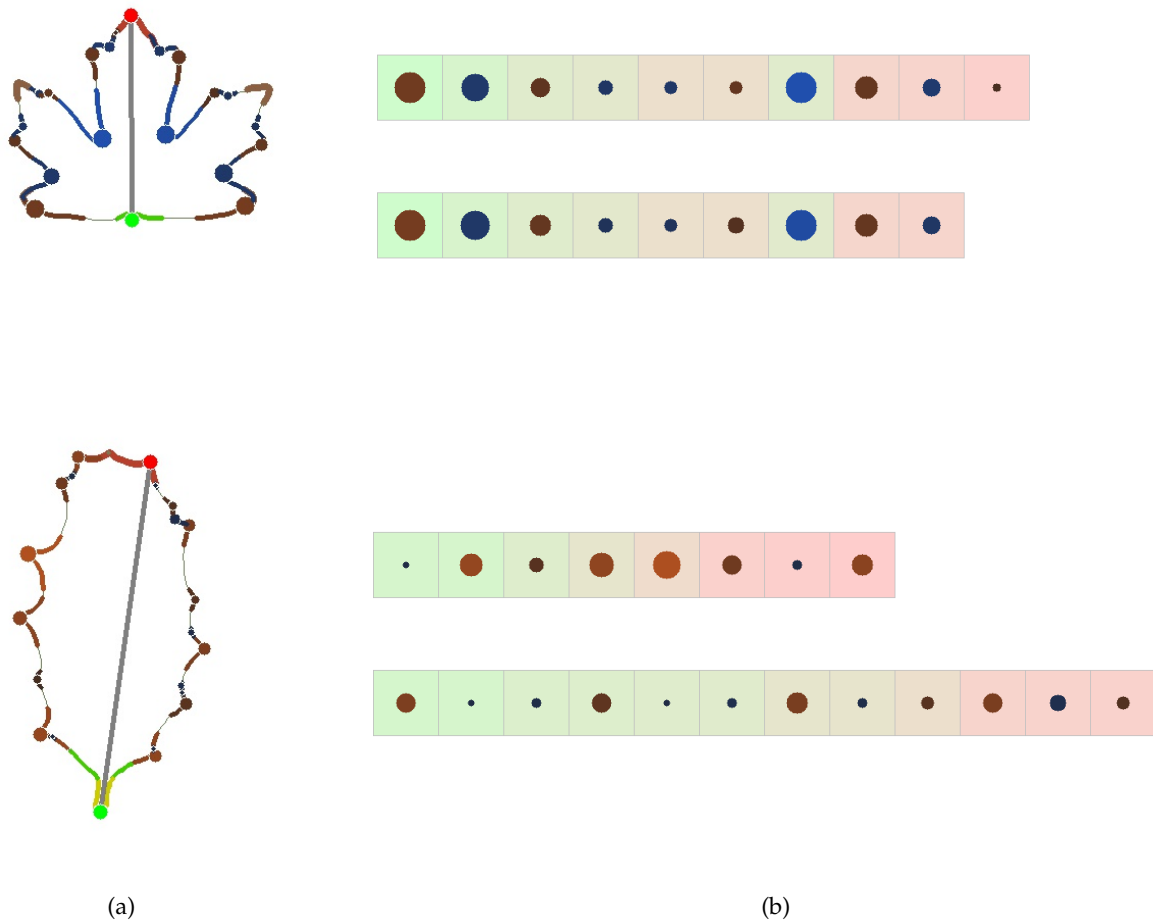


FIGURE 5.7 – Traduction des points détectés sur le contour de la feuille (a) en deux chaînes de symboles (bord gauche en haut et bord droit en bas) (b) représentés par leur taille, leur courbure (courbure positive \blacktriangle et courbure négative \blacktriangleleft) et leur position verticale (base \square apex \square)

trois types :

- ▶ **Suppression** d'un symbole de la première chaîne
- ▶ **Insertion** d'un symbole de la seconde chaîne à un emplacement donné
- ▶ **Substitution** d'un symbole de la première chaîne par un symbole de la seconde

La distance d'édition entre deux chaînes est alors définie comme le coût minimal d'un chemin d'édition entre celles-ci. Tout repose donc sur le coût des opérations unitaires, qui peut être fixe et égal à 1 dans le cas le plus simple, ou varier du tout au tout selon les opérands : on parle alors de distance d'édition pondérée. Par ailleurs, ces distances peuvent être normalisées pour atteindre plus de généralité [MV93, LL07], mais au risque de perdre l'inégalité triangulaire.

En ce qui concerne les ensembles de chaînes, il existe des approches pour trouver des groupes et des motifs dans ce type de données, que ce soit par clustering [SL98] ou via des cartes auto-adaptatives

(*Self Organizing Maps* ou SOM) [KS98]. On trouve aussi des moyens d'estimer la distribution des chaînes dans un ensemble, notamment par l'évaluation de chaînes médianes. On distingue d'ailleurs deux médianes différentes :

- ▶ **Médiane d'ensemble** : chaîne d'un ensemble qui minimise la somme des distances à toutes les chaînes de l'ensemble
- ▶ **Médiane généralisée** : chaîne quelconque générée par l'alphabet, minimisant la même quantité ; cette médiane ne peut généralement qu'être estimée [FZ00]

La variabilité d'un ensemble de chaînes est une notion plus compliquée à définir, bien plus que ne l'est par exemple la variance ou l'écart type d'une donnée numérique, ou la matrice de covariance d'une grandeur vectorielle. On peut néanmoins, à partir des chemins d'édition, définir la **dévi**ation d'un ensemble de chaînes [Jol03], comme la médiane généralisée des chemins d'éditions (eux-mêmes vus comme des chaînes) des chaînes de l'ensemble à sa médiane. Tous ces outils permettent d'envisager de manipuler les objets complexes que sont les chaînes à la manière de grandeurs vectorielles plus familières, et de pouvoir leur appliquer le même genre de méthodes.

Dans notre cas, le principal outil que nous utilisons pour traiter les chaînes de symboles, est ce que nous définissons comme une distance d'édition pondérée normalisée. Le véritable problème lié à la définition vient de l'alphabet par lequel sont décrits les symboles. Contrairement au cadre classique des chaînes de caractères où l'alphabet est discret, fini, et permet donc d'énumérer les différentes opérations possibles, nos symboles sont des vecteurs à valeurs dans un espace tridimensionnel continu. Aussi est-il inconcevable de définir explicitement le coût de toutes les substitutions possibles, comme cela se fait pour les caractères via des matrices de coût. De plus, l'identité entre deux symboles ne se présentera probablement jamais, bien que certains vecteurs puissent être très proches. Il faut donc pour obtenir des résultats plus fins qu'en utilisant des coûts unitaires, adapter cette approche au cas d'un alphabet vectoriel.

Le point central de cette adaptation est la définition des coûts des opérations élémentaires de suppression, d'insertion et de substitution qui doivent dépendre de leurs opérands. On considère alors deux chaînes de points $\mathcal{P}_1 = (\mathcal{P}_{1,i} = (\kappa_{1,i}; s_{1,i}; \pi_{1,i}))_{i=1..|\mathcal{P}_1|}$ et $\mathcal{P}_2 = (\mathcal{P}_{2,j} = (\kappa_{2,j}; s_{2,j}; \pi_{2,j}))_{j=1..|\mathcal{P}_2|}$ et nous nous intéressons aux opérations permettant de passer de l'une à l'autre. Il faut ici introduire le symbole nul et mot vide λ , de telle sorte que l'opération de **suppression** ($\mathcal{P}_i \rightarrow \lambda$) corresponde à la transformation d'un symbole en λ , et l'opération d'**insertion** ($\lambda \rightarrow \mathcal{P}_i$) à la transformation de λ en un nouveau symbole. Nous définissons alors les coûts c des trois opérations élémentaires nécessaires à la détermination du chemin d'édition optimal, les coûts d'insertion et de suppression dépendant directement de la taille de l'élément concerné, et celui de substitution d'une distance pondérée entre

les deux vecteurs utilisant une matrice diagonale de poids W :

$$\begin{aligned}
 c(\mathcal{P}_{1,i} \rightarrow \lambda) &= (s_{1,i})^{\frac{1}{2}} \\
 c(\lambda \rightarrow \mathcal{P}_{2,j}) &= (s_{2,j})^{\frac{1}{2}} \\
 c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) &= \left(((\kappa_{1,i}; s_{1,i}; \pi_{1,i}) - (\kappa_{2,j}; s_{2,j}; \pi_{2,j}))^T W ((\kappa_{1,i}; s_{1,i}; \pi_{1,i}) - (\kappa_{2,j}; s_{2,j}; \pi_{2,j})) \right)^{\frac{1}{2}}
 \end{aligned}
 \tag{5.4}$$

La distance d'édition $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2)$ va alors être simplement évaluée comme le coût ainsi calculé du chemin d'édition optimal $O^*(\mathcal{P}_1, \mathcal{P}_2)$ que l'on obtient en suivant une méthode de programmation dynamique. Ce calcul passe par le remplissage dynamique d'une matrice $(|\mathcal{P}_1| + 1) \times (|\mathcal{P}_2| + 1)$ contenant en chacune de ses cases $(i; j)$ le chemin d'opérations optimal $O_{i,j}$ pour transformer la sous-chaîne de \mathcal{P}_1 jusqu'à l'indice i en la sous-chaîne de \mathcal{P}_2 jusqu'à l'indice j . Ce chemin est directement associé à un coût $c_{i,j}$ qui va permettre de choisir les opérations optimales :

$$c_{i,j} = \sum_{o \in O_{i,j}} c(o)
 \tag{5.5}$$

Initialement, le chemin d'opérations est vide et son coût donc nul. On commence alors par remplir la première colonne correspondant à la suppression successive de tous les symboles de \mathcal{P}_1 , et la première ligne correspondant à l'insertion successive de tous les symboles de \mathcal{P}_2 en première position.

Une fois initialisée, la matrice va ensuite être parcourue en incrémentant les indices, et en sélectionnant localement à chaque étape la meilleure opération donnant le coût minimal. Ainsi, dans la case $(i; j)$, trois opérations sont possibles : l'insertion de $\mathcal{P}_{2,j}$ à la position i de \mathcal{P}_1 , la suppression de $\mathcal{P}_{1,i}$ et la substitution des deux symboles, de façon à obtenir dans chaque cas la sous-chaîne de $\mathcal{P}_{2,j}$ jusqu'à j . Les opérations s'appliquent cependant sur des sous-chaînes différentes, pour lesquelles la méthode assure qu'on a déjà calculé le chemin optimal.

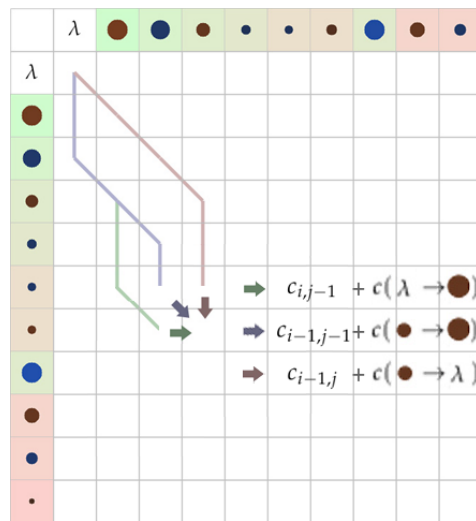


FIGURE 5.8 – Exemple de remplissage d'une case de la matrice contenant les chemins optimaux pour le calcul de la distance d'édition : le chemin choisi sera celui de coût minimal parmi les trois possibles

L'opération choisie est alors celle qui, lorsqu'on l'ajoute au chemin correspondant ($O_{i-1,j}$ pour la suppression, $O_{i,j-1}$ pour l'insertion et $O_{i-1,j-1}$ pour la substitution), donne un coût total minimal. On examine donc les trois coûts d'insertion $c_{i-1,j} + c(\mathcal{P}_{1,i} \rightarrow \lambda)$, de suppression $c_{i,j-1} + c(\lambda \rightarrow \mathcal{P}_{2,j})$ et de substitution $c_{i-1,j-1} + c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j})$, le minimum permet de déterminer l'opération optimale. On conserve alors dans $O_{i,j}$ l'ensemble du chemin d'opérations qui a permis de l'atteindre.

Au final, le chemin amenant à la dernière case ($|\mathcal{P}_1|; |\mathcal{P}_2|$) correspond au chemin le moins coûteux permettant de transformer l'ensemble de \mathcal{P}_1 en \mathcal{P}_2 , soit le chemin d'opérations optimal $O^*(\mathcal{P}_1, \mathcal{P}_2)$. Son coût total est par définition la distance d'édition $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2)$. Le détail de ce processus est donné dans l'Algorithme 6, et la Figure 5.8 illustre une étape de l'algorithme.

Algorithme 6 Calcul du chemin optimal d'édition entre deux chaînes

Entrée : $\mathcal{P}_1, \mathcal{P}_2$ deux chaînes de symboles

Sortie : $O^*(\mathcal{P}_1, \mathcal{P}_2)$ chemin d'édition optimal

```

 $O_{0,0} \leftarrow \{\}$ 
 $c_{0,0} \leftarrow 0$ 
for  $i = 1..|\mathcal{P}_1|$  do
   $O_{i,0} \leftarrow O_{i-1,0} \cup \{\mathcal{P}_{1,i} \rightarrow \lambda\}$ 
   $c_{i,0} \leftarrow c_{i-1,0} + c(\mathcal{P}_{1,i} \rightarrow \lambda)$ 
end for
for  $j = 1..|\mathcal{P}_2|$  do
   $O_{0,j} \leftarrow O_{0,j-1} \cup \{\lambda \rightarrow \mathcal{P}_{2,j}\}$ 
   $c_{0,j} \leftarrow c_{0,j-1} + c(\lambda \rightarrow \mathcal{P}_{2,j})$ 
end for
for  $i = 1..|\mathcal{P}_1|$  do
  for  $j = 1..|\mathcal{P}_2|$  do
     $o_{i,j} \leftarrow \{(\mathcal{P}_{1,i} \rightarrow \lambda); (\lambda \rightarrow \mathcal{P}_{2,j}); (\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j})\}$ 
     $O_{i,j}(\mathcal{P}_{1,i} \rightarrow \lambda) \leftarrow O_{i-1,j} \cup \{(\mathcal{P}_{1,i} \rightarrow \lambda)\}$ 
     $O_{i,j}(\lambda \rightarrow \mathcal{P}_{2,j}) \leftarrow O_{i,j-1} \cup \{(\lambda \rightarrow \mathcal{P}_{2,j})\}$ 
     $O_{i,j}(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) \leftarrow O_{i-1,j-1} \cup \{(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j})\}$ 
     $c_{i,j}(\mathcal{P}_{1,i} \rightarrow \lambda) \leftarrow c(O_{i,j}(\mathcal{P}_{1,i} \rightarrow \lambda)) = c_{i-1,j} + c(\mathcal{P}_{1,i} \rightarrow \lambda)$ 
     $c_{i,j}(\lambda \rightarrow \mathcal{P}_{2,j}) \leftarrow c(O_{i,j}(\lambda \rightarrow \mathcal{P}_{2,j})) = c_{i,j-1} + c(\lambda \rightarrow \mathcal{P}_{2,j})$ 
     $c_{i,j}(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) \leftarrow c(O_{i,j}(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j})) = c_{i-1,j-1} + c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j})$ 
     $O_{i,j} \leftarrow O_{i,j}(\text{argmin}_{o \in o_{i,j}} (c_{i,j}(o)))$ 
     $c_{i,j} \leftarrow \min_{o \in o_{i,j}} (c_{i,j}(o))$ 
  end for
end for
 $O^*(\mathcal{P}_1, \mathcal{P}_2) \leftarrow O_{|\mathcal{P}_1|, |\mathcal{P}_2|}$ 
 $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2) \leftarrow c_{|\mathcal{P}_1|, |\mathcal{P}_2|}$ 

```

La distance d'édition est ensuite normalisée de façon à obtenir des distances qui soient toujours comparables. En effet, on s'aperçoit vite qu'en sommant des coûts d'opérations unitaires, la distance obtenue ne va pas du tout avoir la même signification selon la longueur des chaînes comparées. Nous

nous assurons de maintenir la distance entre 0 et 1 en utilisant une méthode existante de normalisation [LL07] qui fait intervenir la distance maximale $d_{\text{edit,max}}(\mathcal{P}_1, \mathcal{P}_2)$ qui correspond au chemin d'édition trivial supprimant tous les symboles de \mathcal{P}_1 et insérant tous ceux de \mathcal{P}_2 . On obtient ainsi une distance normalisée $d_{\text{edit}}^*(\mathcal{P}_1, \mathcal{P}_2)$ qui est définie de la façon suivante :

$$d_{\text{edit}}^*(\mathcal{P}_1, \mathcal{P}_2) = \frac{2 \times d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2)}{d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2) + d_{\text{edit,max}}(\mathcal{P}_1, \mathcal{P}_2)} \quad (5.6)$$

où $d_{\text{edit,max}}(\mathcal{P}_1, \mathcal{P}_2) = \sum_{i=1}^{|\mathcal{P}_1|} c(p_{1,i} \rightarrow \lambda) + \sum_{j=1}^{|\mathcal{P}_2|} c(\lambda \rightarrow p_{2,j})$

Pour faire des mots de bord de feuilles des objets un peu plus pratiques, ce que nous souhaitons est de pouvoir estimer le représentant canonique d'un ensemble de ces chaînes, autrement dit calculer des moyennes ou des médianes. Une première étape est de calculer la médiane de deux chaînes \mathcal{P}_1 et \mathcal{P}_2 , par exemple pour se retrouver avec une unique chaîne représentant les deux bords d'une feuille. Pour ce faire, nous calculons simplement le chemin d'édition optimal $O^*(\mathcal{P}_1, \mathcal{P}_2)$ comme décrit ci-dessus. Nous définissons alors la chaîne médiane $\text{med}(\mathcal{P}_1, \mathcal{P}_2)$ en ne considérant que les opérations de **substitution** de ce chemin. On forme ainsi la chaîne des vecteurs moyens des symboles substitués à chaque opération :

$$\text{med}(\mathcal{P}_1, \mathcal{P}_2) = \left(\frac{\mathcal{P}_{1,i} + \mathcal{P}_{2,j}}{2} \right)_{(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) \in O^*(\mathcal{P}_1, \mathcal{P}_2)} \quad (5.7)$$

On peut démontrer que $\text{med}(\mathcal{P}_1, \mathcal{P}_2)$ ainsi définie est bel et bien une médiane (l'unicité n'étant pas garantie), c'est-à-dire une chaîne parmi toutes celles définies par le dictionnaire qui minimise la somme des distances d'édition (non-normalisées) à \mathcal{P}_1 et à \mathcal{P}_2 . Ainsi, soit \mathcal{P}_{med} une médiane de \mathcal{P}_1 et \mathcal{P}_2 , la somme $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_{\text{med}}) + d_{\text{edit}}(\mathcal{P}_{\text{med}}, \mathcal{P}_2)$ est donc minimale. Elle est en outre, en vertu de l'inégalité triangulaire, supérieure à $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2)$, qui s'exprime très simplement comme la somme des coûts de suppression, des coûts d'insertion et des coûts de substitution :

$$d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2) = \sum_{(\mathcal{P}_{1,i} \rightarrow \lambda) \in O^*} c(\mathcal{P}_{1,i} \rightarrow \lambda) + \sum_{(\lambda \rightarrow \mathcal{P}_{2,j}) \in O^*} c(\lambda \rightarrow \mathcal{P}_{2,j}) + \sum_{(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) \in O^*} c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{2,j}) \quad (5.8)$$

De façon assez évidente, la somme des distances à \mathcal{P}_1 et à \mathcal{P}_2 doit se rapprocher le plus possible de sa borne inférieure $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_2)$ pour être minimale. Si l'on sépare d'un côté les insertions et les suppressions (dont les coûts ont la même formulation), toutes les chaînes pour lesquelles la réunion des chemins $O^*(\mathcal{P}_1, \mathcal{P}_{\text{med}})$ et $O^*(\mathcal{P}_{\text{med}}, \mathcal{P}_2)$ ne contient que les opérations d'insertion et de suppression déjà dans $O^* = O^*(\mathcal{P}_1, \mathcal{P}_2)$ vont ainsi minimiser cette partie de la somme.

On peut donc en particulier choisir \mathcal{P}_{med} de façon à ce que toutes les insertions de O^* soient dans le chemin $O^*(\mathcal{P}_1, \mathcal{P}_{\text{med}})$, et toutes les suppressions dans le chemin $O^*(\mathcal{P}_{\text{med}}, \mathcal{P}_2)$. C'est notamment le cas avec la définition de $\text{med}(\mathcal{P}_1, \mathcal{P}_2)$. Le reste de la somme comprend alors les opérations de substitution,

qui seront nécessairement en même nombre dans chacun des deux chemins. Ce nombre correspond au nombre de substitutions dans O^* , et donc par définition à la longueur de \mathcal{P}_{med} . On obtient alors les distances d'édition suivantes :

$$\begin{aligned} d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_{\text{med}}) &= \sum_{(\mathcal{P}_{1,i} \rightarrow \lambda) \in O^*} c(\mathcal{P}_{1,i} \rightarrow \lambda) + \sum_{(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{\text{med},k}) \in O^*} c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{\text{med},k}) \\ d_{\text{edit}}(\mathcal{P}_{\text{med}}, \mathcal{P}_2) &= \sum_{(\lambda \rightarrow \mathcal{P}_{2,j}) \in O^*} c(\lambda \rightarrow \mathcal{P}_{2,j}) + \sum_{(\mathcal{P}_{\text{med},k} \rightarrow \mathcal{P}_{2,j}) \in O^*} c(\mathcal{P}_{\text{med},k} \rightarrow \mathcal{P}_{2,j}) \end{aligned} \quad (5.9)$$

La définition sous forme de distance des coûts de substitution permet de garantir que pour chaque symbole substitué, la quantité $c(\mathcal{P}_{1,i} \rightarrow \mathcal{P}_{\text{med},k}) + c(\mathcal{P}_{\text{med},k} \rightarrow \mathcal{P}_{2,j})$ est minimale pour un barycentre (à coefficients positifs) de $\mathcal{P}_{1,i}$ et $\mathcal{P}_{2,j}$, donc en particulier pour leur point moyen. Par conséquent la définition de $\text{med}(\mathcal{P}_1, \mathcal{P}_2)$ permet de minimiser la somme des distances d'édition $d_{\text{edit}}(\mathcal{P}_1, \mathcal{P}_{\text{med}}) + d_{\text{edit}}(\mathcal{P}_{\text{med}}, \mathcal{P}_2)$, ce qui justifie qu'elle est bien une médiane des deux chaînes \mathcal{P}_1 et \mathcal{P}_2 .

Pour un ensemble de chaînes plus conséquent $\mathcal{S} = \{\mathcal{P}_s\}_{s=1..N}$ l'algorithme d'estimation de la médiane généralisée $\text{med}(\mathcal{S})$ que nous utilisons est dérivé de celui introduit dans [FZ00], qui considère comme point de départ la médiane d'ensemble $\text{med}_{\mathcal{S}}(\mathcal{S})$:

$$\mathcal{P}_{\text{med}} \leftarrow \text{med}_{\mathcal{S}}(\mathcal{S}) = \underset{\mathcal{P}_{s_1} \in \mathcal{S}}{\text{argmin}} \left(\sum_{\mathcal{P}_{s_2} \in \mathcal{S}} d_{\text{edit}}(\mathcal{P}_{s_1}, \mathcal{P}_{s_2}) \right)$$

L'idée est alors, à partir de cette première estimation, de mettre à jour de façon itérative la chaîne \mathcal{P}_{med} de façon à s'approcher de la médiane généralisée. Cela passe par le calcul des chemins d'édition optimaux de \mathcal{P}_{med} à toutes les chaînes \mathcal{P}_s (on notera $O_s^* = O^*(\mathcal{P}_{\text{med}}, \mathcal{P}_s)$) et de sélectionner à chaque position de l'estimation courante l'opération la plus fréquente pour la lui appliquer. Malencontreusement, au vu de la nature continue de notre alphabet, on ne peut pas espérer retrouver exactement la même opération plus d'une fois ; nous avons donc choisi de comptabiliser les opérations selon 5 types :

- Suppression du symbole $\mathcal{P}_{\text{med},i}$:

$$D_{\mathcal{S},i} = \{(\mathcal{P}_{\text{med},i} \rightarrow \lambda) \in O_s^*, s = 1..N\}$$

- Substitution de $\mathcal{P}_{\text{med},i}$ par un symbole de même courbure :

$$S_{\mathcal{S},i}^+ = \{(\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{s,j}) \in O_s^*, s = 1..N \mid \kappa_{\text{med},i} \times \kappa_{s,j} > 0\}$$

- Substitution de $\mathcal{P}_{\text{med},i}$ par un symbole de courbure opposée :

$$S_{\mathcal{S},i}^- = \{(\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{s,j}) \in O_s^*, s = 1..N \mid \kappa_{\text{med},i} \times \kappa_{s,j} < 0\}$$

- Insertion après $\mathcal{P}_{\text{med},i}$ d'un symbole de même courbure :

$$I_{\mathcal{S},i}^+ = \{(\lambda \rightarrow \mathcal{P}_{s,j}) \in O_s^*, s = 1..N \mid (\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{s,j-1}) \in O_s^* \wedge \kappa_{\text{med},i} \times \kappa_{s,j} > 0\}$$

- Insertion après $\mathcal{P}_{\text{med},i}$ d'un symbole de courbure opposée :

$$I_{\mathcal{S},i}^- = \{(\lambda \rightarrow \mathcal{P}_{s,j}) \in O_s^*, s = 1..N \mid (\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{s,j-1}) \in O_s^* \wedge \kappa_{\text{med},i} \times \kappa_{s,j} < 0\}$$

Parmi les trois premiers ensembles $D_{\mathcal{S},i}$, $S_{\mathcal{S},i}^+$ et $S_{\mathcal{S},i}^-$, nous en retenons un, le plus représenté, pour déterminer l'opération à appliquer à la position i . En ce qui concerne les insertions, elles peuvent avoir lieu en plus d'une substitution ou d'une suppression, et il faut donc laisser la possibilité de les prendre en compte aussi, dès que la taille d'un des ensembles $I_{\mathcal{S},i}^+$ ou $I_{\mathcal{S},i}^-$ dépasse $N/2$. Nous utilisons alors les opérations retenues pour construire un chemin d'édition moyen $O_{\mathcal{S}}$, associant les opérations choisies mais effectuées avec le vecteur moyen comme deuxième opérande, par exemple si la substitution est retenue à la position i :

$$O_{\mathcal{S},i} \leftarrow \left(\mathcal{P}_{\text{med},i} \rightarrow \frac{1}{|S_{\mathcal{S},i}^+|} \sum_{(\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{s,j}) \in S_{\mathcal{S},i}^+} \mathcal{P}_{s,j} \right) \quad (5.10)$$

Le calcul des opérations moyennes serait le même avec les autres ensembles d'opérations. Le chemin $O_{\mathcal{S}}$ formé par la séquence de ces opérations moyennes est alors appliqué à l'estimation courante pour obtenir un nouveau \mathcal{P}_{med} . Si des changements ont eu lieu, alors le processus est répété avec cette nouvelle estimation, jusqu'à que le chemin à appliquer devienne $O_{\mathcal{S}} = ((\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{\text{med},i}))_{i=1..|\mathcal{P}_{\text{med}}|}$ et que plus aucune transformation ne puisse être effectuée. La chaîne finale $\mathcal{P}_{\text{med}}^*$ correspond alors à l'estimation de la médiane généralisée de l'ensemble des chaînes $\text{med}(\mathcal{S})$.

La Figure 5.9 illustre une première itération de ce processus. Parmi les quatre chaînes de l'ensemble considéré, la médiane d'ensemble est tout d'abord sélectionnée (Figure 5.9 (a)) et utilisée comme première estimation de la médiane. Les chemins d'édition permettant de transformer cette chaîne médiane en chacune des chaînes de l'ensemble sont ensuite calculés (Figure 5.9 (b)). Le chemin permettant de passer de la médiane d'ensemble à elle-même est alors l'identité $((\mathcal{P}_{\text{med},i} \rightarrow \mathcal{P}_{\text{med},i}))_{i=1..|\mathcal{P}_{\text{med}}|}$. En recensant les opérations figurant dans les chemins à chaque position, on construit enfin le chemin moyen (Figure 5.9 (c)). Par exemple en première position, on a trois substitutions et une suppression, c'est donc la substitution (par le symbole moyen des trois) qui va être retenue. Il figure aussi trois insertions de symboles de courbure négative, ce qui pousse à retenir cette opération ici, alors qu'elle ne l'est pas après le troisième symbole, où seules deux insertions interviennent. Ce chemin est ensuite appliqué à la médiane courante pour obtenir une nouvelle estimation de la médiane (Figure 5.9 (d)). Le processus devrait alors se poursuivre sur cette nouvelle chaîne.

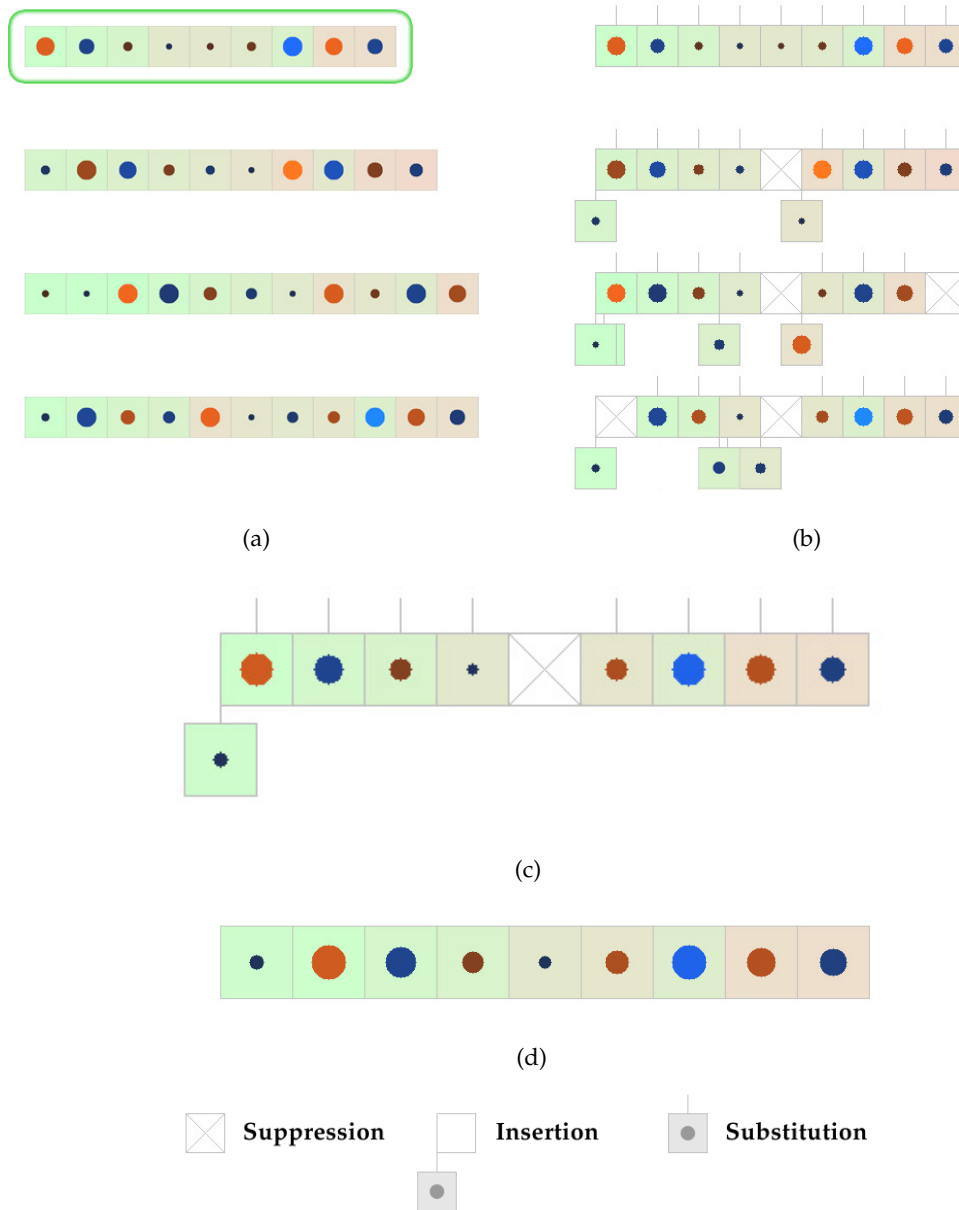


FIGURE 5.9 – Une étape de l’algorithme d’estimation de la médiane généralisée : ensemble de chaînes (a) et calcul de la médiane d’ensemble (entourée en vert), chemins d’édition de la médiane courante aux chaînes de l’ensemble (b), chemin d’édition moyen obtenu à partir de ces chemins (c) et médiane réestimée (d)

De façon intuitive, en sélectionnant itérativement les transformations les plus souvent employées pour atteindre les chaînes de l’ensemble, on réduit nécessairement la distance d’édition (les prochaines opérations seront moins coûteuses pour la majorité des chaînes de \mathcal{S}) et l’estimation se rapproche donc de la médiane généralisée de l’ensemble \mathcal{S} .

Cette notion de chaîne médiane s'avère très utile, puisqu'elle permet de condenser un ensemble de chaînes en un représentant unique, ce qui rend les comparaisons plus simples à mettre en œuvre. Il est en revanche plus difficile de représenter la variabilité d'un ensemble \mathcal{S} . Nous avons essayé de mettre en pratique une théorie de la déviation d'un ensemble de chaînes [Jol03] en estimant, une fois calculée $\text{med}(\mathcal{S})$ la médiane généralisée des chemins d'édition $O^*(\text{med}(\mathcal{S}), \mathcal{P}_s)$. Cela suppose donc de voir les chemins comme des chaînes d'opérations, et de définir des coûts de suppression, insertion et substitution d'opérations, ce qui est une entreprise difficile et très peu intuitive.

Nous pensons toutefois que notre méthode pourrait bénéficier d'une telle information complémentaire, mais pour qu'elle soit robuste, une certaine forme d'apprentissage devrait être employée pour traiter au mieux ces chaînes d'opérations. Sans compter que la prise en compte de la variabilité pour le calcul de distances (un peu à la manière d'une distance de Mahalanobis) constitue un défi supplémentaire, qui demanderait un travail fourni pour être traité correctement.

Ces outils de traitement constituent néanmoins une base suffisante pour se servir des objets extraits du contour des feuilles et les utiliser comme une description de la forme du bord. On dispose en effet de moyens de comparer et établir des similitudes entre les mots décrivant les bords, et de construire des modèles à partir d'un certain nombre de ces mots, de façon à réaliser une généralisation à partir de données individuelles.

5.2.3 Application à l'identification des espèces

La principale application des outils de traitement des chaînes représentant le bord des feuilles va être de les utiliser comme descripteurs pour la reconnaissance de espèces végétales. La première étape dans ce cadre est de former un seul descripteur par feuille, en lieu et place des chaînes \mathcal{P}_g et \mathcal{P}_d représentant les bords gauche et droit.

Pour y parvenir, nous évaluons simplement la chaîne médiane de ces deux bords $\mathcal{P}^* = \text{med}(\mathcal{P}_g, \mathcal{P}_d)$ en utilisant la méthode développée pour le cas de deux chaînes. Ce qui est en fait réalisé par cette opération est une mise en correspondance deux à deux des pointes et des creux détectés sur chaque côté, de telle sorte que seules les structures qui peuvent être reliées à une structure similaire sur le bord opposé soient conservées dans le descripteur. Bien que les feuilles ne soient jamais parfaitement symétriques, la correspondance est généralement bien présente, et il est intéressant de la faire apparaître concrètement.

Cet alignement, et l'élimination des points ne pouvant être appariés, a aussi pour effet, tout comme le moyennage des vecteurs décrivant les points, de lisser les imperfections qui pourraient résulter d'erreurs locales dans la phase de segmentation. La présence d'artéfacts se signalant par une courbure très forte va ainsi voir son influence réduite, puisque si le même genre de structure n'est pas visible sur le bord opposé, ils risquent bien d'être tout bonnement ignorés, ou dans le cas contraire adoucis. En particulier, si l'un des bords est lisse et vide de dents, le mot résultant sera lui-même vide, même si l'autre bord est très bruité. Des exemples de ces mots médians et des appariements qui en découlent sont donnés dans le Figure 5.10.

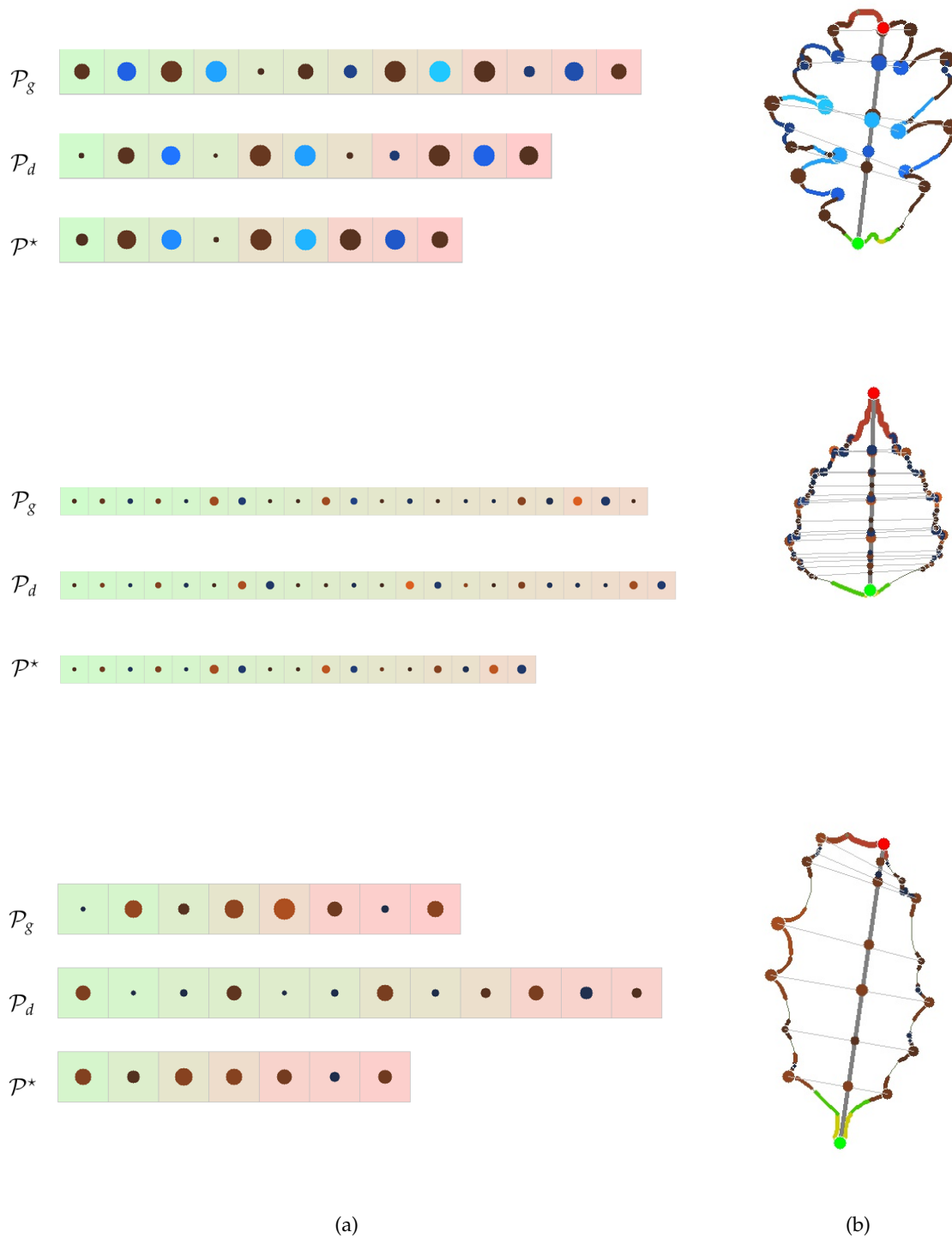


FIGURE 5.10 – Calcul du mot de bord médian à partir des mots des bords gauche et droit (a) et transposition sous forme d'appariement des points détectés (b)

Un tel descripteur de bord constitue en premier lieu une synthèse visuellement très explicite des dents et des creux rencontrés sur le bord de la feuille, interprétable facilement à haut niveau, tout en étant un objet manipulable par la machine. Il permet quasiment une conversion directe en courbe, si l'on retranscrit les convexités et les concavités représentées par les points. Il pourrait ainsi constituer un outil de visualisation intéressant pour une reconstruction automatique de la forme analysée.

Plusieurs approches sont alors possibles pour exploiter ce descripteur dans le cadre de la reconnaissance des espèces. Dans tous les cas, on part d'une base d'exemples de taille importante, constituée de chaînes et de leur étiquette de classe, soit ici leur espèce : $\mathcal{S} = \left\{ \left(\mathcal{P}_n; E_{\epsilon(n)} \right) \right\}$. A chaque espèce E_e va ainsi être associé un ensemble de chaînes $\mathcal{S}_e = \{P_n \mid \epsilon(n) = e\}$.

La méthode la plus basique est la recherche des chaînes les plus proches d'un exemple \mathcal{P} dans l'ensemble de \mathcal{S} . C'est une approche simple mais extrêmement coûteuse en termes de calcul puisqu'elle nécessite de calculer les distances d'édition $d_{\text{edit}}(\mathcal{P}, \mathcal{P}_n)$ pour tous les exemples de la base, puis de classer les résultats par ordre de distance croissante. Reste alors à déterminer l'espèce reconnue, soit comme l'espèce de la première réponse (plus proche voisin), soit comme la plus représentée parmi les k premières (k plus proches voisins), ou encore de classer l'ensemble des espèces, en fonction de leur ordre d'apparition dans le classement des exemples.

Pour éviter d'avoir à effectuer autant de calculs sur des objets moins simples que des vecteurs (le calcul de la distance d'édition s'effectue par exemple en temps quadratique) une solution pour limiter le nombre de comparaisons est de ne conserver qu'un seul représentant par espèce. L'approche la plus évidente est de considérer le mot médian comme modèle de l'espèce. C'est la solution que nous avons retenue, car elle séduit tant par sa légèreté en termes de temps de calcul que par sa capacité à synthétiser une grande quantité d'information dans un modèle qui reste interprétable à haut niveau :

$$\mathcal{P}_e = \text{med}(\mathcal{S}_e) = \text{med}\left(\left\{\left(\mathcal{P}_n; E_{\epsilon(n)}\right)\right\}\right)$$

Les modèles d'espèces qui résultent de ce calcul sont des séquences de points beaucoup plus régulières et synthétiques que les chaînes obtenues sur une feuille donnée, comme le montrent les séquences de la Figure 5.11. Le calcul de la chaîne médiane permet donc une généralisation très intéressante ne serait-ce que d'un point de vue visuel, où il apparaît clairement une synthèse fidèle des propriétés du bord d'une espèce, un bord canonique idéal qui donne un bon aperçu de la forme généralement prise par l'espèce.

Malgré cet aspect intellectuellement très satisfaisant, un modèle canonique n'est généralement pas suffisant pour couvrir la diversité des formes prises par une espèce, et le risque est grand que des individus d'une espèce donnée s'éloignent dans une assez grande mesure de ce représentant idéal. C'est dans la difficulté qu'il y a à capturer la variabilité des ensembles, à la quantifier de manière exploitable, que les méthodes structurelles montrent leurs limitations.

Pour pallier à cette limite intrinsèque des approches structurelles, un recours est de repasser de manière plus ou moins artificielle dans un espace vectoriel, où les classes sont plus facilement

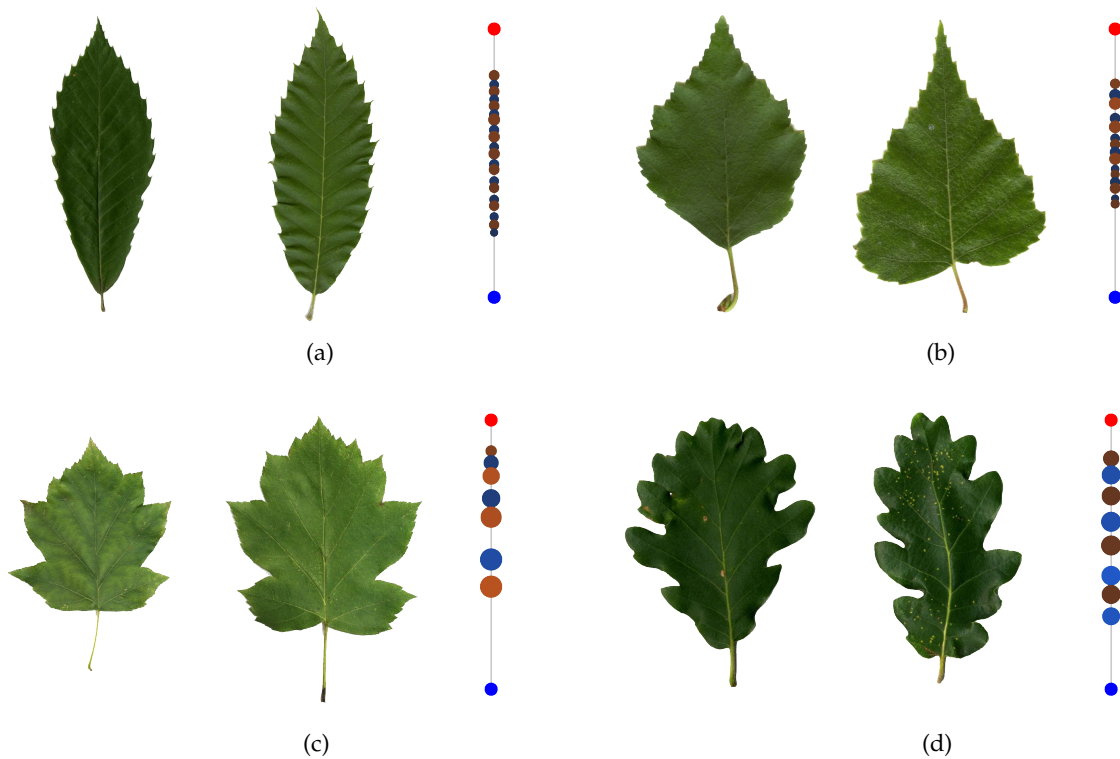


FIGURE 5.11 – Exemples de mots médians représentatifs de quelques espèces calculés sur la base Pl@ntLeaves 2012 : châtaignier (*Castanea sativa* Mill.) (a), bouleau verruqueux (*Betula pendula* Roth) (b), alisier torminal (*Sorbus torminalis* Crantz) (c) et chêne pubescent (*Quercus pubescens* Willd.) (d)

modélisables, et où des algorithmes de classification usuels fonctionnent à merveille. C’est ce que permettent de faire les noyaux de chaînes. Un noyau est une fonction symétrique semi-définie positive qui à deux éléments d’un même domaine quelconque associe une valeur réelle, qui peut s’exprimer comme un produit scalaire. Par définition, la distance d’édition $d_{\text{edit}}(\cdot, \cdot)$ constitue un noyau sur les chaînes telles que nous les avons définies, et il est possible de connaître pour toute paire d’éléments sa valeur $d_{\text{edit}}(\mathcal{P}_n, \mathcal{P}_{n'})$.

Les noyaux permettent alors d’utiliser bon nombre d’algorithmes, de classification notamment, qui ne s’appuient que sur des produits scalaires comme le *Support Vector Machine* (SVM). Avec des données vectorielles, ces produits doivent être calculés à partir des valeurs des points. Avec les noyaux de chaînes, la donnée vectorielle est implicite et seule la valeur du produit scalaire (ici $d_{\text{edit}}(\mathcal{P}_n, \mathcal{P}_{n'})$) va entrer en ligne de compte.

Enfin, il est également possible de revenir à une donnée vectorielle explicite, en calculant des mesures statistiques à partir des points détectés, de façon à retrouver des objets plus pratiques. Le fait de passer par un processus agrégatif entraîne nécessairement une perte d’information quant à la spatialité, et ces mesures ne sauraient donc être parfaitement satisfaisantes. Il s’agira néanmoins d’un objet plus simple, que l’on pourra utiliser dans le cadre d’algorithmes de classification usuels.

Nous définissons ainsi un certain nombre de grandeurs permettant de résumer les propriétés des

dents sur le bord de la feuille. Les premières d'entre elles consistent à mesurer quelle fraction du bord analysé peut être considéré comme faisant partie d'une pointe (w_+), d'un creux (w_-), ou au contraire d'une portion lisse (w_0). Nous utilisons pour cela le mot moyen, où les éléments irréguliers n'étant pas présents sur les deux bords sont ignorés. Ces mesures de bases sont exprimées comme des rapports de longueurs en abscisse curviligne, et font intervenir l'échelle s :

$$w_+ = \frac{2}{|\Gamma_g| + |\Gamma_d|} \sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i > 0}} (s_i) \quad w_- = \frac{2}{|\Gamma_g| + |\Gamma_d|} \sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i < 0}} (s_i) \quad (5.11)$$

$$w_0 = \max(1 - w_+ - w_-, 0)$$

Nous calculons ensuite des mesures statistiques simples pour rendre compte des propriétés des pointes et des creux. Il s'agit ici simplement de moyennes et d'écart-types pour la courbure, l'échelle et l'orientation des points détectés (lorsqu'il y en a), en dissociant courbures positives et négatives. La seule particularité est que ces mesures sont calculées en pondérant chaque point par son échelle, de façon à ce que chaque point du contour contribue de la même façon (une structure d'échelle s représentant un nombre de points du contour proportionnel à s). On obtient ainsi 12 valeurs calculées sur les points concaves ($-$) ou convexes, respectivement $\bar{s}_+, \sigma_{s+}, \bar{s}_-, \sigma_{s-}$ pour l'échelle, $\bar{\kappa}_+, \sigma_{\kappa+}, \bar{\kappa}_-, \sigma_{\kappa-}$ pour la courbure, et $\bar{o}_+, \sigma_{o+}, \bar{o}_-, \sigma_{o-}$ pour l'orientation. Le calcul est le même pour tous les paramètres, par exemple pour la courbure des pointes convexes :

$$\bar{\kappa}_+ = \frac{\sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i > 0}} (s_i \kappa_i)}{\sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i > 0}} (s_i)} \quad \sigma_{\kappa+} = \sqrt{\frac{\sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i > 0}} (s_i \kappa_i^2)}{\sum_{\substack{(s_i; \kappa_i; \pi_i) \in \mathcal{P}^* \\ \kappa_i > 0}} (s_i)} - \bar{\kappa}_+^2} \quad (5.12)$$

Le fait de considérer des valeurs moyennes est une perte d'information, mais cela constitue malgré tout un bon moyen de modéliser les propriétés de la dent représentative du bord en termes de taille, d'acuité et d'orientation. Il s'agit néanmoins d'une approximation, surtout dans le cas de feuilles avec plusieurs niveaux de dentition. On peut penser que les écarts-types amènent assez d'information quant à la variabilité de ces propriétés et suffisent à représenter les feuilles avec plusieurs types de dents. Ils restent cependant un regard global, gommant l'information de répartition et de fréquence.

Pour prendre tout de même en compte cette répartition, nous utilisons des mesures supplémentaires pour garder une trace de la position des points les uns par rapport aux autres. Ces mesures correspondent à l'écart en position de chaque point aux points suivant et précédent de courbure opposée, lorsqu'ils existent :

$$\delta\pi_i \uparrow = \min_{\substack{(s_j; \kappa_j; \pi_j) \in \mathcal{P}^* \\ j > i, \kappa_j \kappa_i < 0}} (\pi_j - \pi_i) \quad \delta\pi_i \downarrow = \min_{\substack{(s_j; \kappa_j; \pi_j) \in \mathcal{P}^* \\ j < i, \kappa_j \kappa_i < 0}} (\pi_j - \pi_i) \quad (5.13)$$

Ces différences sont alors simplement moyennées sur l'ensemble des points, en distinguant les

points de courbure positive et de courbure négative. Nous produisons ainsi 4 paramètres supplémentaires, $\overline{\delta\pi} \uparrow_+$, $\overline{\delta\pi} \downarrow_+$ pour les points convexes, et $\overline{\delta\pi} \uparrow_-$, $\overline{\delta\pi} \downarrow_-$ pour les points concaves, qui reflètent la répartition spatiale des dents sur le bord.

On ne conserve alors que 13 paramètres qui peuvent être utilisés comme une description condensée des propriétés du bord sous forme d'un vecteur \mathbf{p}_{Bord} , et être employés à la place des mots de bord pour la classification des espèces (Chapitre 6). Bien qu'ils correspondent par leur construction aux propriétés que les botanistes cherchent à analyser sur le bord des feuilles, le regard qu'ils portent sur le bord en occulte la dimension spatiale pourtant déterminante, qui nécessite un minimum de structure pour être prise en compte.

$$\mathbf{p}_{Bord} = \left(w_0; \bar{s}_+; \bar{s}_-; \sigma_{s+}; \sigma_{s-}; \bar{\kappa}_+; \bar{\kappa}_-; \sigma_{\kappa+}; \sigma_{\kappa-}; \bar{d}_+; \bar{d}_-; \overline{\delta\pi} \uparrow_+ / \overline{\delta\pi} \downarrow_+; \overline{\delta\pi} \uparrow_- / \overline{\delta\pi} \downarrow_- \right)$$

Les approches possibles en vue de la reconnaissance sont donc multiples à partir de cette représentation explicite et spatialisée du bord des feuilles. Il est tout à fait envisageable de l'utiliser telle quelle, mais il sera intéressant de mesurer l'apport de la représentation structurelle, comparé notamment à une description vectorielle moins riche. La prise en compte de la variabilité est intrinsèquement moins aisée, mais les possibilités existent pour compenser cette difficulté. En revanche, cette description offre l'opportunité d'une interprétation sémantique quasiment directe, visuellement intelligible, qui pourra s'avérer particulièrement intéressante dans le cadre de l'interaction avec un utilisateur.

5.3 Description des formes basales et apicales

L'interprétation du contour de la feuille permet, en plus de l'analyse des structures présentes sur le bord, de repérer les zones d'intérêt particulières que sont la base et l'apex de la feuille. Le traitement basé sur le CSS fournit une localisation précise des points autour desquels la forme de ces zones va pouvoir être caractérisée. Cela offre donc la possibilité d'une description spécifique de ces formes, incorporant une part de connaissance, de façon à évaluer les caractéristiques discriminantes pour la reconnaissance.

Pour prendre en compte cette information *a priori*, nous disposons de modélisations dédiées à la représentations des formes basales et apicales, introduites dans la Section 3.5, et construites pour recréer les formes canoniques utilisées en botanique pour décrire les feuilles et donc mener à la détermination de leur espèce. Ajuster ces modèles au contour obtenu constitue alors un moyen d'extraire facilement l'information pertinente tout en fournissant une représentation visuelle et explicite des formes basales et apicales.

5.3.1 Modèles paramétriques déformables

A la manière de ce que nous avons réalisé avec les différents modèles représentant la forme globale des feuilles simples comme composées, l'idée est d'utiliser une approche de type modèle déformable

optimisant les paramètres de construction des modèles de façon à ce qu'ils correspondent au mieux au contenu de l'image par minimisation d'énergie. Cependant, si l'évolution de ces premiers modèles reposait sur l'information de couleur extraite directement de l'image, nous disposons à ce stade d'une donnée un peu moins brute sous la forme du contour Γ^* de la feuille (ou de contours individuels Γ_f^* dans le cas de folioles).

Les modèles de formes basales et apicales sont très semblables, et ils se résument à une courbe Ψ (on distinguera par la suite Ψ_B et Ψ_A respectivement pour la base et pour l'apex), construite à partir d'un point de référence et de paramètres numériques :

$$\mathbf{p}_{Base} = \left(\alpha_B^*; o_B; \theta_B^g; \theta_B^d; \delta_B^g; \delta_B^d; \theta_g^B; \theta_d^B; \delta_g^B; \delta_d^B \right) \quad \mathbf{p}_{Apex} = \left(\alpha_A^*; o_A; \theta_A^g; \theta_A^d; \delta_A^g; \delta_A^d; \theta_g^A; \theta_d^A; \delta_g^A; \delta_d^A \right)$$

Ce sont les points détectés lors de l'interprétation du contour, respectivement B^* et A^* , qui vont jouer le rôle de points de référence. L'ensemble des modèles étant construit à partir de ces points, il nous a semblé plus judicieux de ne pas les considérer comme mobiles, de façon à ce que l'ajustement du modèle au contour se fasse uniquement par le biais des paramètres descriptifs. La détection correcte des points B^* et A^* sera donc une condition importante de la validité des modèles, mais c'est un choix moins hasardeux que de laisser ces points libres au risque qu'ils se déplacent loin du véritable apex par exemple.

L'objectif est alors d'obtenir des courbes Ψ qui soient représentatives de la forme locale du contour, autrement dit qui se trouvent aussi près que possible des points du contour. Par conséquent, la principale force qui doit guider l'évolution des modèles ne peut être que la distance des points de Ψ aux points de Γ . C'est donc cette distance qui est à la base de la définition de l'énergie d'attache aux données que la courbe doit minimiser au cours de son évolution.

Nous nous basons donc sur la transformée en distance du contour Γ^* qui produit une carte où chaque pixel $(x; y)$ de l'image est associé à sa distance au contour de la feuille :

$$\|(x; y) - \Gamma^*\| = \min_{(x'; y') \in \Gamma^*} \|(x; y) - (x'; y')\|$$

Au sein de la courbe Ψ , deux points ont une influence particulière : ce sont les extrémités des deux branches de la courbe (A_g^* et A_d^* pour l'apex). En décrivant la courbe ouverte Ψ par son abscisse curviligne, les points correspondent alors à $\Psi(0)$ et $\Psi(1)$. Ces extrémités doivent en effet absolument appartenir au contour pour que les valeurs d'angles et de longueurs des tangentes à la courbe en ces points aient un quelconque sens descriptif. Par conséquent, le poids donné à leur distance au contour doit être nettement plus important que pour les autres points de Ψ . On définit ainsi l'énergie d'attache à minimiser :

$$E_{attache}(\Psi) = \int_{\Psi} \|\Psi(u) - \Gamma^*\| du + \eta (\|\Psi(0) - \Gamma^*\| + \|\Psi(1) - \Gamma^*\|) \quad (5.14)$$

Les courbes sont alors libres d'évoluer, et leur définition leur permet de prendre des formes

asymétriques, qui ne manquent pas de se présenter sur les feuilles réelles. Il est toutefois important de limiter leur expressivité à des formes plausibles, et d'éviter des comportements aberrants qui donneraient naissance à des formes irréalistes. C'est pourquoi nous avons ajouté une contrainte souple de symétrie, qui se justifie par le fait que la plupart des feuilles présentent tout de même une forme globalement symétrique. Cette contrainte prend la forme d'un terme d'énergie interne, une somme de différences entre tous les paramètres de \mathbf{p} présidant à la construction du modèle (\mathbf{p}_{Base} ou \mathbf{p}_{Apex}) qui présentent une valeur pour la branche gauche et une valeur pour la branche droite de la courbe Ψ :

$$E_{interne}(\Psi) = \sum_{\{p_g, p_d\} \in \mathbf{p}} |p_g - p_d| \quad (5.15)$$

Le modèle, au cours de son évolution, devra donc minimiser une fonction d'énergie, que l'on formule comme une somme pondérée de son énergie d'attache, le forçant à se rapprocher du contour, et de son énergie interne, le poussant à conserver une forme pseudo-symétrique :

$$E(\Psi) = E_{attache}(\Psi) + \omega_{interne} E_{interne}(\Psi) \quad (5.16)$$

Le processus d'évolution qui permet de réaliser cette minimisation est très exactement le même que celui employé pour les modèles de forme globale, à savoir une descente de gradient de l'énergie. Le modèle Ψ est initialisé avec une valeur d'angle α_B^* ou α_A^* égale à la valeur correspondante α_B ou α_A obtenue à la fin de l'évolution du modèle polygonal Π^* , et son orientation de base est donnée par l'axe BA du même modèle polygonal. Pour le reste, les autres paramètres sont initialisés de façon à obtenir des segments rectilignes comme branches de la courbe.

L'évolution se fait alors de manière itérative en estimant pour chaque paramètre p de \mathbf{p} la dérivée d'énergie $\partial E / \partial p$ (aucune coordonnée de point n'est en revanche considérée cette fois-ci) et en appliquant des déplacements suivant la direction opposée du vecteur gradient d'énergie défini par l'ensemble de ces dérivées. Nous utilisons également une variable de température \mathcal{T} qui suit des cycles, afin d'éviter les minima locaux. L'Algorithme 7 reprend le déroulement du processus d'évolution pour le modèle de forme apicale, et la Figure 5.12 illustre l'évolution du modèle sur quelques images de feuilles en milieu naturel.

Le modèle ainsi optimisé constitue une représentation fidèle de la forme locale de la feuille, tout en portant un regard suffisamment synthétique pour ne pas être influencée par des détails de trop petite échelle (dents, défauts du contour) qui ne correspondraient pas à ce que les botanistes placent derrière les concepts de formes basales et apicales. La description fournie par les paramètres \mathbf{p}_{Base} et \mathbf{p}_{Apex} est un moyen efficace de condenser l'information de forme dans un jeu de descripteurs explicites, qui intègrent dans leur définition une part de savoir concernant ce qui va être déterminant pour la reconnaissance de l'espèce.

Algorithme 7 Optimisation du modèle de forme de l'apex par descente de gradient d'énergie

Entrée : Γ^* , contour obtenu après segmentation, A^* , point d'apex, Π^* modèle polygonal

Sortie : Ψ_A^* modèle d'apex après évolution

```

 $\Psi_A^{(0)} \leftarrow \Psi(A^*, \mathbf{p}_{Apex}^{(0)}(\Pi^*))$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{max}$ 
  while  $\mathcal{T} > \mathcal{T}_{min}$  do
    for all  $\mathbf{p} \in \{\alpha_A^*, o_A, \theta_A^g, \theta_A^d, \delta_A^g, \delta_A^d, \theta_g^A, \theta_d^A, \delta_g^A, \delta_d^A\}$  do
       $\mathbf{p}^{(cycle)} \leftarrow \mathbf{p}^{(cycle)} + \frac{1}{\partial_{max}} \left( \frac{\partial E}{\partial \mathbf{p}}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  OK  $\leftarrow$   $\|\Psi_A^{(cycle)} - \Psi_A^{(cycle-1)}\| < \Delta \Psi_{max}$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Psi_A^* \leftarrow \Psi_A^{(cycle)}$ 

```

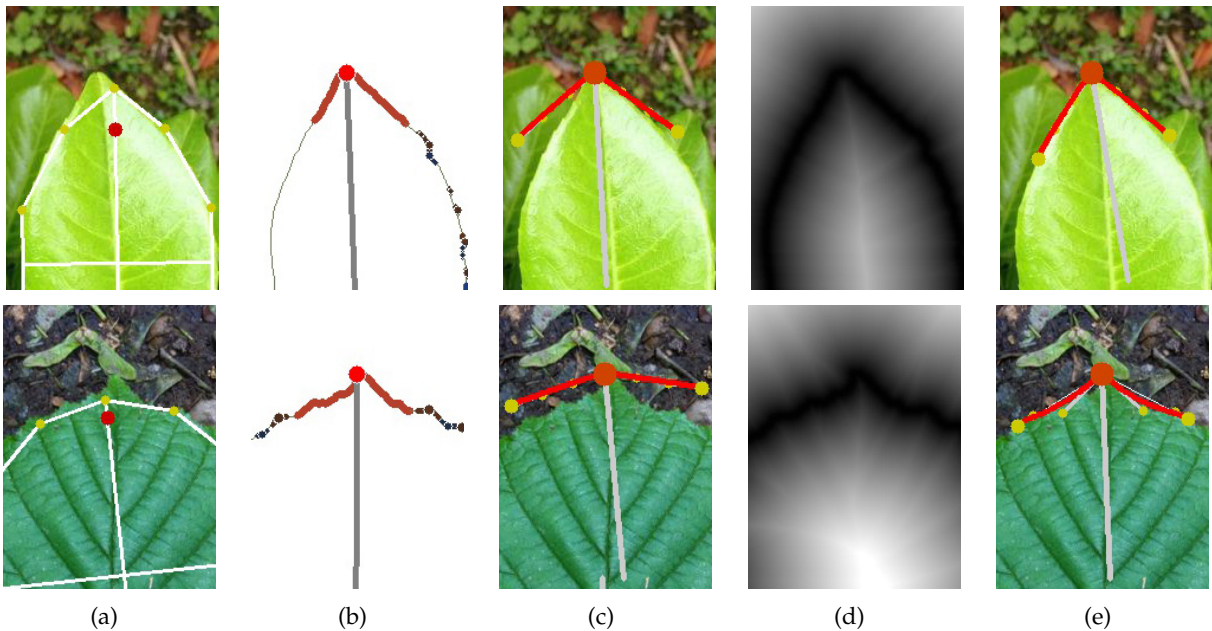


FIGURE 5.12 – Evolution du modèle paramétrique de forme apicale, à partir du modèle polygonal (a) et du contour interprété (b) : position initiale (c), transformée en distance du contour (d) et modèle final (e)

5.3.2 Réponse au problème du pétiole

Avec une telle approche, le pétiole est une vraie source de nuisance pour l'évaluation de la forme de la base, puisque, lorsqu'il est présent dans le contour, il en change radicalement la forme justement à l'endroit où on souhaite l'étudier. Il est pourtant assez rare que celui ci soit présent, généralement la différence de couleur avec le limbe est suffisante pour que le contour ne l'atteigne pas. Cela risque de se produire toutefois dans les images très contrastées, par ailleurs les plus faciles, où il est alors vu comme plus proche de la feuille.

L'interprétation du contour permet de localiser les points appartenant au pétiole, mais l'éliminer à ce stade semble périlleux. En effet le pétiole, bien que gênant, forme généralement une convexité très marquée sur le contour, qui garantit presque à coup sur que le point de base B^* sera placé à son extrémité. C'est ce qui apparaît sur les exemples de la Figure 5.5.

C'est là une certitude intéressante, puisque, s'il ne correspond pas au point d'insertion du pétiole, au moins le point B^* se trouve-t-il sur le bon axe, rendant certaines étapes de l'interprétation (position verticale des dents, construction du bord moyen, placement du modèle de base) plus robustes. Ceci ne serait plus tout à fait aussi certain si l'on supprimait le pétiole avant la détection de B^* . On aurait en effet rarement une convexité importante au point d'insertion, et le risque existerait de détecter un point du bord erroné et relativement éloigné.

Il faut donc faire avec, et ce n'est qu'au moment de l'optimisation de Ψ_B que cet excès de points du contour se révèle problématique. La solution que nous proposons tient en deux idées simples :

- ▶ Ne pas prendre en compte les points du pétiole pour guider l'évolution de la courbe Ψ_B
- ▶ Rendre le point B^* mobile pour que les valeurs des paramètres soient pertinentes

La première partie se traduit assez simplement dans le calcul de l'énergie d'attache minimisée par le modèle. En effet ce n'est simplement plus la distance des points au contour Γ^* qui doit être prise en compte, mais à ce même contour dont on a retiré les points appartenant au pétiole de la feuille, soit $\Gamma^* \setminus \Gamma_P$.

La deuxième condition est moins évidente. Nous ne voulons pas rendre le point B^* parfaitement libre, sans quoi l'information apportée par sa détection précise serait perdue, et la position finale du point sur le contour pourrait ne pas correspondre du tout au point d'insertion. Au lieu de cela, nous laissons le point libre de se déplacer dans une direction seulement, celle de l'axe B^*A^* ; pour cela nous rajoutons un paramètre p_B de position relative du point d'insertion réel B_B^* qui va venir remplacer B^* dans la construction du reste du modèle, tel que $\overrightarrow{B^*B_B^*} = p_B \overrightarrow{B^*A^*}$. Les paramètres à prendre en compte lors de la descente de gradient sont alors :

$$\mathbf{p}_{Base} = \left(p_B; \alpha_B^*; o_B; \theta_B^g; \theta_B^d; \delta_B^g; \delta_B^d; \theta_g^B; \theta_d^B; \delta_g^B; \delta_d^B \right)$$

La prise en compte de la distance au contour sans pétiole fait que le point B_B^* sera poussé à bouger, et qu'il devra en fait coulisser le long du pétiole pour atteindre son point d'insertion. On s'en

assure en forçant le point B_B^* à appartenir au contour amputé, de la même façon que pour les points aux extrémités, par le biais de l'énergie d'attache. Par conséquent, pour le modèle de forme basale, l'expression de l'énergie d'attache devient :

$$E_{attache}(\Psi) = \int_{\Psi} \|\Psi(u) - \Gamma^* \setminus \Gamma_P\| du + \eta (\|B_B^* - \Gamma^* \setminus \Gamma_P\| + \|\Psi(0) - \Gamma^* \setminus \Gamma_P\| + \|\Psi(1) - \Gamma^* \setminus \Gamma_P\|) \quad (5.17)$$

Ainsi, comme le montre la Figure 5.13, même en partant d'un point détecté qui n'est pas le point d'insertion autour duquel on définit la forme basale, le modèle parvient quand même à se rapprocher du contour et à coïncider avec la base de la feuille. La description produite est alors nettement plus significative et représentative de la forme de la feuille qu'elle ne l'aurait été si le point était demeuré fixe.

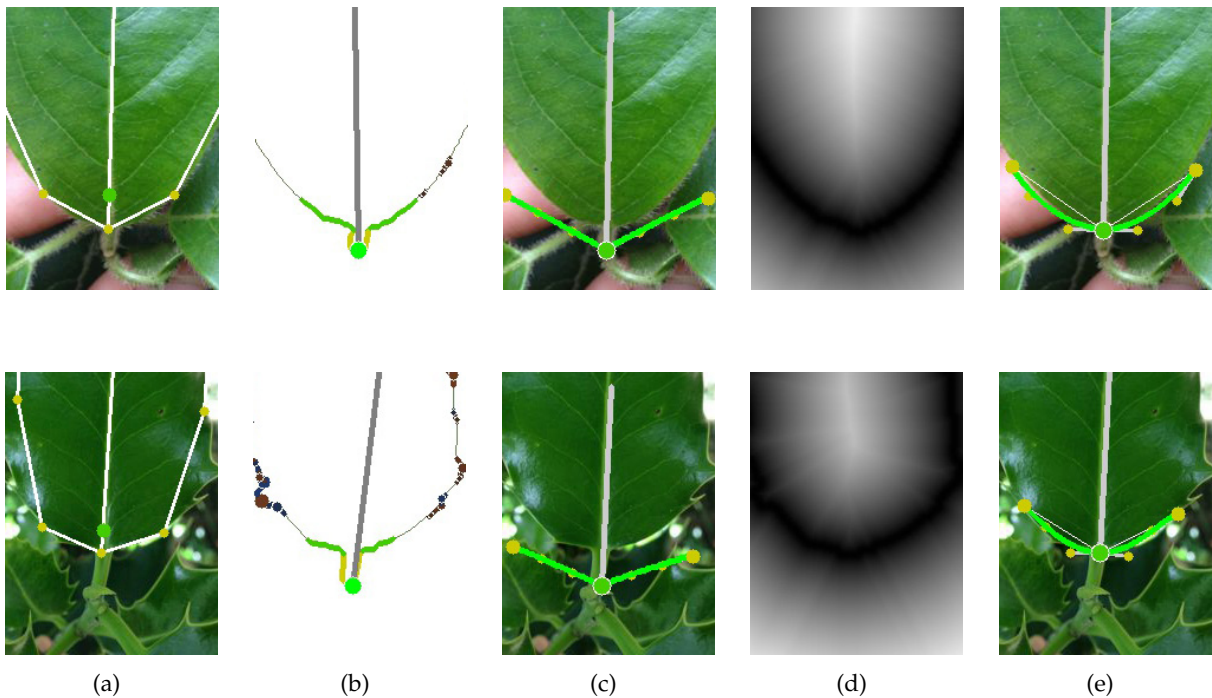


FIGURE 5.13 – Traitement du pétiole par le modèle paramétrique de forme basale, à partir du modèle polygonal (a) et du contour interprété (b) : position initiale à l'extrémité du pétiole (c), transformée en distance du contour amputé (d) et modèle final (e)

Cette évaluation d'un point à la signification forte comme le point d'insertion se produit donc conjointement avec d'évaluation d'une forme locale, dans un unique processus d'optimisation. Le détail de cette méthode, très proche de celle utilisée pour la forme de l'apex, est donnée dans l'Algorithme 8. D'un point de vue de l'implémentation, ces deux évolutions se font en parallèle, les cycles de températures étant les mêmes pour les modèles Ψ_A et Ψ_B . Le résultat de cette évolution conjointe est donnée à la Figure 5.14.

Algorithme 8 Optimisation du modèle de forme de la base par descente de gradient d'énergie

Entrée : Γ^* , contour obtenu après segmentation, B^* , point de base, Π^* modèle polygonal

Sortie : Ψ_A^* modèle de base après évolution

```

 $\Psi_B^{(0)} \leftarrow \Psi(B^*, \mathbf{p}_{Base}^{(0)}(\Pi^*))$ 
OK  $\leftarrow$  false
cycle  $\leftarrow$  1
while  $\neg$ OK do
   $\mathcal{T} \leftarrow \mathcal{T}_{max}$ 
  while  $\mathcal{T} > \mathcal{T}_{min}$  do
    for all  $\mathbf{p} \in \{p_B, \alpha_B^*, o_B, \theta_B^s, \theta_B^d, \delta_B^s, \delta_B^d, \theta_g^B, \theta_d^B, \delta_g^B, \delta_d^B\}$  do
       $\mathbf{p}^{(cycle)} \leftarrow \mathbf{p}^{(cycle)} + \frac{1}{\partial_{max}} \left( \frac{\partial E}{\partial \mathbf{p}}(\mathcal{T}) \times \Delta \mathbf{p}(\mathcal{T}) \right)$ 
    end for
     $\mathcal{T} \leftarrow \lambda \times \mathcal{T}$ 
  end while
  OK  $\leftarrow$   $\|\Psi_B^{(cycle)} - \Psi_B^{(cycle-1)}\| < \Delta \Psi_{max}$ 
  if  $\neg$ OK then
    cycle  $\leftarrow$  cycle + 1
  end if
end while
 $\Psi_B^* \leftarrow \Psi_B^{(cycle)}$ 

```

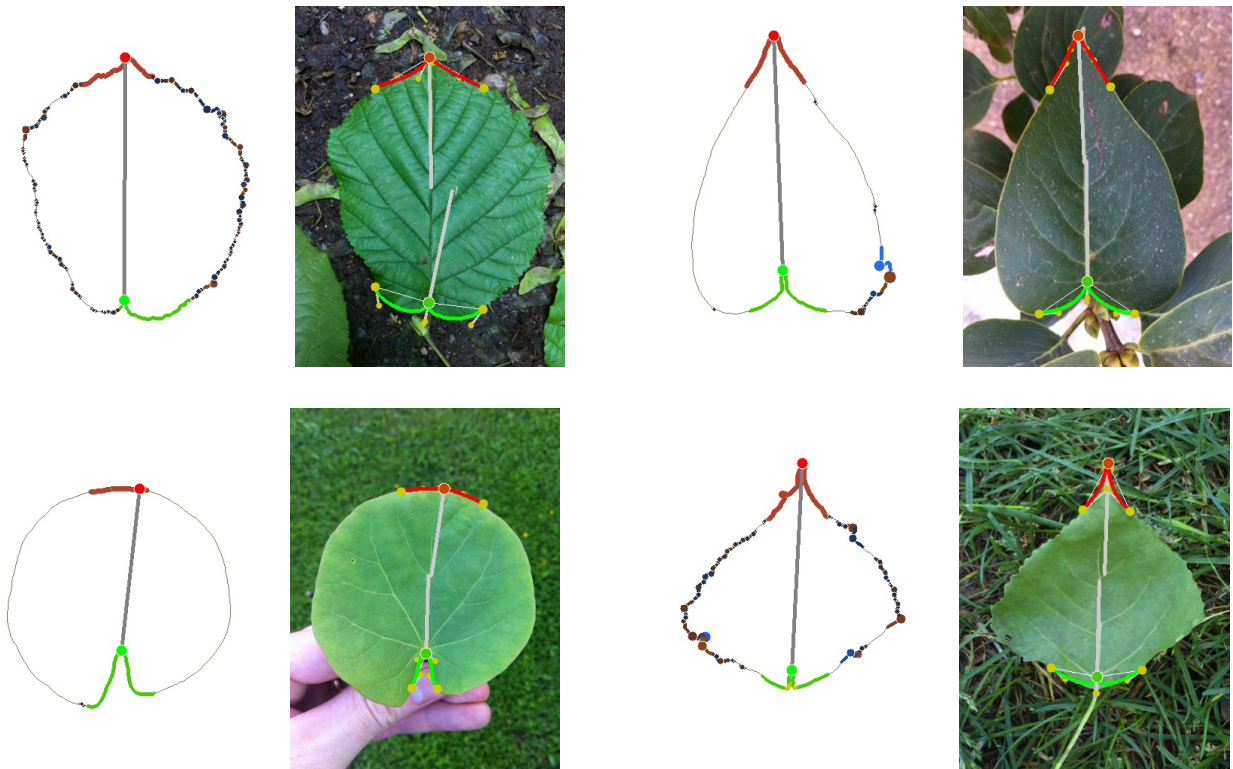


FIGURE 5.14 – Contours interprétés et évolution conjointe des modèles de base et d'apex sur plusieurs images de feuilles en environnement naturel

Dans les deux cas, l'optimisation des modèles paramétriques Ψ_A^* et Ψ_B^* produit un ensemble de valeurs de paramètres, celles qui donnent la forme la mieux adaptée au contour de la feuille, qui sont représentatives des spécificités de la forme représentée. Pour une meilleure généralisation, et un peu à la manière de ce qui est fait pour les dents du bord, nous allons considérer comme paramètres descriptifs non pas l'ensemble des valeurs, mais une moyenne de ce qui a été trouvé sur le côté gauche et le côté droit, puisque l'interprétabilité des représentation que nous utilisons nous permet de faire cette distinction. De plus, certains paramètres sont purement des paramètres de construction, et par conséquent assez peu porteurs de sens. Nous conservons ainsi au final deux vecteurs de paramètres qui fournissent une description des formes basales et apicales :

$$\mathbf{p}_{Base}^* = \left(\alpha_B^*; \frac{\theta_B^g + \theta_B^d}{2}; \frac{\delta_B^g + \delta_B^d}{2}; \frac{\theta_g^B + \theta_d^B}{2}; \frac{\delta_g^B + \delta_d^B}{2} \right)$$

$$\mathbf{p}_{Apex}^* = \left(\alpha_A^*; \frac{\theta_A^g + \theta_A^d}{2}; \frac{\delta_A^g + \delta_A^d}{2}; \frac{\theta_g^A + \theta_d^A}{2}; \frac{\delta_g^A + \delta_d^A}{2} \right)$$

5.4 Pertinence de la description

Les descripteurs de formes foliaires locales introduits dans ce chapitre, tout comme les paramètres des formes globales de feuilles évalués au cours de la première phase de la segmentation (Chapitre 4), constituent une interprétation de haut niveau du contenu de l'image. Leur traduction en concepts intelligibles par un utilisateur humain est quasiment immédiate, et ils permettent également une réversibilité vers la forme assez intéressante. On peut toutefois s'interroger sur l'efficacité d'une description située à un aussi haut niveau sémantique dans un contexte d'apprentissage automatique, et sur la pertinence de l'introduction d'une telle quantité de connaissances expertes à un stade si anticipé du processus de reconnaissance.

Une telle conception de la description des formes se trouve en effet en opposition relativement frontale avec l'approche plus classique des descripteurs de formes génériques et souvent statistiques qui s'attachent à représenter les propriétés d'une forme digitale sans se risquer à la moindre interprétation. Des mesures comme les moments qualifient les formes de façon objective et uniforme, en capturant toujours de la même manière leurs propriétés globales, et se faisant ainsi des témoins fidèles de leurs caractéristiques. De telles méthodes laissent alors la tâche d'interprétation aux phases suivantes, se cantonnant à un rôle purement descriptif.

Cela peut apparaître comme une faiblesse des descripteurs génériques, qui semblent ainsi capturer moins d'information, ou du moins une information moins pertinente, sur la forme. Ce n'est pas tout à fait exact, et de bons descripteurs de forme vont en fait représenter la même quantité d'information, mais de façon brute et non-interprétée, sans pour autant en être moins discriminante. C'est même bien souvent un gage de robustesse, puisque qu'en ne supposant rien de l'objet étudié, on évite de faire

des hypothèses pour l'extraction de mesures qui pourraient se révéler fausses.

A l'inverse, les descripteurs des formes locales et du contour, on ne peut plus spécifiques, supposent énormément de choses sur la forme sur laquelle ils s'appliquent, à savoir en premier lieu qu'il s'agit d'une feuille bien segmentée ; que chaque concavité ou convexité est bel et bien le reflet d'un creux ou d'une dent sur le bord de la feuille ; que le point sur lequel s'attache le modèle de sommet correspond exactement à la pointe de la feuille et qu'il va pouvoir s'ajuster sur les autres points du contour, qui suivent effectivement le bord réel du limbe. Plus la description intervient tard dans le processus d'interprétation, plus la somme des hypothèses faites en amont accroît le risque d'une interprétation erronée.

En tout cas, ces représentations sont initialement conçues avec une part probable de naïveté, en les supposant capables de superposer trois aspects *a priori* difficilement conciliables :

- ▶ **Construction** : le fait de pouvoir passer d'un jeu de valeurs directement à une forme (comme c'est le cas notamment avec les modèles paramétriques) et donc l'expressivité des paramètres, leur capacité à créer des formes variées et plausibles, avec la contrainte que différentes valeurs ne permettent pas de générer la même forme.
- ▶ **Description** : la dimension la plus évidente, la capacité des valeurs des différents paramètres à représenter fidèlement la forme sur laquelle ils ont été extraits, c'est-à-dire que des formes proches produisent des valeurs proches et que des valeurs très différentes ne puissent que correspondre à des formes éloignées.
- ▶ **Signification** : la possibilité de tirer une information intelligible de la valeur d'un ou plusieurs paramètres, le lien direct qui peut exister entre un paramètre et un élément mesurable par un utilisateur, autrement dit la cohérence de la description non seulement avec la forme mais avec les concepts sémantiques qu'on peut y apposer.

Pourtant, une description sémantiquement riche et basée sur une connaissance de l'objet possède plusieurs attraits, à commencer par le fait qu'elle soit plus pertinente car focalisée sur les critères qui vont permettre de prendre la meilleure décision. Les informations à extraire d'une feuille pour identifier son espèce sont depuis bien longtemps étudiées, et on ne saurait probablement trouver mieux pour mener à bien cette tâche que ce que la science botanique nous a laissé en héritage au fil des siècles. Le risque serait alors moins grand d'égarer en chemin des informations cruciales et ponctuelles lorsque l'on sait à l'avance quoi chercher, mais cela est aussi valable pour une description générique exhaustive.

Là où la véritable différence peut se faire en interprétant la forme est au contraire dans la possibilité d'ignorer volontairement certaines de ses propriétés qui ne correspondraient pas à une préconception de l'objet. C'est une chose que sont incapables d'effectuer des descripteurs "transparents" de la forme, puisqu'ils reflèteront toujours son contenu quel qu'il soit. Interpréter le contour permet de décorrélérer la description de certains aspects de la forme, par exemple en imposant de trouver une dent sur les deux bords de la feuille pour la prendre en compte, ou de s'appuyer sur un modèle rigide pour

atteindre un certain niveau de généralisation.

Le but visé par cette description est en lien avec l'identification des espèces et la volonté de modéliser au mieux leurs propriétés de forme. Le point crucial qui pose le plus de problèmes au vu de la reconnaissance est celui de la variabilité intra-classe. Plus précisément, cette variabilité va s'exprimer différemment selon les critères de forme. Certains pourront être très variables au sein d'une espèce et seront à ignorer. D'autres cependant resteront stables seront donc particulièrement intéressants pour la reconnaissance. Ces distinctions se font en outre différemment selon les espèces, mais il semble déterminant de pouvoir le faire. C'est avec cette idée que les descripteurs que nous utilisons ont été pensés de façon à décorrélérer l'information relative aux différentes formes foliaires, et à ne pas traiter la forme en un bloc.

Enfin, la charge sémantique des différents descripteurs, et la prise en compte d'un seul critère de forme à la fois permet d'envisager une interprétation intermédiaire, avant la reconnaissance des espèces, une reconnaissance de concepts botaniques. En introduisant cette couche sémantique sur les valeurs extraites, il serait possible d'offrir une interprétation de l'information extraite de l'image dans un langage intelligible pour un utilisateur, et de réaliser ainsi un lien entre la feuille comme objet concret, les représentations visuelles des mesures, et les concepts utilisés en botanique, par ailleurs pas toujours évidents à appréhender.

La description des formes foliaires locales proposée dans ce chapitre permet d'envisager le problème de reconnaissance des espèces d'arbres sous un angle assez inhabituel puisqu'une grande part de connaissance sur l'objet d'intérêt se voit incluse dans la définition des descripteurs. Il en ressort une représentation très explicite des formes, évaluées de manière relativement indépendante, de façon à faciliter l'identification qui va suivre et sera détaillée au chapitre suivant. Les outils définis pour manipuler ces descripteurs offrent plusieurs possibilités pour l'identification des espèces, passant par la comparaison ou l'apprentissage automatique.

Il sera dans tous les cas intéressant d'évaluer l'influence de l'introduction de connaissance dans une description de haut-niveau des formes, en confrontant cette description à des mesures plus génériques dans le cadre de la reconnaissance. Mesurer explicitement la pertinence de chacun des descripteurs n'est pas une tâche aisée, mais des outils existent pour y aider. Au premier abord, il semble que tant l'identification des espèces que l'étiquetage par des concepts botaniques doivent pouvoir bénéficier de descripteurs conçus explicitement pour représenter les formes foliaires. Mais encore faut-il trouver les moyens de traiter au mieux les problèmes spécifiques liés à la variabilité des objets considérés, raison pour laquelle a été pensée cette description dédiée des formes des feuilles.

Bibliographie

- [AMK97] Sadegh Abbasi, Farzin Mokhtarian, and Josef Kittler. Reliable classification of chrysanthemum leaves through curvature scale space. In *International Conference on Scale-Space Theory in Computer Vision*, pages 284–295, 1997.
- [CA10] Carlos Caballero and M. Carmen Aranda. Plant species identification using leaf image retrieval. In *Proceedings of the ACM International Conference on Image and Video Retrieval, CIVR 2010*, pages 327–334, 2010.
- [FZ00] Igor Fischer and Andreas Zell. String averages and self-organizing maps for strings. In *Proceedings of the Neural Computation 2000*, pages 208–215, 2000.
- [Jol03] Jean-Michel Jolion. The deviation of a set of strings. *Pattern Analysis and Applications*, 6(3) :224–231, 2003.
- [KBB⁺12] Neeraj Kumar, Peter Belhumeur, Arijit Biswas, D. Jacobs, I. Kress, W. Lopez, and Soares Joao V. B. Leafsnap : A computer vision system for automatic plant species identification. In *European Conference on Computer Vision*, volume 2, pages 502–516, 2012.
- [KS98] Teuvo Kohonen and Panu Somervuo. Self-organizing maps of symbol strings. *Neurocomputing*, 21(1-3) :19–30, 1998.
- [Lev66] Vladimir I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8) :707–710, 1966.
- [LL07] Yujian Li and Bo Liu. A normalized levenshtein distance metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6) :1091–1095, 2007.
- [MA04] F. Mokhtarian and S. Abbasi. Matching shapes with self-intersections : Application to leaf classification. *IEEE Transactions on Image Processing*, 13(5) :653–661, 2004.
- [MM92] F. Mokhtarian and A.K. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14 :789–805, 1992.

- [MV93] Adrés Marzal and Enrique Vidal. Computation of normalized edit distance and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9) :926–932, 1993.
- [PL92] S. C. Pei and C. N. Lin. The detection of dominant points on digital curves by scale-space filtering. *Pattern Recognition*, 25(11) :1307–1314, 1992.
- [SL98] Marc Sebban and Anne M. Landraud. String clustering and statistical validation of clusters. In *Proceedings of the 12th Biennial Conference of the Canadian Society for Computational Studies of Intelligence on Advances in Artificial Intelligence, AI '98*, pages 298–309, 1998.
- [TC89] C. H. Teh and R. T. Chin. On the detection of dominant points on digital curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8) :859–872, 1989.
- [WF74] Robert A. Wagner and Michael J. Fischer. The string-to-string correction problem. *Journal of the ACM*, 21(1) :168–173, 1974.
- [WHD⁺08] X.-F. Wang, D.-S. Huang, J.-X. Du, X. Huan, and L. Heutte. Classification of plant leaf images with complicated background. *Applied Mathematics and Computation*, 205(2) :916–926, 2008.

Chapitre 6

Classification et fusion de données

Modèles de décision pour l'identification des espèces




Populus alba L.

Résumé du Chapitre

Dans ce chapitre, nous nous intéressons à la façon dont sont utilisés les descripteurs extraits des images pour véritablement identifier l'espèce de la feuille qu'ils représentent. Nous détaillons donc les différentes méthodes d'apprentissage automatique et de décision que nous avons employées pour parvenir à cette détermination. Un cadre basique de classification utilisant des modèles gaussiens est tout d'abord utilisé pour juger de la pertinence des attributs de forme introduits et de leur compétitivité face à d'autres descripteurs plus génériques. Nous présentons également un travail original de classification des feuilles non pas en espèces mais en vocabulaire botanique de formes, et ses possibles applications. Enfin une étude approfondie de la classification par des forêts aléatoires, donnant des résultats supérieurs aux autres méthodes, et ses résultats dans le cadre de la participation au challenge ImageCLEF viennent conclure le chapitre.

La Section 6.2 présente des travaux menés en parallèle avec Honghui LIU, post-doctorant au LISTIC.

La Section 6.3 présente des résultats développés par Violaine ANTOINE, post-doctorante au LISTIC.

ES paramètres extraits et les descripteurs représentatifs des feuilles contenues dans les images ont pour vocation de constituer une base d'exemples dont l'espèce est connue, permettant par la suite d'identifier l'espèce de nouveaux individus. Le processus de classification qui vise à attribuer une classe, ici une espèce, à un objet représenté par ses descripteurs va s'appuyer sur cet ensemble de données pour prendre une décision.

Il s'agit dans tous les cas d'un problème supervisé, puisque les classes qui nous intéressent ne sont pas directement extractibles à partir des données mais proviennent d'une définition préexistante. Chaque feuille correspond en effet à une espèce et une seule, et l'ensemble de ces catégories forme une partition de l'univers des feuilles qui est celle que nous souhaitons capturer le plus fidèlement possible. L'objectif est donc de localiser un individu dans cet univers bien partitionné des feuilles, en se basant sur les valeurs des descripteurs représentant la forme de la feuille.

Les approches pour y parvenir sont multiples et on peut les séparer en deux grandes classes de méthodes. La première consiste à utiliser la base d'exemples telle quelle et à comparer un nouvel individu à l'ensemble des exemples considérés pour prendre une décision. La seconde relève de l'apprentissage automatique et de la création d'un modèle à partir des données d'exemples, de façon à obtenir un outil de décision plus compact, plus rapide, et offrant une généralisation des données fournies.

Les méthodes détaillées dans la suite relèvent de la deuxième catégorie, pour laquelle l'essentiel du processus (l'apprentissage) se déroule hors-ligne, la décision elle-même s'effectuant de manière très rapide. C'est une condition indispensable pour envisager un traitement complet sur un smartphone dont les capacités de mémoire et de calcul sont limitées.

6.1 Modèles d'espèces par agrégation

De façon à évaluer le plus simplement possible la pertinence des descripteurs de forme introduits dans les chapitres précédents, nous avons mis en place une méthode de classification naïve visant à représenter chaque espèce par un unique modèle moyen. La construction de ces modèles peut être facilement envisagée à partir des valeurs vectorielles des paramètres qui peuvent être agrégées, moyennées, de façon à obtenir une représentation canonique de la classe dans l'espace des paramètres.

Les modèles d'espèces que l'on obtient alors vont pouvoir servir de références auxquelles comparer un nouvel individu en vue de sa classification. Il s'agit d'une idée des plus basiques qui permet *a priori* de mesurer à quel point les paramètres extraits sont représentatifs des espèces. Néanmoins, pour que la modélisation soit complète, il faut que la représentation des espèces prenne aussi en compte leur variabilité qui est, on l'a vu, une caractéristique inévitable du problème d'identification des espèces d'arbres (Chapitre 1).

6.1.1 Modèles gaussiens d'espèces par descripteur

Une manière intuitive de représenter les valeurs prises par les paramètres vectoriels d'une classe est d'utiliser une modélisation gaussienne de leur distribution. Cela suppose que les descripteurs suivent une loi gaussienne multivariée qui peut alors simplement être caractérisée par un point moyen et une matrice de variance-covariance. Cette modélisation présente l'avantage de pouvoir être estimée très simplement à partir d'un ensemble de vecteurs. Le calcul de ces deux grandeurs statistiques définit une distribution qui permet d'évaluer la vraisemblance de tout point de l'espace pour savoir s'il correspond à la classe considérée.

Les attributs vectoriels que nous utilisons pour la description des formes de feuilles sont constitués de paramètres numériques issus des différents modèles de formes foliaires, et l'on peut ainsi distinguer plusieurs vecteurs :

- ▶ $\mathbf{p}_{Polygone}$: les paramètres du modèle polygonal de forme globale, permettant aussi de décrire la forme des lobes des feuilles palmatilobées et des folioles de feuilles composées (5 attributs).
- ▶ \mathbf{p}_{Lobes} : les paramètres spécifiques aux lobes décrivant leur agencement ; étant en nombre variable ($2 \times n_L$) nous fixons la taille du vecteur à 6 attributs en tronquant dans le cas où $n_L > 3$.
- ▶ $\mathbf{p}_{Folioles}$: les paramètres du modèle d'organisation de feuille composée, uniquement pour les espèces à feuilles composées, décrivant le nombre et l'agencement des folioles (4 attributs).
- ▶ \mathbf{p}_{Base} et \mathbf{p}_{Apex} : les paramètres obtenus après évolution des modèles de base et d'apex, et représentant la forme locale de la feuille à ces emplacements (5 et 5 attributs).
- ▶ \mathbf{p}_{Bord} : la version numérique de la description du bord de la feuille, qui peut être utilisée à la place de la représentation structurale, et traitée de la même façon que les autres descripteurs (13 attributs).

Le point fondamental est l'hypothèse selon laquelle ces vecteurs de descripteurs peuvent être considérés comme indépendants entre eux. En effet ils représentent des propriétés des formes foliaires qui ne sont pas supposées dépendre les unes des autres (on peut selon toute vraisemblance trouver la même forme au sommet quelle que soit la forme globale de la feuille, ou celle de son bord) et chaque vecteur d'attributs peut être considéré comme une source d'information à part. Il est donc légitime de ne pas prendre en compte la covariance d'attributs de vecteurs différents (supposément nulle) ce qui revient à ne pas considérer tous les attributs comme un seul vecteur par concaténation, mais bien à conserver un vecteur par jeu de descripteurs.

Tous les attributs correspondant généralement à des grandeurs géométriques ou à des mesures diverses, ils ne sont nécessairement pas définis sur la même échelle. Pour ne pas que cette disparité influe sur la modélisation des classes, la première opération du traitement consiste à normaliser l'ensemble des données de la base d'exemples \mathbf{S} . Pour chaque attribut numérique p , nous calculons ainsi sa moyenne μ_p et sa variance σ_p^2 sur l'ensemble de la base d'apprentissage, de façon à produire une nouvelle valeur centrée-réduite :

$$p^* = \frac{p - \mu_p}{\sigma_p}$$

Dans la suite on considérera que tous les vecteurs d'exemples (par exemple $\mathbf{p}_{Polygone}$) contiennent des valeurs centrées-réduites. Nous considérons alors dans le cadre de l'apprentissage automatique plusieurs bases d'exemples, une par ensemble d'attributs, étiquetées par l'espèce de la feuille. Les espèces considérées forment un ensemble $\mathbb{E} = \{E_e\}_{e=1..|\mathbb{E}|}$, et chaque exemple numéroté par un indice x peut être rattaché à une espèce E_e par la donnée d'une étiquette $\epsilon(x) = e$. Par exemple, pour la forme globale, les exemples forment ainsi une base $\mathbf{S}_{Polygone} = \left\{ \left(\mathbf{p}_{Polygone,x}; E_{\epsilon(x)} \right) \right\}_{x=1..|\mathbf{S}|}$. On peut diviser cette base d'apprentissage en autant de sous-ensembles que de classes, en considérant les vecteurs d'exemples associés à chaque espèce E_e de \mathbb{E} dans un sous-ensemble $\mathbf{S}_{Polygone,e} = \left\{ \mathbf{p}_{Polygone,x} \mid \epsilon(x) = e \right\}_{x=1..|\mathbf{S}|}$.

Les différents modèles gaussiens représentant les espèces (un par espèce E_e et par jeu de descripteurs F) peuvent alors être estimés très simplement en calculant moyennes et matrices de covariance sur ces sous-ensembles $\mathbf{S}_{F,e}$ de la base. Pour chaque descripteur de forme F et pour chaque espèce E_e , on obtient ainsi un modèle $\mathbf{G}_{F,e} = (\mathbf{P}_{F,e}; \mathbf{\Sigma}_{F,e})$ qui représente la distribution des attributs $\mathbf{p}_{F,x}$ décrivant la forme considérée pour l'espèce en question.

$$\mathbf{P}_{F,e} = \frac{1}{|\mathbf{S}_{F,e}|} \sum_{\mathbf{S}_{F,e}} \mathbf{p}_{F,x} \quad \mathbf{\Sigma}_{F,e} = \frac{1}{|\mathbf{S}_{F,e}|} \sum_{\mathbf{S}_{F,e}} (\mathbf{p}_{F,x} - \mathbf{P}_{F,e}) (\mathbf{p}_{F,x} - \mathbf{P}_{F,e})^T \quad (6.1)$$

La représentation gaussienne des espèces permet de condenser sous une forme particulièrement compacte la distribution des attributs de forme des feuilles d'une espèce. La prise en compte de la covariance des attributs au sein d'un jeu de descripteurs permet une quantification de la variabilité de l'espèce sans occulter la corrélation qui peut exister entre les différents paramètres de description. Il s'agit d'une méthode d'apprentissage très simple qui permet surtout de mesurer la pertinence des descripteurs utilisés puisque le modèle gaussien constitue une conversion directe des points définis par les valeurs des paramètres en une distribution moyenne.

6.1.2 Une distance adaptée à la variabilité des classes

La classification d'un nouvel individu dans ce processus passe par la comparaison de ses descripteurs avec les différents modèles représentant chacune des espèces possibles. La connaissance d'un point moyen, représentant canonique de l'espèce, permet d'envisager un calcul de distance pour déterminer l'espèce la plus proche, à la manière d'un algorithme de plus proche voisin basé sur les modèles d'espèces. Cependant, la variabilité intra-classe, ainsi que la similarité inter-classe, jouent un rôle important et viennent compliquer cette tâche de comparaison. Elles doivent en tout cas être prises en compte d'une manière ou d'une autre pour ne pas s'exposer à des erreurs de classification.

Pour comparer un vecteur de descripteurs \mathbf{p}_F avec le modèle gaussien $(\mathbf{P}_{F,e}; \mathbf{\Sigma}_{F,e})$ d'une espèce, plusieurs possibilités existent pour la définition de la fonction de distance à calculer. La plus simple

est sans doute la distance euclidienne basée sur la norme 2 et qui s'écrit simplement :

$$\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_2 = \left((\mathbf{p}_F - \mathbf{P}_{F,e})^T (\mathbf{p}_F - \mathbf{P}_{F,e}) \right)^{\frac{1}{2}}$$

Cette distance ne prend pas du tout en compte la variabilité des classes, et se contente de mesurer la distance d'un nouvel individu aux représentants canoniques des espèces. Il semble évident que ces modèles idéaux seuls ne sauraient suffire à rendre compte de toutes les formes possibles prises par une espèce, d'autant que celles-ci peuvent se montrer très variables pour certains critères, ou au contraire très stables pour d'autres. Il apparaît donc comme une nécessité d'utiliser l'information de variance-covariance pour évaluer au mieux la distance d'un individu à une classe.

Une distance permettant de prendre en compte la variabilité des classes serait la distance de Mahalanobis [Mah36], qui introduit la matrice de covariance (ou plus exactement son inverse, qui existe toujours dès lors que la matrice est définie positive) dans le calcul de la distance :

$$\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M = \left((\mathbf{p}_F - \mathbf{P}_{F,e})^T \boldsymbol{\Sigma}_{F,e}^{-1} (\mathbf{p}_F - \mathbf{P}_{F,e}) \right)^{\frac{1}{2}}$$

Le but est de donner moins de poids aux directions de l'espace des paramètres où la classe est la plus variable (où les données sont potentiellement les plus bruitées) et plus de poids là où elle l'est le moins (où les valeurs vont être le plus représentatives de la classe). Cette distance convient parfaitement pour comparer deux points provenant d'une même distribution, mais elle pose problème dès lors que l'on compare un point, dont la distribution est inconnue, à plusieurs distributions différentes.

En effet, la distance de Mahalanobis présente le gros désavantage de distordre, de contracter ou d'étirer, l'espace des paramètres de manière bien différente pour les différentes classes. Cela constitue un gros obstacle pour notre cas d'étude où les espèces peuvent présenter des variabilités bien distinctes selon les critères. Le risque principal est qu'un vecteur se trouvant très proche du point moyen d'une classe apparaisse encore moins distant, avec la distance de Mahalanobis, d'un autre point plus éloigné si sa classe se trouve être très variable dans cette direction. Les classes les plus variables tendent ainsi à "absorber" les plus stables, qui du fait de la distorsion apparaîtront toujours comme plus lointaines. Il s'agit d'une limitation très pénalisante pour les classes peu variables (ou celles avec peu d'exemples) qui nuit forcément à l'identification des espèces.

Pour éviter cette distorsion, il faut absolument que les distances d'un point à toutes les distributions soient calculées sur la même échelle, comme le fait la distance euclidienne. Pour autant la variabilité des classes, sous la forme de la matrice $\boldsymbol{\Sigma}_{F,e}$ doit être prise en compte. Nous proposons donc de représenter une espèce E_e non pas par son seul point moyen $\mathbf{P}_{F,e}$ mais par l'hyper-ellipsoïde centré sur $\mathbf{P}_{F,e}$ et dont les axes et les dimensions sont donnés par la matrice $\boldsymbol{\Sigma}_{F,e}$ (par la donnée de ses valeurs et vecteurs propres par exemple). Cet ellipsoïde $\mathcal{E}_{F,e}$ forme un hyper-volume dont la définition peut

en fait être donnée via la distance de Mahalanobis :

$$\mathcal{E}_{F,e}(\mathbf{P}_{F,e}, \boldsymbol{\Sigma}_{F,e}) = \left\{ p_F \mid (p_F - \mathbf{P}_{F,e})^T \boldsymbol{\Sigma}_{F,e}^{-1} (p_F - \mathbf{P}_{F,e}) \leq 1 \right\} = \left\{ p_F \mid \|p_F - \mathbf{P}_{F,e}\|_M \leq 1 \right\} \quad (6.2)$$

Si un vecteur de paramètres se trouve à l'intérieur de l'ellipsoïde ainsi défini, il appartient à la région de l'espace délimitant la classe représentée. Par conséquent, on peut considérer que sa distance au modèle de l'espèce est nulle. Pour le reste, nous considérons que la distance d'un point \mathbf{p}_F à la classe E_e peut être donnée par la distance du point à la **surface de l'ellipsoïde**. On note cette dernière $\mathcal{S}_{F,e}$:

$$\mathcal{S}_{F,e}(\mathbf{P}_{F,e}, \boldsymbol{\Sigma}_{F,e}) = \left\{ p_F \mid (p_F - \mathbf{P}_{F,e})^T \boldsymbol{\Sigma}_{F,e}^{-1} (p_F - \mathbf{P}_{F,e}) = 1 \right\} = \left\{ p_F \mid \|p_F - \mathbf{P}_{F,e}\|_M = 1 \right\} \quad (6.3)$$

La distance que nous choisissons d'utiliser est la distance euclidienne, donc sans distorsion inconstante de l'espace des paramètres, à un point de la surface $\mathcal{S}_{F,e}$. Pour des raisons de simplicité, nous avons décidé de considérer la distance du point \mathbf{p}_F au point $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ de la surface $\mathcal{S}_{F,e}$ situé dans la même direction que \mathbf{p}_F depuis le point moyen $\mathbf{P}_{F,e}$. La Figure 6.1 illustre cette définition dans un cas de figure simple à deux dimensions.

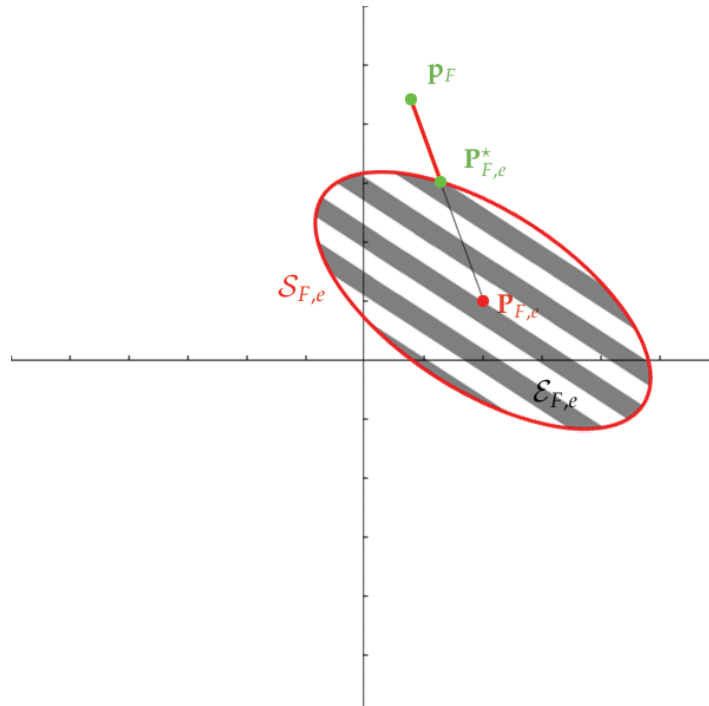


FIGURE 6.1 – Exemple de distance d'un point à la surface de l'ellipsoïde représentant une classe, dans un espace bi-dimensionnel

La distance considérée comme représentative de la proximité d'un point à la classe peut alors être appelée distance à la l'ellipsoïde et notée $\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}}$ et s'écrit simplement :

$$\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}} = \|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_2 \quad (6.4)$$

Il ne s'agit généralement pas de la distance euclidienne à la surface $S_{F,e}$ comme on le conçoit habituellement, à savoir :

$$\|\mathbf{p}_F - S_{F,e}\|_2 = \min_{p_F \in S_{F,e}} \|\mathbf{p}_F - p_F\|_2$$

Le point de la surface qui serait alors considéré serait en effet le projeté orthogonal de \mathbf{p}_F sur la surface, ce qui n'est ici pas le cas. En revanche, par définition, le point $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ vérifie de manière générale les deux propriétés suivantes :

$$\begin{cases} (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})^T \Sigma_{F,e}^{-1} (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e}) = 1 \\ \exists k \in \mathbb{R}^{+*}, (\mathbf{p}_F - \mathbf{P}_{F,e}) = k (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e}) \end{cases} \quad (6.5)$$

On peut par conséquent écrire :

$$\begin{aligned} (\mathbf{p}_F - \mathbf{P}_{F,e})^T \Sigma_{F,e}^{-1} (\mathbf{p}_F - \mathbf{P}_{F,e}) &= (k (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e}))^T \Sigma_{F,e}^{-1} (k (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})) \\ &= k^2 (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})^T \Sigma_{F,e}^{-1} (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e}) \\ &= k^2 \end{aligned} \quad (6.6)$$

Donc le coefficient réel k définissant le point $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ vaut :

$$k = \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M \quad (6.7)$$

Autrement dit, la distance de Mahalanobis $\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M$ permet de définir facilement le point d'intersection entre la surface $S_{F,e}$ de l'ellipsoïde et l'axe défini par le point \mathbf{p}_F et le centroïde de la classe $\mathbf{P}_{F,e}$. En effet l'équation précédente donne :

$$\mathbf{P}_{F,e}^*(\mathbf{p}_F) = \mathbf{P}_{F,e} + \frac{(\mathbf{p}_F - \mathbf{P}_{F,e})}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M} \quad (6.8)$$

On peut démontrer que le point ainsi défini n'est pas celui qui minimise la distance euclidienne de \mathbf{p}_F à la surface $S_{F,e}$ (celui considéré pour la distance classique à une surface) mais celui minimisant la distance de Mahalanobis de \mathbf{p}_F à $S_{F,e}$.

En effet soit $p_{F,e}^*$ le point minimisant cette distance $p_{F,e}^* = \operatorname{argmin}_{p_F \in S_{F,e}} \|\mathbf{p}_F - p_F\|_M$. Puisque la distance de Mahalanobis vérifie l'inégalité triangulaire, et comme $p_{F,e}^* \in S_{F,e}$, on peut écrire l'inéquation suivante :

$$\begin{aligned} \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M &= \|(\mathbf{p}_F - p_{F,e}^*) + (p_{F,e}^* - \mathbf{P}_{F,e})\|_M \\ &\leq \|\mathbf{p}_F - p_{F,e}^*\|_M + \|p_{F,e}^* - \mathbf{P}_{F,e}\|_M \\ &\leq \|\mathbf{p}_F - p_{F,e}^*\|_M + 1 \end{aligned} \quad (6.9)$$

Par conséquent, on obtient que :

$$\|\mathbf{p}_F - \mathbf{p}_{F,e}^*\|_M \geq \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M - 1 \quad (6.10)$$

Par ailleurs, puisque les vecteurs $(\mathbf{p}_F - \mathbf{P}_{F,e})$ et $(\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})$ sont linéairement dépendants, il en va de même pour les vecteurs $(\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F))$ et $(\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})$, et on se trouve donc dans le cas d'égalité de l'inégalité triangulaire. En d'autres termes, on sait que :

$$\begin{aligned} \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M &= \|(\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)) + (\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e})\|_M \\ &= \|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_M + \|\mathbf{P}_{F,e}^*(\mathbf{p}_F) - \mathbf{P}_{F,e}\|_M \\ &= \|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_M + 1 \end{aligned} \quad (6.11)$$

Ce qui donne donc que $\|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_M = \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M - 1$ et donc par conséquent que :

$$\|\mathbf{p}_F - \mathbf{p}_{F,e}^*\|_M \geq \|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_M \quad (6.12)$$

Autrement dit la distance de Mahalanobis du point $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ à \mathbf{p}_F est inférieure ou égale au minimum possible de cette distance sur la surface $S_{F,e}$. Or, par définition, $\mathbf{P}_{F,e}^*(\mathbf{p}_F) \in S_{F,e}$, ce qui montre donc bien que le point $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ est celui qui minimise la distance de Mahalanobis du point \mathbf{p}_F à la surface de l'ellipsoïde.

Finalement, si l'on veut obtenir une expression plus simple de la distance $\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}}$ à la surface de l'ellipsoïde, on peut remplacer $\mathbf{P}_{F,e}^*(\mathbf{p}_F)$ par son expression pour obtenir :

$$\begin{aligned} \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}} &= \|\mathbf{p}_F - \mathbf{P}_{F,e}^*(\mathbf{p}_F)\|_2 \\ &= \left\| \mathbf{p}_F - \mathbf{P}_{F,e} - \frac{(\mathbf{p}_F - \mathbf{P}_{F,e})}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M} \right\|_2 \\ &= \left\| (\mathbf{p}_F - \mathbf{P}_{F,e}) \left(1 - \frac{1}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M} \right) \right\|_2 \\ &= \left(1 - \frac{1}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M} \right) \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_2 \end{aligned} \quad (6.13)$$

Cette distance est valable pour les points situés à l'extérieur de l'ellipsoïde $\mathcal{E}_{F,e}$, puisque pour les autres, on considère que leur mesure de distance à la classe doit être nulle. Les points de l'intérieur de l'ellipsoïde vérifient $\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M \leq 1$, et donc le terme $\left(1 - \frac{1}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M} \right)$ est pour eux négatif. Comme nous voulons que $\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}} = 0$ si $\mathbf{p}_F \in \mathcal{E}_{F,e}$, l'expression finale du terme de distance à la classe devient :

$$\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}} = \max \left(1 - \frac{1}{\|\mathbf{p}_F - \mathbf{P}_{F,e}\|_M}, 0 \right) \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_2 \quad (6.14)$$

Cette mesure ne définit pas une distance à proprement parler dans l'espace des paramètres, mais une mesure de l'éloignement d'un point à un modèle de classe. Etant homogène à la distance

euclidienne, elle évite les distorsions problématiques de l'espace des paramètres, tout en favorisant sensiblement les classes très variables. L'important est que les points très distants du centroïde d'une classe le soient encore suffisamment de la surface de l'ellipsoïde, même lorsqu'ils se situent dans son axe principal. Cela assure qu'ils pourront être considérés comme plus proches d'autres centroïdes même si les classes qu'ils représentent ont une moindre variance, ce qui n'aurait pas été le cas avec la distance de Mahalanobis.

En calculant les valeurs de cette distance d'un point \mathbf{p}_F à chacun des modèles gaussiens des différentes espèces de \mathbb{E} , il est possible d'établir un classement, puisque les valeurs sont évaluées sur une même échelle et sont donc comparables entre elles. Pour chaque jeu de descripteurs F , les espèces vont pouvoir être classées par ordre croissant de distances, avec une possibilité d'un certain nombre de distances nulles. En effet sur un critère donné, une feuille peut parfaitement correspondre à plusieurs espèces à la fois, étant donné la plage de variations possibles.

Ces classements, indexés par les différentes mesures de distance, ont pour vocation d'être combinés, puisque l'ensemble des critères de forme est à prendre en compte pour identifier l'espèce. On peut donc légitimement espérer que si plusieurs espèces sont "reconnues" par un critère, un autre permettra de faire la décision. Outre les descripteurs introduits dans le cadre de ce travail, tout descripteur vectoriel de forme peut être utilisé pour produire un classement d'espèces par agrégation en modèles gaussiens et calcul de distances aux modèles, ce qui rend possible la comparaison de nos méthodes avec des descriptions classiques de formes, présentée au paragraphe 6.1.5.

6.1.3 Combinaison de descripteurs par fusion de confiances

Les différents jeux de descripteurs forment des sources indépendantes d'information qui produisent toutes une classification en espèces sous la forme d'une distance de l'individu à chaque espèce considérée E_e de l'ensemble \mathbb{E} . Pour parvenir à identifier l'espèce de l'individu représenté par l'ensemble des descripteurs, il faut combiner ces résultats intermédiaires de façon à obtenir un unique classement des espèces.

Chacun de ces classements donne donc pour chaque descripteur F une mesure de la distance de l'individu représenté par son vecteur d'attributs \mathbf{p}_F aux différentes espèces, distance que l'on notera $d_F(\mathbf{p}_F, E_e)$ pour l'espèce E_e . Dans le cas d'un modèle gaussien, on aura ainsi :

$$d_F(\mathbf{p}_F, E_e) = \|\mathbf{p}_F - \mathbf{P}_{F,e}\|_{\mathcal{E}} \quad (6.15)$$

Cela peut toutefois s'étendre à d'autres représentations. Par exemple pour les chaînes décrivant le bord des feuilles, on pourra ainsi utiliser la distance d'édition de la chaîne \mathcal{P} représentant l'individu, au mot moyen de l'espèce (Paragraphe 5.2.2) pour définir une mesure de distance spécifique :

$$d_{MotBord}(\mathbf{p}_{MotBord}, E_e) = d_{edit}^*(\mathcal{P}, \mathcal{P}_e) \quad (6.16)$$

Plutôt que de manipuler des distances dont les valeurs s'étendent potentiellement entre 0 et l'infini, nous préférons travailler dans un cadre plus stable avec des valeurs duales de confiance c_F qui vont rester entre 0 et 1. Ces mesures sont directement dérivées des distances d_F par une exponentielle négative. Néanmoins de façon à obtenir des confiances significatives, nous utilisons un paramètre de normalisation des distances δ_F , différent pour chaque jeu de descripteurs, qui permet d'obtenir des grandeurs comparables en vue de leur combinaison :

$$c_F(\mathbf{p}_F, E_e) = e^{-\frac{1}{2} \left(\frac{d_F(\mathbf{p}_F, E_e)}{\delta_F} \right)^2} \quad (6.17)$$

Dans les faits, ces différents facteurs de normalisation sont appris automatiquement, via l'apprentissage réalisé pour chacun des jeux de descripteurs. Ils correspondent à la valeur moyenne, calculée sur l'ensemble de la base d'apprentissage, de la mesure de distance d'un exemple appris au modèle représentant son espèce :

$$\delta_F = \frac{1}{|\mathbf{S}_F|} \sum_{\mathbf{s}_F} d_F(\mathbf{p}_{F,x}, E_{e(x)}) \quad (6.18)$$

Pour obtenir le classement final des espèces, nous calculons un score de confiance combiné correspondant à la similarité du vecteur de descripteurs complet $\mathbf{p} = (\mathbf{p}_F)_{F \in \mathcal{F}}$ formé par l'ensemble des différents jeux de descripteurs sélectionnés dans \mathcal{F} . Par exemple pour une feuille simple, on peut choisir de travailler avec $\mathcal{F} = \{Polygone, Lobes, Base, Apex, MotBord\}$.

Nous proposons deux méthodes voisines pour combiner les mesures de confiances. La première est une fusion **additive** de confiances et consiste à effectuer une somme des valeurs de confiance accordées à l'espèce par les différents jeux de descripteurs avant de la normaliser pour retrouver une valeur entre 0 et 1 :

$$c_{\oplus}(\mathbf{p}, E_e) = \frac{1}{|\mathcal{F}|} \sum_{F \in \mathcal{F}} c_F(\mathbf{p}_F, E_e) \quad (6.19)$$

Dans cette vision, chaque jeu de descripteurs vient placer de la confiance sur les différentes espèces de manière indépendante, et c'est simplement l'espèce qui sera le plus souvent citée avec une forte confiance qui se trouvera en tête du classement. La décision se fait en quelque sorte en comptant les critères correspondant à l'espèce, mais sans donner énormément d'importance aux critères qui n'y correspondent pas.

Au contraire, la seconde méthode proposée va plutôt procéder par élimination en réalisant une fusion **multiplicative** de confiances. Il est intéressant de noter que dans ce cas, étant donné la définition exponentielle de la confiance, cette combinaison revient au même que d'effectuer une fusion **additive** des distances (au carré) :

$$c_{\otimes}(\mathbf{p}, E_e) = \left(\prod_{F \in \mathcal{F}} c_F(\mathbf{p}_F, E_e) \right)^{\frac{1}{|\mathcal{F}|}} \quad (6.20)$$

Dans ce schéma en effet, le moindre critère ne correspondant pas à une espèce la pénalisera définitivement dans le classement final, l'espèce reconnue étant définie par son portrait en creux. C'est

alors l'espèce qui aura le moins de points de divergence avec l'individu considéré qui apparaîtra à la première place du classement.

Les deux approches peuvent se justifier du point de vue de l'identification des espèces. La voie classique des clés de détermination qui pousse à exclure les espèces qui diffèrent sur un critère particulier, par le choix d'embranchements dans une arborescence menant à la décision, s'apparenterait plutôt à la seconde approche. Mais pour l'observateur expérimenté, l'identification, de façon plus ou moins consciente, s'apparente plus à la convergence d'un faisceau d'indices sans que l'un occulte les autres de façon définitive, vision plus proche de notre première méthode de combinaison. La Figure 6.2 résume ce processus de combinaison des résultats des différents classifieurs associés à chaque jeu de descripteurs dans le cas des feuilles simples.

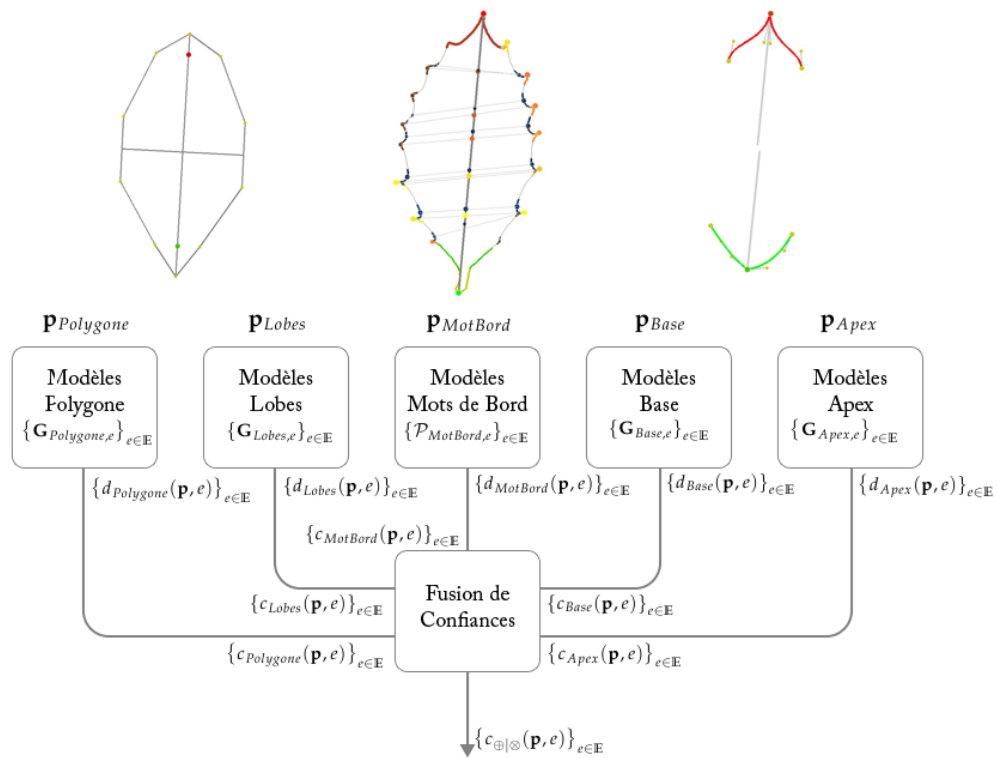


FIGURE 6.2 – Fusion de confiances issues des différents descripteurs pour les feuilles simples

Au final, en triant les mesures de confiances obtenues par ordre décroissant, on affecte à chacune des espèces E_e de \mathbb{E} un rang entre 1 et $|\mathbb{E}|$, noté $\tau(\mathbf{p}, E_e)$. En particulier l'espèce reconnue E_e^* est celle qui vérifie $\tau(\mathbf{p}, E_e^*) = 1$, mais on peut voir le résultat de la classification comme un classement plus complet en listant les espèces par rang croissant. On considérera toutefois que ces rangs peuvent être identiques en cas d'égalité de confiances pour plusieurs espèces.

Des approches plus fines pour combiner les résultats obtenus via les différents critères auraient pu être employées, notamment l'utilisation de la théorie des fonctions de croyance. Cette théorie peut en effet être vue comme une extension de la théorie probabiliste [Dem68, Sha76], permettant notamment

de mieux appréhender l'incertain, en quantifiant explicitement l'absence d'information.

Les différentes confiances pourraient ainsi permettre d'affecter des masses de croyance (*basic belief assignments*) aux espèces. Le principal problème vient du fait que ces masses s'exprime dans l'espace $\mathcal{P}(\mathbb{E})$ des **parties** de l'ensemble des espèces, dont la cardinalité est tout bonnement $2^{|\mathbb{E}|}$. Pour un nombre d'espèces de l'ordre de la centaine, la taille du cadre de détermination devient gigantesque, et les besoins en termes de mémoire et de temps de calcul deviennent irréalisables. C'est cette explosion combinatoire qui nous a poussé à renoncer à cette approche pourtant séduisante.

6.1.4 Contexte d'évaluation

Afin de mesurer la pertinence de notre description des formes foliaires, et valider les choix effectués dans ce travail, nous avons évalué les performances de classification atteintes en utilisant l'approche décrite précédemment sur une base conséquente d'exemples. Cette base est constituée à partir des images de la base **Pl@ntLeaves II** [GBJ⁺12] contenant initialement 8422 images de feuilles de 126 espèces d'arbres. Ces images sont toutes étiquetées par l'espèce qu'elles représentent, et sont réparties en trois catégories :

- ▶ **Scan** : feuilles scannées, sur un fond blanc uniforme, généralement orientées avec l'apex vers le haut, présentant une surface plane et mate.
- ▶ **Pseudoscan** : feuilles photographiées sur un fond uniforme, pas toujours blanc, souvent orientées correctement, mais pouvant présenter ombres (propres et portées) et reflets.
- ▶ **Photograph** : feuilles photographiées, généralement en environnement naturel, avec très peu de stabilité sur l'orientation, la localisation, la taille voire le nombre des feuilles.

De façon à se retrouver dans un cadre correspondant à nos prérequis, nous avons recadré et retourné certaines images (notamment de la catégorie Photograph), éliminé d'autres trop peu exploitables (photographies de feuillage par exemple), de façon à obtenir une base de 7383 images d'espèces de feuillus, contenant une seule feuille, simple ou composée (6258 simples, 1125 composées) orientée de façon à avoir globalement sa base vers le bas de l'image. Pour interpréter les résultats, nous avons considéré de la même façon les catégories Scan et Pseudoscan, sur lesquelles le comportement de nos méthodes ne diffère quasiment pas. Nous ne considérons ainsi plus que deux catégories : **FondUni** (6307 exemples) et **FondNaturel** (1076 exemples).

Au niveau du nombre de classes, nous distinguons en fait deux processus de classification bien séparés pour les feuilles simples et les feuilles composées, la décision de l'appartenance d'une feuille à l'un de ces deux ensembles étant faite bien en amont de la phase de classification. Il ne s'agit, par conséquent, pas d'un problème de classification à 126 classes, d'autant plus que certaines espèces présentes dans la base ne correspondent pas à notre contexte d'arbres et arbustes, que ce soit parce qu'il s'agit de plantes herbacées, d'espèces ne faisant pas parties des Angiospermes (conifères ou ginkgoacées), ou dont les formes ne correspondent à aucune espèce native européenne. Au final

nous distinguons donc dans la base **Pl@ntLeaves II** deux cadres de détermination distincts pour la classification :

- ▶ L'ensemble des espèces à feuilles simples : **86** classes
- ▶ L'ensemble des espèces à feuilles composées : **22** classes

De manière générale, l'évaluation de la classification se fait en validation croisée, en divisant la base d'exemples \mathbf{S} en deux sous-ensembles, un ensemble \mathbf{S}_A pour l'apprentissage et l'autre \mathbf{S}_V pour la validation (avec un rapport deux tiers, un tiers). Dans notre contexte, seuls les exemples du premier ensemble sont pris en compte pour l'estimation des modèles d'espèces, et si tous sont ensuite classifiés, les taux de classification pertinents ne sont que ceux évalués sur le second ensemble.

La mesure de performance la plus évidente est simplement le taux de bonne classification, à savoir le pourcentage d'exemples de la base \mathbf{S}_V classifiés correctement. Ce taux peut être étendu en considérant plusieurs réponses, et pas seulement la première. On pourra ainsi mesurer le pourcentage d'exemples pour lesquels l'espèce correcte figure parmi les k premières espèces du classement, le taux de bonne classification correspondant alors à une valeur de k égale à 1. On définit ainsi une mesure dépendant d'un rang maximal k de la façon suivante :

$$\mathcal{T}(\mathbf{S}_V, k) = \frac{1}{|\mathbf{S}_V|} \sum_{\mathbf{p}_x \in \mathbf{S}_V} \left(\tau(\mathbf{p}_x, E_{e(n)}) \leq k \right) \quad (6.21)$$

Une des particularités de la base considérée est que le nombre d'individus par espèce est extrêmement variable, pouvant aller de moins de 10 à environ 200. Cela pose quelques problèmes pour l'apprentissage, mais il s'agit surtout d'un biais important pour l'évaluation de la classification puisque certaines espèces vont avoir un poids considérable dans les taux de classification, alors que d'autres n'auront pour ainsi dire aucune incidence. Nous utilisons donc comme mesure un taux harmonisé de façon à donner le même poids à chacune des espèces :

$$\mathcal{T}_E(\mathbf{S}_V, k) = \frac{1}{|\mathbb{E}|} \sum_{E_e \in \mathbb{E}} \frac{1}{|\mathbf{S}_{V,e}|} \sum_{\mathbf{p}_x \in \mathbf{S}_{V,e}} \left(\tau(\mathbf{p}_x, E_e) \leq k \right) \quad (6.22)$$

Cette mesure est ainsi moins dépendante de la distribution de la base, et revient en fait au même que le taux de classification brut, dès lors que la base d'exemples est équilibrée. Dans la suite, les performances de classification seront donc données en évaluant cette dernière mesure.

Si seule les valeurs de $\mathcal{T}_E(\mathbf{S}_V, k)$ permettent de juger de la performance "réelle" de la méthode, il peut tout de même être intéressant de les comparer aux valeurs $\mathcal{T}_E(\mathbf{S}_A, k)$ obtenues sur les exemples utilisés pour l'apprentissage. L'écart potentiel entre les deux donne en effet une mesure de la généralisabilité des descripteurs utilisés, de leur capacité à être stables au sein d'une espèce et pouvoir ainsi apporter beaucoup d'information même avec peu d'exemples.

Dans les deux cas, on représentera les performances de classification par des courbes comme celles de la Figure 6.3, donnant pour différentes valeurs croissantes de k la valeur de $\mathcal{T}_E(\mathbf{S}_V, k)$, qui va

évidemment elle aussi en augmentant. Lorsque le taux $\mathcal{T}_E(\mathbf{S}_A, k)$ sur la base d'apprentissage est lui aussi représenté, il l'est par une courbe plus fine et plus claire que les valeurs sur la base de validation, mais toujours d'une même teinte pour une même méthode, permettant ainsi de faire des comparaisons. La courbe idéale est évidemment constante à 1, mais de façon générale, plus une courbe se rapproche du coin supérieur gauche, plus la méthode est performante.

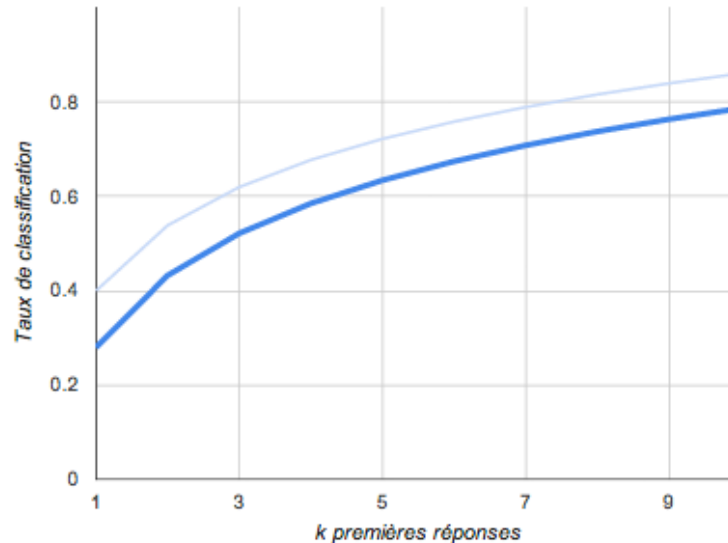


FIGURE 6.3 – Exemple de courbes de taux de classification sur la base d'apprentissage (ligne **claire**) et sur la base de validation (ligne **foncée**) pour une même méthode

6.1.5 Résultats et comparaisons

Les résultats présentés dans ce paragraphe détaillent les performances en termes de classification des méthodes employées pour la description des feuilles. Les différents descripteurs seront donc d'abord évalués séparément, et comparés à d'autres descripteurs de la littérature aux objectifs similaires, en utilisant les mêmes méthodes de classification. Dans un second temps, leur combinaison sera évaluée et mise en rapport avec d'autres méthodes de la littérature visant à une description complète de la forme des objets.

La première description de la forme d'une feuille provient de la modélisation de sa forme globale, et sa représentation par les paramètres numériques permettant la construction des modèles. Cette description se compose en fait de deux jeux de descripteurs distincts, à savoir :

- ▶ $(\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes})$ dans le cas des feuilles simples
- ▶ $(\mathbf{p}_{Folioles}; \mathbf{p}_{Polygone})$ dans le cas des feuilles composées

Dans chaque cas, la donnée seule de $\mathbf{p}_{Polygone}$ fournit des informations, sur la forme des éventuels lobes ou folioles, mais la description n'est vraiment complète que lorsque l'on prend aussi en compte l'agencement de ces derniers.

Pour les feuilles simples, qu'elles soient à lobe unique ou palmatilobées, la Figure 6.4 donne les résultats de classification et montre l'apport des paramètres concernant les lobes dans la description de la forme globale. Les deux graphiques représentent les résultats respectivement sur les images à fond uni et à fond naturel, bien que l'apprentissage soit réalisé indifféremment sur l'ensemble de ces images. Les images à fond uni étant largement majoritaires, le score sur la base d'apprentissage n'est pas véritablement pertinent pour les images à fond naturel.

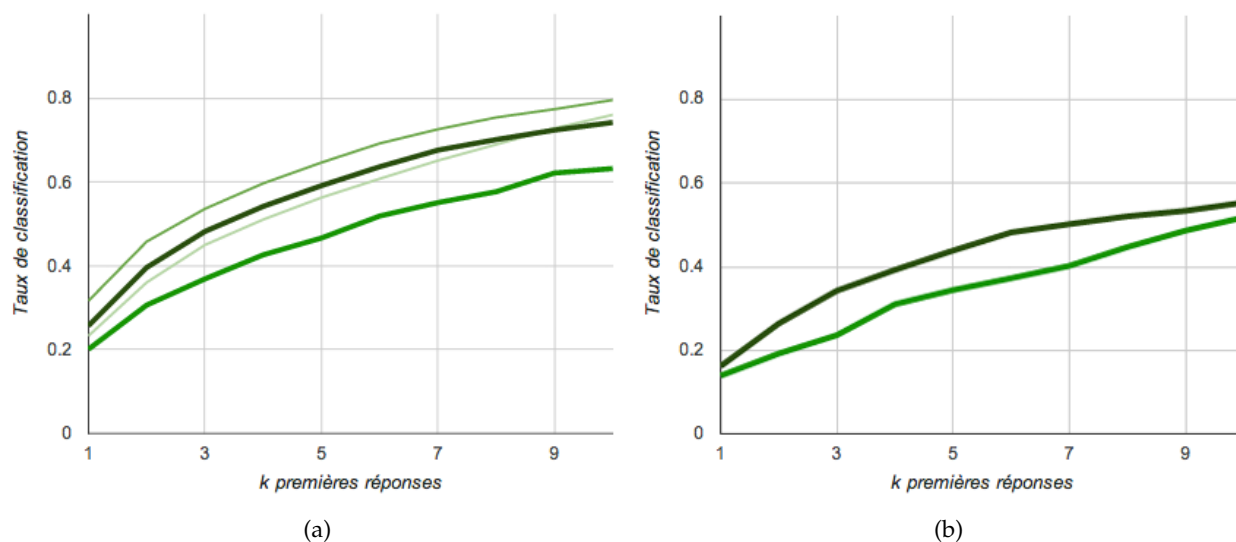


FIGURE 6.4 – Résultats de classification des feuilles simples utilisant les paramètres du modèle polygonal $\mathbf{p}_{Polygone}$, et $\mathbf{p}_{Polygone; p_{Lobes}}$: feuilles sur fond uni (a) et sur fond naturel (b)

La description de la feuille fournie par les paramètres de forme globale n'est bien évidemment pas suffisante pour identifier l'espèce à coup sûr, puisque celle-ci n'est reconnue en utilisant cette seule donnée que dans 25% des cas pour les images sur fond uni, score qui descend à 16 % dans le cas des images naturelles. Cela semble en tout cas cohérent avec les caractéristiques de forme des espèces, et connaître uniquement la forme générale laisse généralement bien trois chances sur quatre de se tromper. L'apport des caractéristiques des lobes est ici mesurable, et si l'amélioration semble de l'ordre de 3 à 5% sur la première réponse, on est plus proche en moyenne des 10% sur le reste des réponses.

Nous avons choisi de comparer cette description avec des descripteurs usuels permettant de représenter les propriétés globales de la forme d'un objet, à savoir les moments, souvent employés dans le cadre de la reconnaissance des feuilles et détaillés dans le Chapitre 2. Parmi les différentes définitions existantes, nous avons choisi deux ensembles de valeurs :

- Les **moments centrés** d'ordre 2 et 3 : 7 valeurs représentant la distribution des points d'une forme autour de son point central, calculés sur la région Ω^* correspondant à la feuille segmentée :

$$\mathbf{p}_{Moments} = \left(\mu_{20}(\Omega^*); \mu_{11}(\Omega^*); \mu_{02}(\Omega^*); \mu_{30}(\Omega^*); \mu_{21}(\Omega^*); \mu_{12}(\Omega^*); \mu_{03}(\Omega^*) \right)$$

- Les **moments de Hu** : 7 moments géométriques obtenus par combinaison des moments centrés

normalisés [Hu62], permettant une invariance à la rotation, et calculés également sur Ω^* :

$$\mathbf{p}_{HuMoments} = \left(\phi_1(\Omega^*); \phi_2(\Omega^*); \phi_3(\Omega^*); \phi_4(\Omega^*); \phi_5(\Omega^*); \phi_6(\Omega^*); \phi_7(\Omega^*) \right)$$

Ces descripteurs ont été extraits sur la même base d'exemples, en utilisant la même méthode de segmentation. Pour la classification nous appliquons la même méthode d'apprentissage de modèles gaussiens par espèce et de calcul de distance. Ce qui est donc comparé est véritablement le pouvoir discriminant des descripteurs et leur capacité à présenter une distribution régulière au sein d'une espèce. La Figure 6.5 illustre cette comparaison de descripteurs de forme globale pour les feuilles simples.

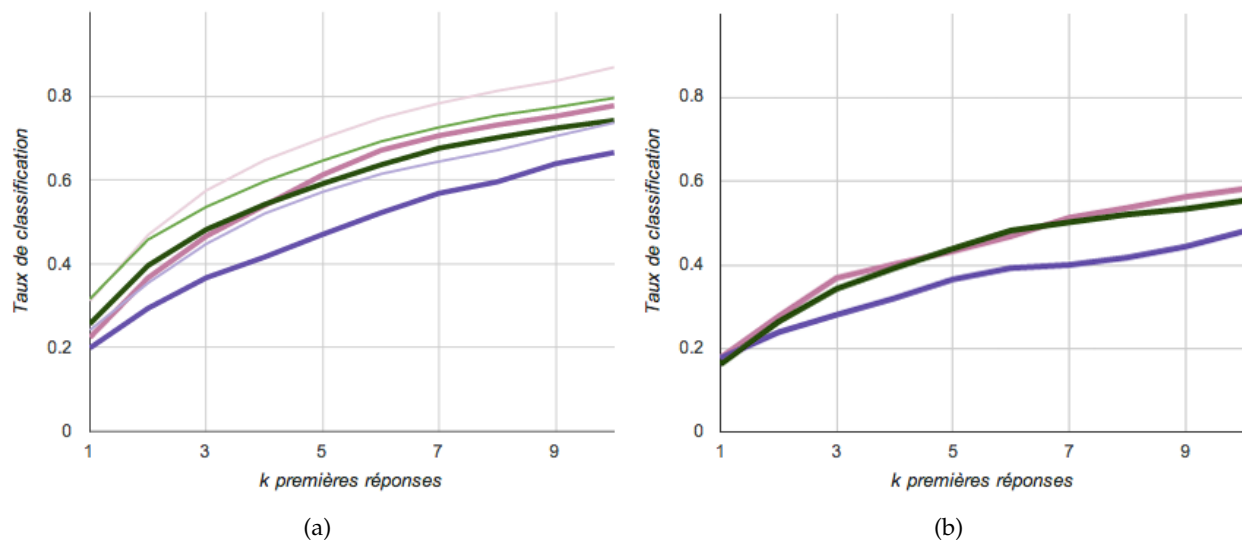


FIGURE 6.5 – Résultats de classification des feuilles simples utilisant les descripteurs de forme globale \blacksquare ($\mathbf{p}_{Polygone}$; \mathbf{p}_{Lobes}), \blacksquare ($\mathbf{p}_{Moments}$), et \blacksquare ($\mathbf{p}_{HuMoments}$) : feuilles sur fond uni (a) et sur fond naturel (b)

Il apparaît assez clairement que les moments centrés fournissent une description plus efficace que les moments de Hu, peut-être parce que, les images étant généralement toutes orientées de la même façon, l'invariance à la rotation risque plus d'apporter de la confusion entre des formes différentes que de fournir une véritable aide. Surtout, notre description se comporte globalement aussi bien que ces caractéristiques de forme bien établies, voire même un peu mieux pour les premières réponses sur fond uni.

Le point important est que nous obtenons des performances comparables avec des attributs qui offrent une interprétation directe en termes de forme, ce que ne permettent généralement pas les moments dont les valeurs restent assez difficiles à déchiffrer telles quelles. Ces performances similaires prouvent tout l'intérêt d'une description sémantique adaptée qui, pour être visiblement tout aussi représentative que des descripteurs statistiques, possède une réversibilité qui en fait un outil tout à fait intéressant pour la transparence du processus d'identification.

La comparaison avec les courbes de classification obtenues sur les exemples de la base d'appren-

tissage montre un écart nettement plus faible pour les paramètres des modèles de feuilles (5% en moyenne) que pour les moments (10% en moyenne). Cette différence suggère que notre description permet une meilleure généralisation que des descripteurs moins spécifiques. Cela indique que la connaissance experte apportée par la définition des descripteurs, dédiés explicitement à la représentation des objets considérés, permet de se focaliser sur des mesures plus stables à l'échelle de l'espèce que ne peuvent l'être des grandeurs statistiques qui ne tiennent aucun compte des spécificités de l'objet décrit.

Du point de vue de la méthode de classification utilisée, nous avons également voulu mettre en évidence la pertinence de la mesure de distance que nous avons introduite. Pour ce faire, nous avons comparé les résultats de notre distance à l'ellipsoïde représentant le modèle gaussien, avec les scores de classification obtenus en utilisant des distances plus conventionnelles, à savoir la distance euclidienne et la distance de Mahalanobis. Il nous a aussi semblé intéressant de comparer ici les performances obtenues avec celle provenant d'une recherche de plus proche voisin, où les exemples sont comparés à tous ceux de la base d'apprentissage, le rang d'une espèce correspondant à son ordre d'occurrence dans le classement des exemples. Cette comparaison de méthodes est présentée dans la Figure 6.6.

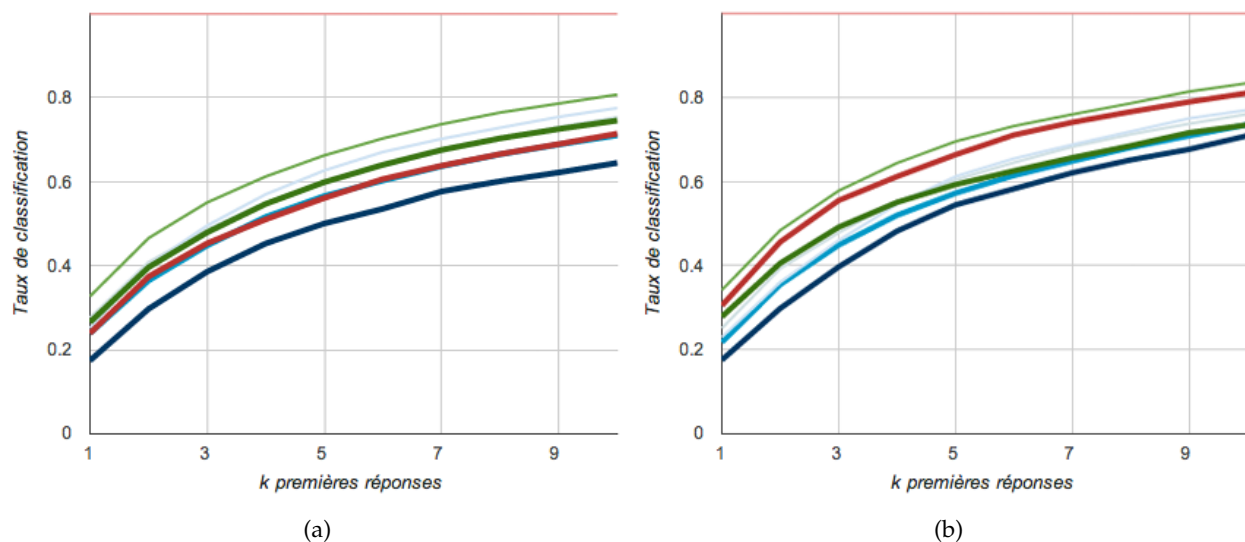


FIGURE 6.6 – Résultats de différentes classifications des feuilles simples utilisant les descripteurs de forme globale ($\mathbf{p}_{Polygone}$; \mathbf{p}_{Lobes}) : ■ distance euclidienne, ■ distance de Mahalanobis, ■ distance à la surface de l'ellipsoïde, et ■ plus proches voisins : taux de classification normalisé (a) et taux brut (b)

Ces courbes mettent bien en évidence l'importance d'avoir des distances homogènes pour toutes les classes. En effet la distance de Mahalanobis souffre beaucoup de son inconstance selon la variabilité des classes. Le fait que certaines classes, peu variables, ne puissent ainsi jamais être reconnues pénalise grandement cette approche, notamment avec une mesure normalisée du taux de classification. La variabilité des espèces apparaît tout de même comme un point important à prendre en compte, ce qui se traduit par la supériorité de la distance euclidienne à l'enveloppe des classes sur la distance

euclidienne basique.

Enfin, le peu d'écart de performance entre la comparaison et la classification par des modèles d'espèces démontre l'adéquation de la modélisation gaussienne avec la distribution effective des attributs. Au vu de la relative faiblesse de la recherche d'exemple dès lors que l'on considère le taux normalisé par le nombre d'exemples par espèce, il est fort probable que les classes se recouvrent trop (du point de vue des attributs de forme globale du moins) pour qu'une méthode de plus proche voisin soit véritablement efficace.

En ce qui concerne les attributs utilisés pour les formes locales de la feuille, notamment pour les paramètres \mathbf{p}_{Base} et \mathbf{p}_{Apex} des modèles de formes basales et apicales, il est difficile d'établir une comparaison avec des descripteurs génériques qui ne captureraient que ces formes. Nous nous contenterons donc de mesurer la performance de ces descriptions prises séparément et d'évaluer leur influence dans le processus d'identification complet.

Nous avons donc évalué les taux de classification obtenus en n'utilisant que ces deux jeux de paramètres. Il n'est bien entendu pas escompté de résultats faramineux (il serait bien difficile d'identifier une feuille uniquement par sa base et son apex) mais il peuvent receler un certain nombre d'information. En particulier, il est intéressant de voir comment ces deux facteurs *a priori* indépendants se combinent. La Figure 6.7 donne les courbes de ces résultats sur les images sur fonds unis et naturels.

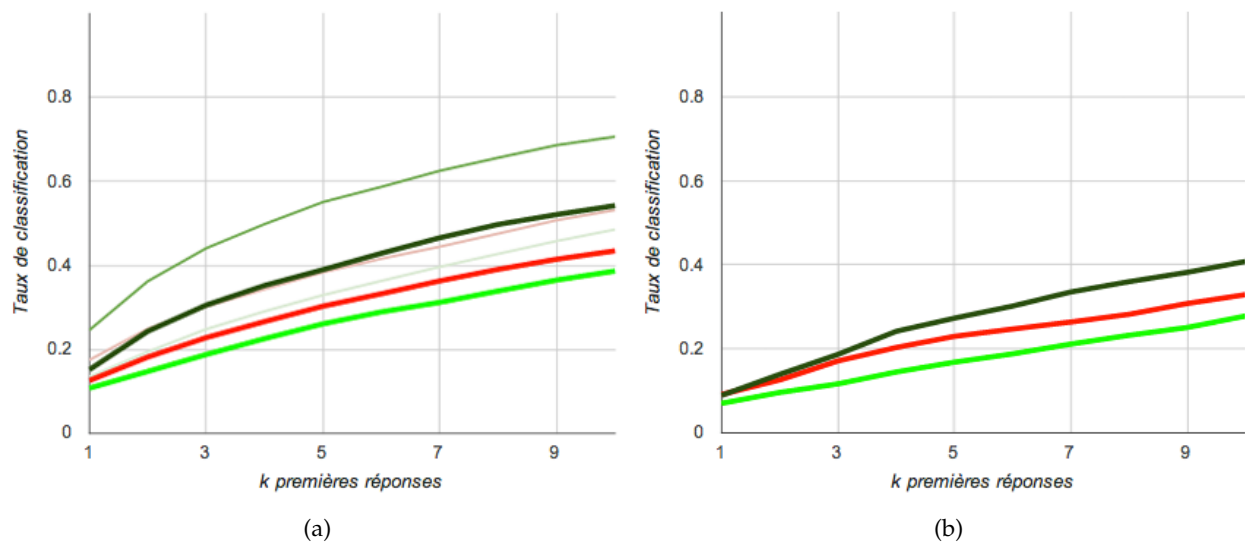


FIGURE 6.7 – Résultats de classification des feuilles simples utilisant les paramètres de formes basales et apicales \mathbf{p}_{Base} (bleu), \mathbf{p}_{Apex} (rouge), et $(\mathbf{p}_{Base}; \mathbf{p}_{Apex})$ (vert foncé) : feuilles sur fond uni (a) et sur fond naturel (b)

Sans surprise, les taux de bonne classification pour chaque descripteur sont relativement faibles, aux alentours de 15% sur les fonds unis, et peinant à atteindre les 10% sur les fonds naturels. Ils permettent pourtant de remarquer que les paramètres de l'apex sont plus performants que ceux de la base, et ce de manière assez constante. Plusieurs explications sont possibles, d'une part, les formes apicales étant plus variées, il est possible que ce critère soit tout simplement plus discriminant pour

l'identification des espèces. Il est également possible d'envisager que l'évaluation des paramètres à la base soit moins robuste que celle au sommet, puisque la présence du pétiole et la détection du point d'insertion posent généralement plus de problèmes que la caractérisation de l'apex. Ce sont là deux facteurs influençant probablement la performance.

La combinaison des deux critères ($\mathbf{p}_{Base}; \mathbf{p}_{Apex}$) permet d'améliorer de façon assez nette les résultats de classification, et, point remarquable, de plus en plus à mesure qu'on progresse dans les réponses. Cela signifie qu'il y a moins de réponses aberrantes, puisque l'espèce correcte se trouve parmi les dix premières dans plus de 54% des cas pour les images sur fond uni, 41% sur fond naturel. La réunion des deux ne permet toujours pas véritablement de faire un choix (la performance en classification pure n'est d'ailleurs pas bien plus élevée que pour les critères pris séparément) mais elle élimine déjà des réponses incorrectes qui pourraient figurer dans le haut de l'un ou l'autre des classements.

Enfin, si l'on considère l'écart apprentissage-validation (sur les fonds unis) la différence est ici flagrante, près de 15% en moyenne, ce qui remet légèrement en cause la robustesse de ces paramètres. L'une des raisons de cette moindre stabilité vient de la définition même des paramètres utilisés. En effet il s'agit à la base de paramètres de construction d'un modèle géométrique, qui sont supposés, en plus de produire une forme, être capables de représentativité dans la description, c'est-à-dire que des valeurs différentes renvoient à des formes différentes. Et dans le cas de la base et de l'apex, il est difficile de garantir ce dernier point, moins en tout cas que pour les modèles de feuilles où chaque paramètre possède une signification géométrique directement interprétable. Ici, les paramètres ayant un rôle indirect de placement de points, il se peut que leurs valeurs se trouvent moins stables, ce qui aurait pour conséquence une moins bonne capacité de généralisation.

Les descripteurs utilisés pour la forme du bord peuvent quant à eux être plus aisément comparés à des descripteurs de contour que l'on rencontre assez communément. Notre description spatialisée, dénotée par $\mathbf{p}_{MotBord}$, ainsi que sa transposition vectorielle \mathbf{p}_{Bord} qui en gomme justement la dimension spatiale, constituent des descripteurs spécialisés, voués à représenter une information bien précise sur les dents du bord, et uniquement celle-ci. En revanche, d'autres descripteurs de contour risquent bien de capturer des informations avec un spectre plus large, contenant incidemment une description plus ou moins complète de la forme globale de la feuille, et ne faisant pas la différence entre les pointes sur le bord et celles correspondant à la base ou à l'apex.

Nous nous intéressons donc dans un premier temps à l'évaluation de nos descripteurs de bord, auxquels nous ajoutons un descripteur que nous avons également utilisé pour capturer cette même caractéristique, quoique de manière moins fine. Il s'agit de moyennes $\bar{\kappa}_s$ à l'échelle du contour de valeurs de courbure calculées à des voisinages de tailles s différentes, ainsi que les écarts-types $\sigma_{\kappa,s}$ correspondants. Ce descripteur, noté $\mathbf{p}_{Courbure}$ est donc une mesure agrégative, assez naïve, du contenu du bord, qui ne fait pas vraiment la distinction entre les dents et l'apex ou les lobes. Il s'agit néanmoins d'une référence de base à laquelle il est intéressant de comparer nos descripteurs spécifiques.

$$\mathbf{p}_{Courbure} = \left(\bar{\kappa}_1; \sigma_{\kappa,1}; \bar{\kappa}_3; \sigma_{\kappa,3}; \bar{\kappa}_5; \sigma_{\kappa,5}; \bar{\kappa}_{10}; \sigma_{\kappa,10} \right)$$

Les descripteurs vectoriels $\mathbf{p}_{\text{Courbure}}$ et \mathbf{p}_{Bord} sont traités de la même façon que ceux présentés jusqu'ici à l'aide de modèles gaussiens, et permettent donc de prendre en compte la variabilité des espèces. En revanche, le descripteur structurel $\mathbf{p}_{\text{MotBord}}$ utilise des chaînes médianes comme modèles d'espèces et une distance d'édition qui ne prend pas en compte la variabilité des classes. Ces méthodes peuvent tout de même être comparées, et leurs performances de classification sont données dans la Figure 6.8.

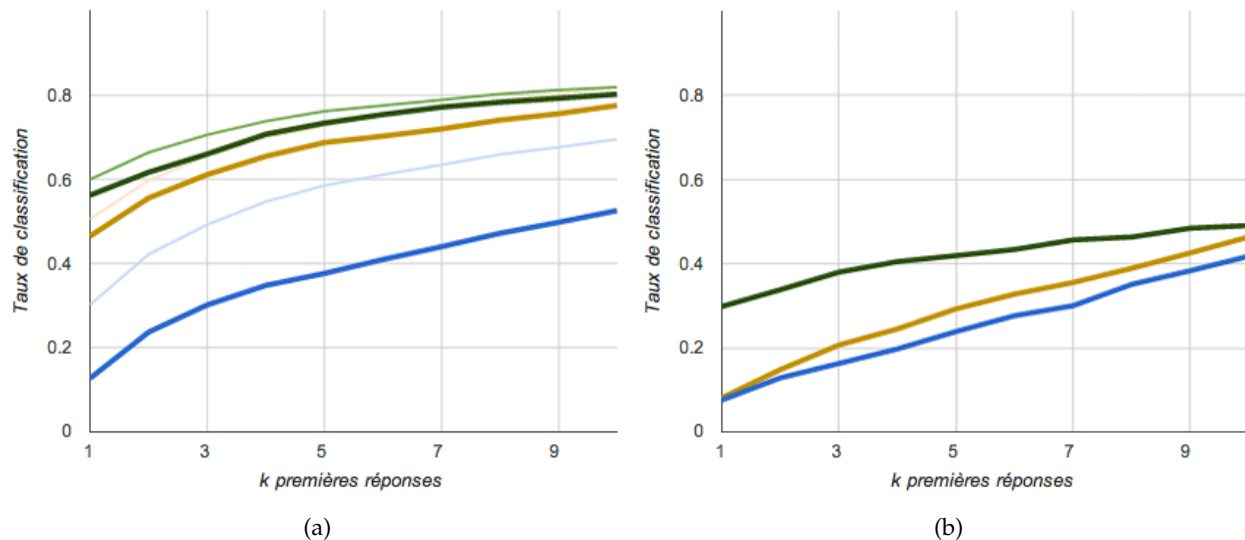


FIGURE 6.8 – Résultats de classification des feuilles simples utilisant les descripteurs de forme du bord $\mathbf{p}_{\text{MotBord}}$ (■), \mathbf{p}_{Bord} (■), et $\mathbf{p}_{\text{Courbure}}$ (■) : feuilles sur fond uni (a) et sur fond naturel (b)

Les résultats obtenus sur les fonds unis sont extrêmement intéressants et confirment si besoin était que la forme du bord de la feuille constitue un indice majeur pour l'identification de l'espèce. La représentation spatialisée du bord des feuilles semble apporter un avantage indéniable, et ce malgré l'impossibilité de prendre en compte la variabilité des classes. Avec un avantage moyen de 5% en termes de classification, conserver la répartition spatiale des dents et des creux du bord dans la description semble donc véritablement profitable.

Les valeurs exactes du score de classification, et notamment les 56% de bonne classification utilisant les seuls mots de bord, sont cependant quelque peu trop optimistes. En effet il faut pour s'en apercevoir revenir à la définition même de nos descripteurs qui sont basés sur la détection de points caractéristiques sur le contour. Dans le cas des feuilles à bord entier, pour lesquelles aucun point n'est détecté, les descripteurs, tant $\mathbf{p}_{\text{MotBord}}$ que \mathbf{p}_{Bord} seront toujours rigoureusement les mêmes. Par conséquent, les distances d'une feuille lisse à toutes les espèces à bord entier risquent bien d'être identiquement nulles, et celles-ci premières ex-æquo dans le classement. Cette particularité gonfle certes artificiellement les scores de classification, mais il n'en demeure pas moins vrai que ces espèces sont bien reconnues comme les plus proches, et que les autres critères de forme devraient permettre de faire la décision.

En dépit de ces résultats encourageants, la perte de performance lorsque l'on passe aux images naturelles est extrêmement sévère. Il y a plusieurs explications possibles à cette chute, à commencer par la sensibilité des méthodes utilisées. En effet, les images sur fond naturel sont de manière évidente celles sur lesquelles il est le plus facile de commettre des erreurs de segmentation, erreurs qui sont bien plus significatives lorsque l'on s'intéresse à la forme locale du contour que lorsque l'on évalue la forme globale de la région. Le fait d'utiliser l'information brute de courbure, sans passer par un modèle de dent garantit que le moindre artéfact sera pris en compte dans le processus.

Dans les images de feuilles sur fond naturel, il est très probable que, localement, le contour obtenu lors de la segmentation ne colle pas exactement au bord de la feuille, faute de gradient marqué, ou suite à un débordement. Dans ce cas, il apparaît une concavité ou une convexité marquée qui sera détectée. Et pour peu que ce genre de problème se reproduise sur le contour, le point aberrant aura toutes les chances d'être pris en compte, et aura une influence bien plus importante sur la description du bord que sur des méthodes de description de contour qui considèrent tous les points de la même façon, comme les moyennes de courbure.

Enfin, le fait que l'apprentissage soit réalisé indifféremment à partir d'exemples sur fond uni ou naturel vient ici rajouter des problèmes difficilement surmontables. Le contour de la région a en effet toutes les chances d'être bien plus différent entre les exemples de ces deux catégories que les autres formes qui sont estimées avec un regard plus global. La base d'apprentissage étant très majoritairement constituée d'exemples sur fond uni, il n'est sans doute pas possible de suffisamment généraliser pour correspondre aux descripteurs extraits sur les images naturelles. La solution serait de différencier les apprentissages, mais le nombre d'exemples par classe dans le cas des fonds naturels serait alors trop réduit.

Il est difficile de comparer nos descripteurs du bord des feuilles avec d'autres descripteurs génériques de contour qui ne sont pas dédiés à la représentation des mêmes objets, et qui, s'ils sont capables d'en capturer les particularités, ne le font qu'en même temps que d'autres propriétés de la forme décrite. Les taux de classification obtenus ne seraient en effet pas représentatifs de l'information apportée.

Pour identifier l'espèce, nous effectuons une combinaison des descripteurs correspondant aux différents aspects de la forme de la feuille. C'est cette fusion d'information qui va permettre de prendre une décision lorsque certains critères pourraient correspondre à plusieurs espèces, l'information apportée par les autres critères guidant le choix vers l'espèce la plus probable. Pour mettre en évidence l'intérêt de cette combinaison, nous mesurons la performance obtenue par l'ensemble des descripteurs, en appliquant les deux stratégies de fusion de confiances explicitées précédemment. Cette performance est comparée avec celle obtenue en ne considérant qu'un descripteur $\mathbf{p}_{Polygone,Lobes,Base,Apex,Bord}$ formé par la concaténation de tous les descripteurs vectoriels, et traité de la même façon, par évaluation d'un modèle gaussien par espèces. Les résultats obtenus sont donnés dans la Figure 6.9.

Ces résultats confirment l'intérêt de voir les différents critères de forme comme des indices indépendants, puisque le fait de les considérer en bloc, et donc de prendre en compte la supposée

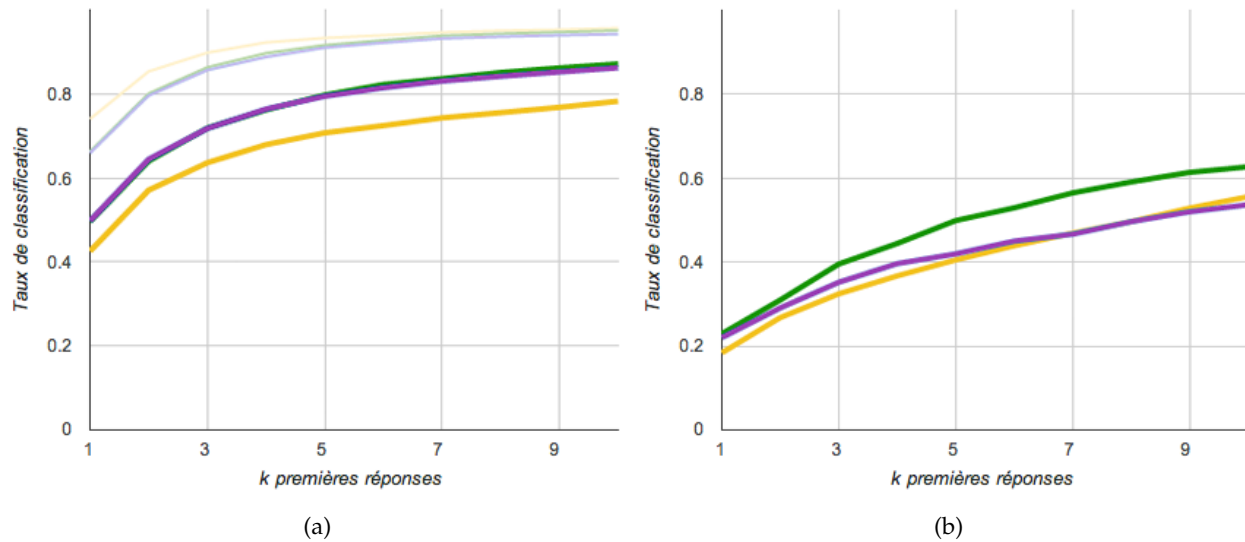


FIGURE 6.9 – Résultats de classification des feuilles simples utilisant l'ensemble des descripteurs \blacksquare vecteur concaténé $(\mathbf{p}_{Polygone, Lobes, Base, Apex, Bord})$, \blacksquare $(\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{Bord})$ avec fusion multiplicative de confiances c_{\otimes} , et \blacksquare $(\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{Bord})$ avec fusion additive de confiances c_{\oplus} : feuilles sur fond uni (a) et sur fond naturel (b)

dépendance des uns par apport aux autres, donne tout simplement de moins bons résultats. La fusion de ces descripteurs dans un second temps offre une bien meilleure performance, arrivant à prédire correctement l'espèce dans environ 50% des cas dans les images sur fond uni, et assurant surtout la présence de l'espèce correcte parmi les 10 premières réponse dans près de 90% des cas. C'est donc la performance que l'on peut attendre dans le cas d'une segmentation parfaite. Il est par ailleurs intéressant de noter que les attributs rassemblés en un seul vecteur sont excellents sur la base d'apprentissage, mais offrent visiblement une capacité de généralisation bien moins efficace que lorsqu'on les considère indépendamment.

En ce qui concerne la stratégie de combinaison, si les fusions par addition et multiplication de confiances font jeu égal sur les images sur fond uni, où l'on est en droit de supposer que les descripteurs sont corrects, vu le peu d'erreurs de segmentation possibles, il en va tout autrement sur fond naturel. C'est en effet la fusion additive de confiances qui se détache dès lors que les descripteurs ne sont plus tout à fait aussi certains, et qu'il peut y avoir du conflit entre les informations provenant des différentes sources. Cette stratégie est en quelque sorte la plus prudente, puisqu'elle n'exclut pas les espèces ne correspondant pas à un critère, mais tente au contraire d'arriver à un compromis entre toutes les information recueillies, ce qui s'avère être la meilleure solution dans les cas difficiles. C'est donc cette stratégie ainsi validée que nous retiendrons pour toute reconnaissance par la suite.

Cette stratégie s'applique également aux descripteurs des feuilles composées, qui se structurent de la même manière. Ces dernières n'ont pas été évoquées jusqu'ici puisqu'elles font l'objet d'un processus de classification à part, moins pertinent pour évaluer l'apport des descripteurs puisque limité à 22 classes. Il est tout de même intéressant d'observer la performance de nos méthodes et

l'apport de chaque descripteur dans ce cadre également. La Figure 6.10 résume donc les performances de classification sur les deux ensembles d'espèces en distinguant :

- ▶ Les descripteurs de forme globale, respectivement ($\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}$) et ($\mathbf{p}_{Folioles}; \mathbf{p}_{Polygone}$)
- ▶ Les descripteurs de formes basale et apicale ($\mathbf{p}_{Base}; \mathbf{p}_{Apex}$)
- ▶ Le descripteur spatialisé de forme du bord ($\mathbf{p}_{MotBord}$)
- ▶ L'ensemble de ces descripteurs combinés par fusion additive de confiances

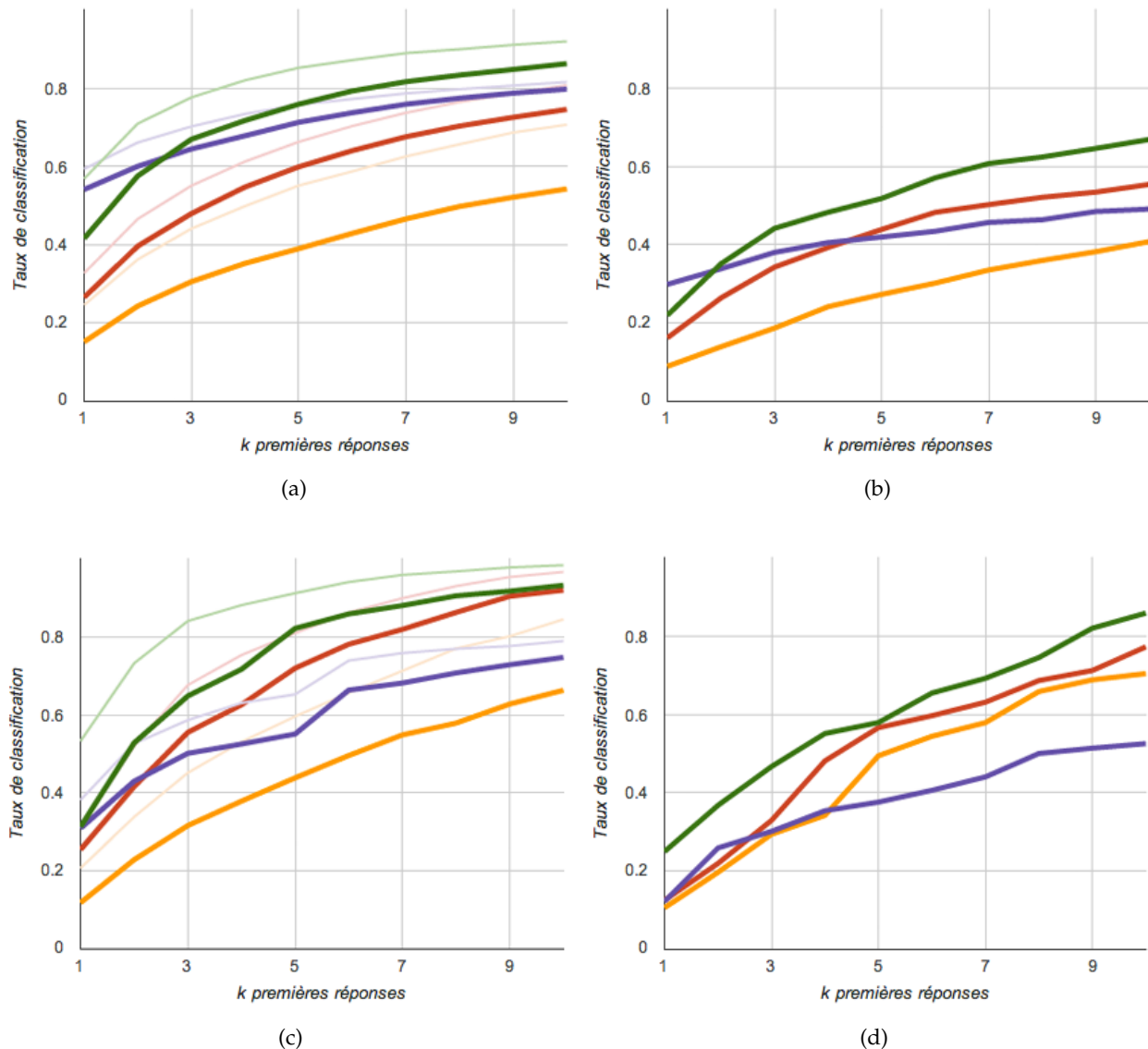


FIGURE 6.10 – Résultats de classification des feuilles simples et composées utilisant l'ensemble des descripteurs \blacksquare ($\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}$) ou ($\mathbf{p}_{Folioles}; \mathbf{p}_{Polygone}$), \blacksquare ($\mathbf{p}_{Base}; \mathbf{p}_{Apex}$), \blacksquare ($\mathbf{p}_{MotBord}$), et \blacksquare ($\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{Bord}$) ou ($\mathbf{p}_{Folioles}; \mathbf{p}_{Polygone}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{Bord}$) : feuilles simples sur fond uni (a), feuilles simples sur fond naturel (b), feuilles composées sur fond uni (c) et feuilles composées sur fond naturel (d)

La performance sur les espèces à feuilles composées semble globalement meilleure, puisque, si les 31 % de reconnaissance sur fond uni restent en dessous des 42% atteints pour les feuilles simples, les 82% de présence de l'espèce correcte parmi les 5 premières réponses offrent une garantie plus fiable que les 75% des feuilles simples. Cela est malgré tout avec quatre fois moins de classes (22 contre 86), ce qui limite les possibilités d'erreur.

Les taux obtenus sur fond naturel sont moins intéressants puisqu'ils ne dépassent pas les 30% dans les deux cas, avec néanmoins respectivement 58% et 52% de présence de l'espèce réelle parmi les 5 meilleures réponses. Cela peut sembler faible, mais il ne faut pas oublier que la base est composée d'images très difficiles, où la segmentation constitue un défi, et que dans des cas d'utilisation réelle, il sera beaucoup plus facile d'obtenir une segmentation plus fiable, et des taux de reconnaissance se rapprochant plus de ceux obtenus sur fond uni.

Concernant l'apport des descripteurs, la modélisation de la forme globale de la feuille apparaît bien souvent comme la source d'informations la plus fiable, et comme l'essentiel contributeur de la performance globale du système. Les formes basales et apicales sont quant à elles nettement moins informatives, du fait notamment des problèmes évoqués lorsqu'il s'agit de les évaluer correctement. Enfin, le bord est une source potentiellement très riche, comme le montre sa performance sur les feuilles simples sur fond uni, mais il s'avère être le descripteur le plus sensible en cas de difficultés, notamment lorsque l'on observe l'écart important entre fond uni et fond naturel.

Enfin, nous avons souhaité comparer notre méthode complète de description avec d'autres descripteurs plus génériques de la littérature, généralement évalués sur le contour, mais conçus comme des représentations suffisamment riches de la forme pour être utilisés seuls. Certains de ces descripteurs sont notamment employés dans le cadre de la description de feuilles pour la reconnaissance d'espèces d'arbres (Section 2.3) et sont à ce titre représentatifs de l'état de l'art en la matière [BCF⁺08, KBB⁺12].

Les premiers de ces descripteurs sont les *Inner-Distance Shape Context* (IDSC), construits à partir de points échantillonnés sur le contours, et qui recensent en chaque point dans un histogramme log-polaire à 64 bins la direction et la distance (calculée en demeurant à l'intérieur du contour) de tous les autres points considérés. Il s'agit d'un descripteur dont la précision dépend énormément du nombre de points retenus, qui a aussi une influence non négligeable sur la taille et le temps de calcul. Nous avons calculé ces descripteurs avec 64 points, ce qui semble être le meilleur compromis fournissant une description extensive tout en restant suffisamment léger. Le descripteur complet est donc formé de 64 histogrammes, un pour chaque point p , comptant le nombre d'occurrences SC des autres points dans des bins définis par un angle θ et une distance d :

$$\mathbf{p}_{IDSC} = \left((SC_p(d_i, \theta_j))_{i=1..8, j=1..8} \right)_{p=1..64}$$

Les IDSC présentent malgré tout un obstacle puisqu'il faut, lorsque l'on en compare deux, trouver l'alignement optimal des points du contour qui correspond aux formes se présentant sous le même angle. Par conséquent, nous apportons deux modifications dans ce cas particulier. La première est que

lors du calcul de la distance, tous les alignements possibles sont testés, et celui donnant une distance minimale est retenu. La seconde concerne la création d'histogrammes moyens par espèce, et consiste à aligner successivement les IDSC d'exemples d'apprentissage appartenant à la même espèce, et à choisir pour chaque nouvel exemple l'alignement minimisant la somme des distances aux exemples déjà rencontrés. De cette manière, on s'assure que les contours sont tous alignés globalement de la même façon, et que le calcul d'histogrammes moyens à chaque point a un sens.

Le second descripteur utilise, tout comme notre méthode, la représentation du contour par sa transformation dans l'espace courbure-échelle, mais se contente de construire un histogramme des valeurs de courbure à différentes échelles (*histogram of curvature over scale*, HoCS). Il s'agit donc d'une méthode agrégative à la manière de nos paramètres de contour, à la différence qu'elle ne cherche pas à interpréter l'information de courbure et se contente de la restituer de façon brute, en prenant donc moins de risques de faire de fausses hypothèses. Ce descripteur renferme des informations sur le contour, mais puisqu'il prend en compte des échelles assez élevées, il capture aussi des propriétés de la forme globale, et peut ainsi être utilisé comme un descripteur complet de la forme de la feuille. Le descripteur est donc formé par les valeurs H de l'histogramme dont les bins sont définis par la courbure κ et l'échelle s :

$$\mathbf{p}_{HoCS} = (H(\kappa_i, s_j))_{i=1..25, j=1..25}$$

Enfin, nous avons choisi d'évaluer des descripteurs locaux de points d'intérêts, en l'occurrence en extrayant des descripteurs SURF sur les images binaires obtenues suite à la segmentation. Le but étant d'obtenir un descripteur de forme, ces images segmentées semblaient plus appropriées que les images brutes où des détails de texture à l'intérieur de la feuille, ou des points non-pertinents à l'extérieur, auraient pu être détectés. Ces descripteurs de points d'intérêts permettent de construire un descripteur unique par image sous forme de *bag of words* (BoW), suite à l'apprentissage d'un vocabulaire de mots visuels. L'image est alors représentée par un histogramme dont la taille est celle du vocabulaire (fixée empiriquement à 1000) recensant le nombre d'occurrences W de chaque mot visuel w parmi les points détectés.

$$\mathbf{p}_{BoW} = (W_w)_{w=1..1000}$$

Ces descripteurs sont généralement utilisés dans un contexte de recherche d'images et sont donc faits pour être comparés individuellement. Cependant, pour garder le même esprit de classification nous avons choisi de leur appliquer le même traitement de création de modèles moyens d'espèces. Seulement, les données étant structurellement différentes, nous n'utilisons pas la même distance pour comparer un exemple à l'histogramme moyen d'une espèce, mais une distance euclidienne, qui nous a donné de meilleurs résultats que la distance χ^2 pourtant utilisée par les concepteurs de ces descripteurs.

En les adaptant de la sorte, il est possible de comparer ces descripteurs permettant de décrire la forme d'une feuille dans son intégralité avec la combinaison de nos différents descripteurs spécifiques de forme, utilisant la méthode de fusion additive de confiances introduite précédemment. Les résultats

de cette comparaison sont donnés dans la Figure 6.11.

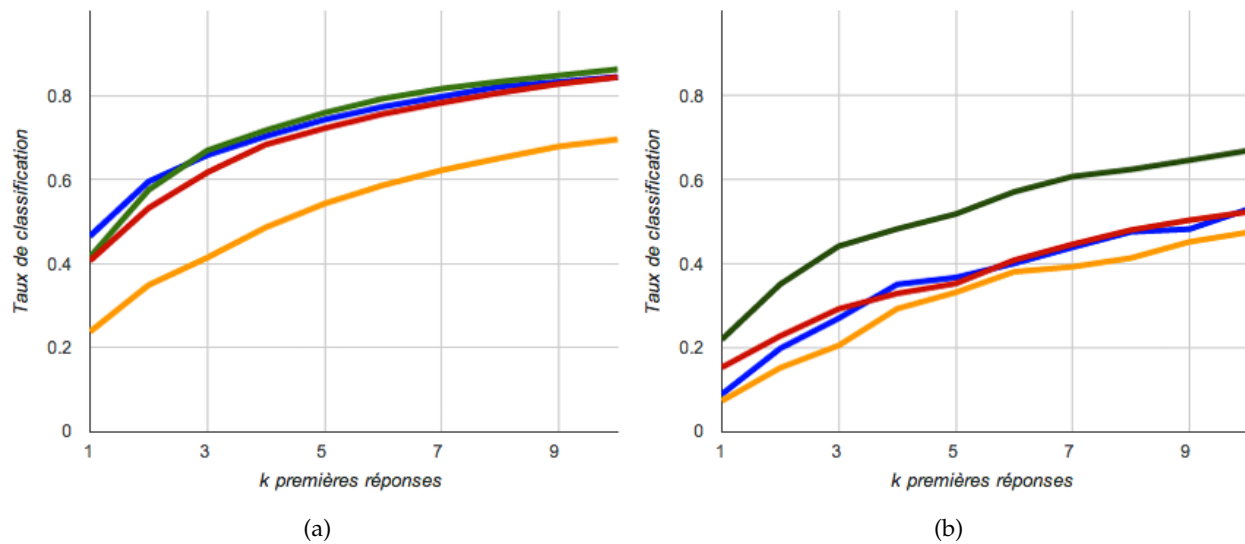


FIGURE 6.11 – Résultats de classification des feuilles simples utilisant les descripteurs de forme complets \blacksquare ($p_{Polygone}$; p_{Lobes} ; p_{Base} ; p_{Apex} ; $p_{MotBord}$), \blacksquare (p_{IDSC}), \blacksquare (p_{HoCS}), et \blacksquare (p_{BoW}) : feuilles sur fond uni (a) et sur fond naturel (b)

Le principal enseignement de cette comparaison est que nos méthodes de description se situent, dans le cadre d'une classification en espèces, au même niveau que les descripteurs les plus performants employés dans le cadre de la reconnaissance de végétaux avec des images de feuilles sur fond blanc. Il s'agit là d'une validation importante de notre approche à la fois inspirée par la botanique, fortement dédiée à l'objet d'étude, et d'assez haut niveau sémantique, dont la performance est comparable à des descripteurs de forme très complets, tout en fournissant une source d'information supplémentaire par l'interprétation directe qui peut en être faite.

Dans un contexte de recherche d'images, il est fort possible que ces descripteurs génériques s'avèrent supérieurs, par leur capacité à retranscrire fidèlement toutes les propriétés d'une forme sans se risquer à des interprétations, mais ce n'est pas un cadre envisageable. En effet l'espace nécessaire pour embarquer une base complète de descripteurs, par ailleurs assez volumineux, et les temps de calcul impliqués par les comparaisons avec un nombre conséquent d'exemples interdiraient tout simplement le traitement local sur les appareils peu adaptés à ce genre d'exercice que sont les smartphones.

Nous avons aussi mis en évidence cette difficulté en termes de temps de calcul, en comparant les temps nécessaires à la classification d'un individu à l'aide des descripteurs les plus performants. Ces tests ont été réalisés sur un processeur Intel Core i7 de 2.7 GHz, avec 4Go de RAM. Les résultats sont présentés dans la Table 6.1.

En prenant en compte le facteur 10 qui peut généralement être mesuré entre les temps d'exécution sur ordinateur et sur smartphone, il apparaît que le matching ne peut être une solution satisfaisante, en particulier pour les *shape contexts* qui donnent pourtant les meilleurs résultats. Concernant notre

Méthode	Temps de classification
(\mathbf{p}_{IDSC}) (gaussienne)	15.53ms
(\mathbf{p}_{HoCS}) (gaussienne)	0.27ms
($\mathbf{p}_{MotBord}$) (médiane)	8.71ms
($\mathbf{p}_{Polygone}; \mathbf{p}_{Lobes}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{MotBord}$) (gaussienne)	9.38ms
(\mathbf{p}_{IDSC}) (matching)	653.61ms
(\mathbf{p}_{HoCS}) (matching)	111.42ms

TABLE 6.1 – Comparaison des temps de classification d’un individu obtenus avec les différents descripteurs de forme

méthode, l’essentiel du temps de classification provient du calcul des distances d’édition nécessaires à la classification de $\mathbf{p}_{MotBord}$, le reste des calculs (distances aux modèles gaussiens et fusion) se situant sous la milliseconde. Le temps total est alors encore inférieur à celui nécessaire à la classification de \mathbf{p}_{IDSC} , ce qui place notre approche comme un compromis intéressant entre performance et légèreté.

Les résultats de classification obtenus sur les images sur fond naturel sont eux aussi particulièrement intéressants. Dans ce cadre où la segmentation est nettement moins fiable, et où les formes risquent de présenter des excroissances et des reliefs erronés, une description qui retranscrit toutes les propriétés de la forme avec une exacte fidélité risque bien de se voir pénalisée. Au contraire, notre description, par la somme de connaissance experte qu’elle contient dans sa définition, fait en quelque sorte une sélection des aspects pertinents pour l’identification. C’est sans doute cette capacité à gommer certains défauts de la forme, à effectuer une généralisation dans la description (par l’utilisation de modèles, par certaines contraintes de symétrie) qui permet d’obtenir une meilleure robustesse en dépit des problèmes de segmentation. L’apprentissage étant réalisé sur l’ensemble des images, cela souligne surtout la capacité de nos descripteurs dédiés à demeurer stable sur différents types d’images, ce qui constitue un réel avantage étant donné le contexte d’acquisition peu contrôlé que suppose son utilisation dans une application mobile.

6.2 Etiquetage par forme et sous-classification

Le fait de disposer de descripteurs représentant explicitement les critères indépendants de forme utilisés en botanique pour décrire les espèces permet d’envisager une transcription directe en concepts botaniques. Une telle interprétation sémantique des valeurs extraites sur l’image au cours du traitement constituerait un moyen attrayant de faire le lien entre le résultat de l’analyse fourni par nos algorithmes et les données théoriques sur les espèces végétales disponibles dans les flores. Il s’agirait ainsi pour l’utilisateur d’un outil d’explication du processus d’identification et de transmission d’un vocabulaire par ailleurs peu accessible, mais on peut également s’interroger sur l’intérêt potentiel à introduire ce niveau sémantique dans le processus de classification, comme une étape intermédiaire permettant de mieux représenter l’information tirée de l’image.

6.2.1 Apprentissage de classifieurs de formes foliaires

Le premier objectif pour parvenir à une interprétation sémantique des descripteurs de formes extraits de l'image consiste à mettre en place un processus de classification où les classes ne seront non pas les espèces mais les formes possibles, autrement dit les termes utilisés pour décrire les formes dans la terminologie botanique. Cela suppose qu'à toute feuille il est possible d'associer un terme par critère de forme retenu, qui sont dans notre cas au nombre de quatre : la forme globale, la forme à la base, la forme à l'apex et la forme du bord. Pour chaque critère $K \in \mathcal{K} = \{Globale; Bord; Base; Apex\}$, l'idée va donc être d'associer à chaque exemple un terme parmi l'ensemble des formes possibles \mathbb{F}_K , dont le nombre se situe généralement autour de 10. Les classes ainsi considérées sont les suivantes :

$$\mathbb{F}_{Globale} = \{ \text{Lineaire; Lanceolee; Oblongue; Ovale; Elliptique; Obovale; Suborbiculaire; Triangulaire; Orbiculaire; Cordee} \}$$

$$\mathbb{F}_{Bord} = \{ \text{Entiere; Denticulee; Ondulee; Crenelee; Dentee en scie; Dentee; Doublement dentee; Sinuee; Epineuse; Anguleuse; Lobee; Pennatifide} \}$$

$$\mathbb{F}_{Base} = \{ \text{Echancree; En coeur; Auriculee; Tronquee; Arrondie; En coin; Attenuée} \}$$

$$\mathbb{F}_{Apex} = \{ \text{Echancree; Tronquee; Arrondie; Obtuse; Aigue; Mucronee; Acuminee; Epineuse} \}$$

Malheureusement, cette hypothèse de forme unique est pour le moins abusive dans bon nombre de cas. Comme on l'a vu dans le Chapitre 1, le vocabulaire utilisé dans les ouvrages de flores renvoie à des formes canoniques, idéales, qui jalonnent l'univers des formes possibles bien plus qu'elles ne le partitionnent en classes bien délimitées. Il sera donc très fréquent qu'une feuille ne puisse pas être "rangée" sous un terme mais se trouve plutôt "à cheval" entre deux. Il s'agit donc moins d'un problème de classification tel qu'on se le représente habituellement mais plutôt d'une tâche d'annotation par des concepts. Les méthodes que nous employons pour y parvenir ont néanmoins tout d'algorithmes de classification, mais il faudra porter une attention toute particulière à l'interprétation de leurs résultats.

Le principal obstacle à l'entraînement de classifieurs vient alors du fait qu'il est très difficile de disposer d'une base d'exemples tous étiquetés par leurs différentes formes foliaires. En effet, il existe d'une part des cas où la feuille ne rentre tout simplement pas dans une seule de ces classes. Et d'autre part un étiquetage manuel, même effectué par un expert, comporterait une dose bien trop importante de subjectivité pour être utilisé comme vérité. La seule information objective dont nous disposons est en fait l'espèce de la feuille considérée.

Par conséquent, la solution la moins hasardeuse consiste à étiqueter les feuilles de façon aussi objective que possible en utilisant les informations de formes propres à leur espèce. Une espèce n'est toutefois pas associée à un unique terme par critère de forme, mais parfois à deux ou trois, ou bien à aucune. L'annexe A recense ces associations formes-espèces, en se basant sur les données théoriques tirées essentiellement de la flore de Coste [Cos06]. Nous considérons donc que chaque exemple numéroté x est associé, pour chaque critère de forme K à un sous-ensemble éventuellement

vide \mathbf{f}_K de l'ensemble \mathbb{F}_K des formes possibles pour le critère, dépendant uniquement de l'espèce $E_{\epsilon(x)}$ à laquelle appartient la feuille. La Figure 6.12 donne un exemple d'un tel étiquetage.



$$E_{\epsilon(x)} = \textit{AlnusGlutinosa}$$

$$\mathbf{f}_{Globale}(\epsilon(x)) = \{\textit{Suborbiculaire}; \textit{Obovale}\}$$

$$\mathbf{f}_{Base}(\epsilon(x)) = \{\textit{En coin}\}$$

$$\mathbf{f}_{Apex}(\epsilon(x)) = \{\textit{Tronquee}; \textit{Echancree}\}$$

$$\mathbf{f}_{Bord}(\epsilon(x)) = \{\textit{Sinuee}; \textit{Dentee}\}$$

FIGURE 6.12 – Exemple d'étiquetage d'un individu par des formes multiples à partir des données de son espèce

Une autre possibilité aurait pu être d'envisager un apprentissage non-supervisé, en cherchant à faire émerger des catégories de formes à partir des données elles-mêmes. Il s'agit d'un angle d'attaque pertinent puisque, si les botanistes ont dû employer plusieurs termes différents pour décrire les formes prises par les feuilles, c'est bien que des formes bien distinctes existaient, que l'on pourrait espérer voir apparaître comme *clusters*. La seule véritable objection à cette voie provient de l'utilisation que nous souhaitons faire de cet étiquetage par la forme, à savoir en faire un résultat intelligible pour un utilisateur, qui impose que les classes soient parfaitement en adéquation avec le vocabulaire consacré. Une forme de supervision est donc indispensable.

Les deux directions possibles qui demeurent pour l'apprentissage sont donc soit des méthodes supervisées, soit des approches semi-supervisées. La première catégorie nécessite que les données soient toutes étiquetées par une classe unique, et seuls pourront donc être utilisés des exemples pour lesquels l'ensemble \mathbf{f}_K ne contient qu'un élément. Il faut alors soit n'en conserver qu'un par exemple (reste alors à définir la méthode pour effectuer ce choix) soit ne prendre en compte que les feuilles d'espèces ne présentant qu'une forme possible. C'est cette dernière solution que nous avons privilégiée, encore une fois avec le souci d'introduire aussi peu de subjectivité que possible.

Dans ce cadre supervisé, nous avons donc utilisé des forêts aléatoires [Bre01], qui ont par ailleurs donné des résultats intéressants dans le cadre de la reconnaissance espèces (Section 6.3), pour classifier les descripteurs numériques en termes botaniques. Les forêts, contenant un nombre d'arbres fixé à 200, ont ainsi été entraînées à l'aide de données vectorielles bien étiquetées par une seule classe de forme. Le fait de ne pas prendre en compte l'ensemble des données mais seulement celles dont l'étiquette ne fait pas de doute laisse à craindre quant à la capacité de généralisation du classifieur, mais la

particularité des forêts aléatoires de prendre en compte la distribution effective des données permet de limiter ce risque. Pour certaines formes particulières en revanche, il est impossible de trouver des exemples pouvant être étiquetés par cet unique terme, et elles resteront donc en dehors du spectre couvert par le classifieur, ce qui constitue tout de même une limite importante.

La seconde catégorie de méthodes semi-supervisées permet quant à elle d'utiliser telles quelles ces données d'étiquetage imparfaites, en laissant des exemples non-étiquetés, ou en utilisant les probabilités ou tout autre type de raisonnement laissant une place à l'incertain dans l'apprentissage. Le but de ce genre d'approches est de capturer la distribution des exemples non seulement en se basant sur leurs étiquettes mais également en s'appuyant sur la répartition effective des descripteurs les représentant. Ces méthodes permettraient ainsi d'utiliser l'ensemble des données mais il est plus difficile d'assurer la parfaite adéquation des résultats avec les termes botaniques puisqu'elles mettent nécessairement en œuvre une forme de *clustering* qui risque de s'éloigner des concepts que l'on souhaite modéliser.

La méthode que nous avons choisie dans ce second cadre se base sur l'estimation de modèles gaussiens représentant les différentes formes par un algorithme s'apparentant à l'algorithme *Expectation-Maximization* (EM) [DLR77]. Plus précisément, pour chaque critère K nous cherchons à estimer les paramètres d'une somme pondérée de distributions gaussiennes (ou **GMM** pour *Gaussian mixture models*, soit plus simplement mélanges de gaussiennes) dont le nombre de composantes est fixé à $|\mathbb{F}_K|$, en maximisant la vraisemblance de l'ensemble des exemples. La supervision intervient lors de la phase E d'estimation des vraisemblances, où nous forçons les exemples à avoir une vraisemblance nulle pour les composantes gaussiennes correspondant aux formes absentes de leur étiquette \mathbf{f}_K .

L'algorithme d'estimation des modèles gaussiens $\mathbf{F}_{K,f} = (\mathbf{P}_{K,f}, \mathbf{\Sigma}_{K,f})$, ainsi que de leurs poids $w_{K,f}$ dans la distribution mixte, se déroule donc de la façon suivante. Le processus consiste en une alternance de deux phases, une étape d'estimation des vraisemblances (E), suivi de la détermination des paramètres des modèles permettant de maximiser ces vraisemblances (M). Lors de l'initialisation, on ne sait pas quelle forme parmi toutes celles possibles est la plus vraisemblable pour l'exemple représenté par $\mathbf{p}_{K,x}$ (on considèrera que $\mathbf{p}_{Globale,x} = \mathbf{p}_{Polygone,x}$) et on suppose donc leurs vraisemblances initiales $v_{K,f,x}^{(0)}$ toutes égales. En revanche, pour les formes que l'on sait inatteignables par l'espèce de l'exemple considéré, la vraisemblance doit rester nulle.

$$v_{K,f,x}^{(0)} = \begin{cases} \frac{1}{|\mathbf{f}_K(\epsilon(x))|} & \text{si } F_{K,f} \in \mathbf{f}_K(\epsilon(x)) \\ 0 & \text{sinon} \end{cases} \quad (6.23)$$

Une fois passée l'initialisation, les vraisemblances sont calculées directement à partir des modèles gaussiens obtenus jusque là. La seule contrainte est que $v_{K,f,n}^{(i)}$ reste nulle dès lors que la forme f n'est pas possible pour l'espèce $\epsilon(n)$. Ces vraisemblances devant être normalisées pour voir leur somme égale à un, le calcul à effectuer à chaque étape peut s'écrire :

$$v_{K,f,x}^{(i)} = \begin{cases} \frac{w_{K,f}^{(i-1)} e^{-\frac{1}{2}(\mathbf{p}_{K,x} - \mathbf{P}_{K,f}^{(i-1)})^T \boldsymbol{\Sigma}_{K,f}^{(i-1)} (\mathbf{p}_{K,x} - \mathbf{P}_{K,f}^{(i-1)})}}{\sum_{\phi \in \mathbf{f}_K(\epsilon(x))} w_{F,\phi}^{(i-1)} e^{-\frac{1}{2}(\mathbf{p}_{K,x} - \mathbf{P}_{F,\phi}^{(i-1)})^T \boldsymbol{\Sigma}_{F,\phi}^{(i-1)} (\mathbf{p}_{K,x} - \mathbf{P}_{F,\phi}^{(i-1)})}} & \text{si } F_{K,f} \in \mathbf{f}_K(\epsilon(x)) \\ 0 & \text{sinon} \end{cases} \quad (6.24)$$

Dans le cas gaussien, l'étape de maximisation revient simplement à mettre à jour les points moyens et matrices de covariance définissant chaque composante en utilisant les points pondérés par leurs vraisemblances. Les poids des gaussiennes sont aussi actualisés dans la même phase, comme la somme des vraisemblances des exemples contribuant à leur composante respective, en s'assurant que la somme de ces poids reste égale à 1 :

$$\begin{aligned} \mathbf{P}_{K,f}^{(i)} &= \frac{\sum_{\mathbf{S}_F} v_{K,f,x}^{(i)} \mathbf{P}_{K,x}}{\sum_{\mathbf{S}_F} v_{K,f,x}^{(i)}} \\ \boldsymbol{\Sigma}_{K,f}^{(i)} &= \frac{\sum_{\mathbf{S}_F} v_{K,f,x}^{(i)} (\mathbf{p}_{K,x} - \mathbf{P}_{K,f}^{(i)}) (\mathbf{p}_{K,x} - \mathbf{P}_{K,f}^{(i)})^T}{\sum_{\mathbf{S}_F} v_{K,f,x}^{(i)}} \\ w_{K,f}^{(i)} &= \frac{1}{|\mathbf{S}_F|} \sum_{\mathbf{S}_F} v_{K,f,x}^{(i)} \end{aligned} \quad (6.25)$$

Ces étapes sont successivement répétées jusqu'à ce qu'aucun changement significatif des modèles gaussiens ne soit plus apporté. On obtient alors un mélange de gaussiennes modélisant la distribution de l'ensemble des exemples, mais dont les composantes $\mathbf{F}_{K,f} = (\mathbf{P}_{K,f}, \boldsymbol{\Sigma}_{K,f})$ prises indépendamment vont pouvoir être utilisées pour représenter chacune des formes, du fait des contraintes du calcul des vraisemblances. Ce sont en effet ces vraisemblances nulles qui garantissent la cohérence des résultats d'un algorithme qui s'apparente à du *clustering* avec les étiquettes de supervision. Les poids $w_{K,f}$ sont quant à eux trop dépendants de la répartition des exemples pour être pris en compte par la suite, et seront donc ignorés, permettant de ne conserver que l'information sur la distribution de chaque forme dans l'espace décrit par les descripteurs vectoriels.

En termes de classification, une fois obtenus les modèles gaussiens représentant les formes, il est possible de comparer le descripteur \mathbf{p}_K d'un nouvel exemple avec chacun de ces modèles, en utilisant une mesure de distance comme par exemple celle introduite à la section précédente, et classer les formes en calculant les différentes distances $\|\mathbf{p}_K - \mathbf{P}_{K,f}\|_{\mathcal{E}}$, distances qui vont pouvoir être converties en mesures de confiance $c_K(\mathbf{p}_K, F_{K,f})$ de la même façon que les distances aux modèles d'espèces. Si le classifieur utilisé pour les formes est une forêt aléatoire entraînée par un apprentissage supervisé, le résultat en pourcentage de votes correspond directement à une mesure de confiance.

Pour mesurer la précision des différentes approches de sous-classification, nous avons évalué des matrices de confusion \mathcal{M}_K des différentes classes de formes. La valeur obtenue à chaque position $(f_1; f_2)$ de la matrice correspond au pourcentage d'exemples appartenant à la forme F_{K,f_1} associés

par le classifieur à la forme F_{K,f_2} . Le problème avec cette définition vient des exemples étiquetés avec plus d'une forme pour un critère, ce qui n'est pas rare. Dans ce cas, nous avons simplifié les choses en considérant que si la forme reconnue F_K^* fait partie des formes possibles \mathbf{f}_K de l'exemple, elle est considérée comme bien classifiée, et donc comptabilisée dans $\mathcal{M}_K(F_K^*, F_K^*)$. En revanche, si elle n'en fait pas partie, la confusion est équitablement répartie sur tous les $\mathcal{M}_K(F_{K,f}, F_K^*)$ avec $F_{K,f} \in \mathbf{f}_K$.

Les scores ont été évalués en validation croisée, en retenant deux tiers des exemples pour l'apprentissage, puis en utilisant le tiers restant dans le calcul des matrices de confusion. Dans le cas supervisé, seuls les exemples correspondant aux formes pouvant être apprises ont été conservés. De manière générale, ce sont les lignes des matrices qui représentent les réponses probables du système à une forme donnée en entrée, et permettent d'avoir une bonne idée de son comportement. Les Figures 6.13, 6.14, 6.15 et 6.16 donnent les résultats de classification en formes pour les deux approches présentées.

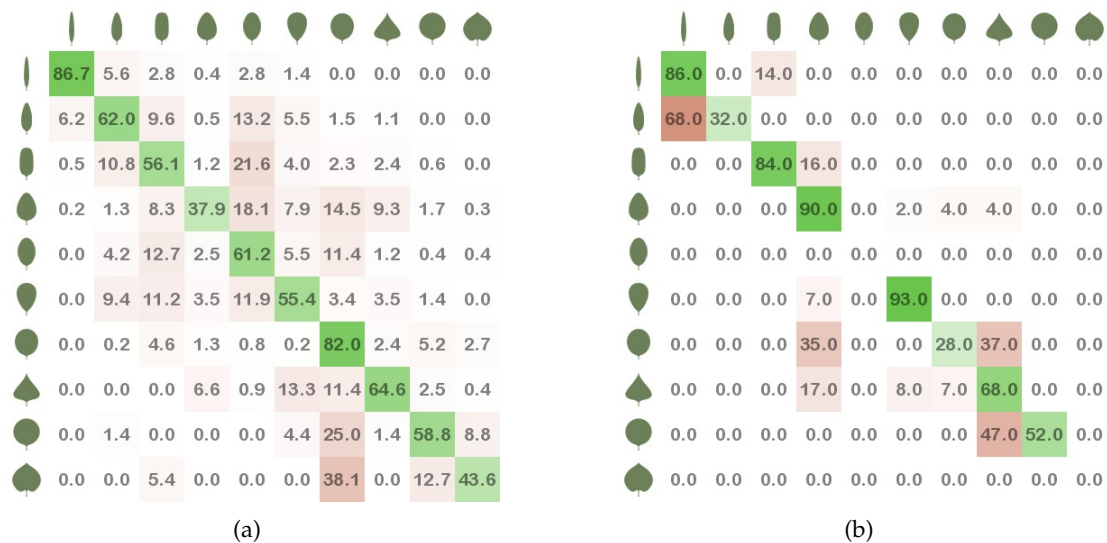


FIGURE 6.13 – Matrices de confusion des sous-classifieurs de formes globale de feuille : classification EM semi-supervisée (a) et forêts aléatoires supervisées (b)

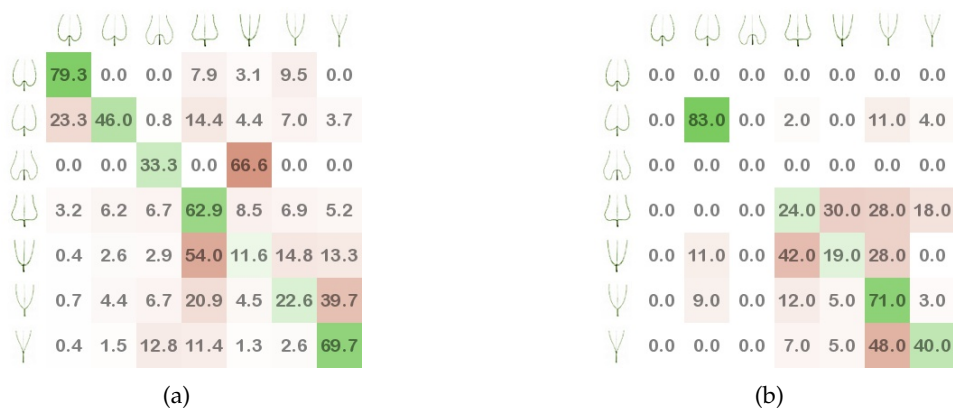


FIGURE 6.14 – Matrices de confusion des sous-classifieurs de formes de base de feuille : classification EM semi-supervisée (a) et forêts aléatoires supervisées (b)

La diagonale des matrices de confusion représente les classifications correctes, et ses valeurs devraient idéalement se rapprocher des 100%. Les valeurs diagonales proches de ce score idéal se signalent par un vert foncé. Toutes les valeurs non nulles en dehors de cet axe représentent donc les erreurs. Celles-ci peuvent dans certains cas dépasser la valeur de la diagonale sur la même ligne, signe que le classifieur, pour la classe correspondante, se trompe plus souvent qu’il ne donne une réponse correcte. La somme des scores sur une ligne est toujours de 100%. Plus le pourcentage d’erreur est important, plus le score apparaît en rouge foncé. Les colonnes présentant souvent un score d’erreur élevé correspondent à des classes attractives souvent citées comme réponses.



FIGURE 6.15 – Matrices de confusion des sous-classifieurs de formes d’apex de feuille : classification EM semi-supervisée (a) et forêts aléatoires supervisées (b)

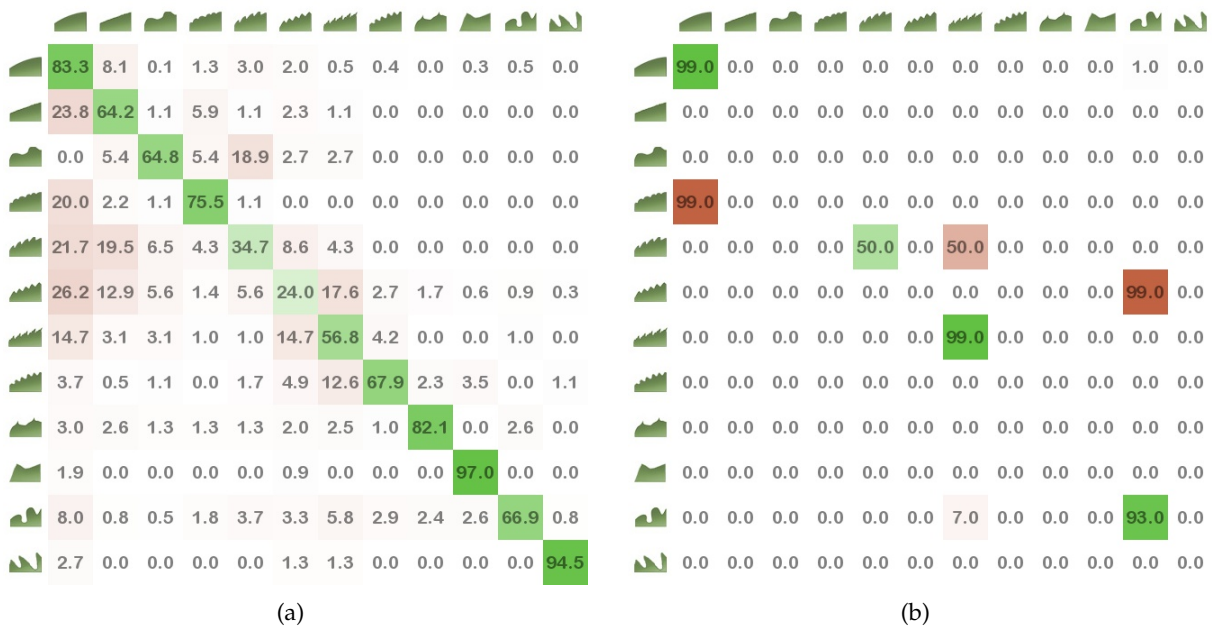


FIGURE 6.16 – Matrices de confusion des sous-classifieurs de formes de bord de feuille : classification EM semi-supervisée (a) et forêts aléatoires supervisées (b)

Les difficultés posées par l'apprentissage supervisé apparaissent ici clairement, et notamment le fait que les classes les plus représentées (forme globale ovale, base en coin, apex acuminé) ont tendance avec les forêts aléatoires à absorber les autres. Cela est dû notamment au grand déséquilibre de l'ensemble d'apprentissage, écueil difficile à contourner. Ce problème se retrouve aussi dans les scores obtenus par la méthode semi-supervisée, mais à l'inverse, puisque ces mêmes classes font partie de celles les moins bien reconnues.

En dehors de ce problème de déséquilibre, les forêts aléatoires semblent être plus performantes sur les classes où l'on dispose des données bien étiquetées, mais la perte de généralité (notamment dans le cas des formes de bord) est pour le moins préjudiciable, bien qu'inévitable avec la stratégie de non-ambiguïté choisie. D'un autre côté, la classification semi-supervisée s'avère capable de reconnaître la plupart des classes de façon satisfaisante, mais l'erreur est malgré tout bien présente. En particulier, les formes de base et d'apex semblent particulièrement difficiles à reconnaître correctement.

Il est ici important de s'interroger sur la pertinence de la description que nous utilisons dans le cadre de l'étiquetage par les formes. En effet, les descripteurs que nous utilisons comportent déjà une part de sémantique et intègrent en quelque sorte une part de connaissance experte dans leur définition. On peut alors se demander si la connaissance intégrée dans le calcul des descripteurs n'entre pas en conflit avec celle que l'on souhaite apposer aux formes pour rajouter un étiquetage sémantique. Cette question est difficile à trancher, mais il serait probablement intéressant d'étudier si des descripteurs plus génériques se comportent mieux ou moins bien étant donné qu'ils n'embarquent pas avec eux une charge sémantique.

Dans tous les cas, la sous-classification permet de fournir un étiquetage de la feuille par des termes botanique qui, s'il est largement perfectible, constitue une possibilité unique de lien entre le fonctionnement du système et des données intelligibles par un utilisateur. L'imperfection de la sous-classification n'est pas non-plus rédhibitoire pour l'utilisation dans le processus de reconnaissance, à condition de la prendre elle-même en compte dans un mécanisme de fusion adapté.

6.2.2 Modèles de fusion de sous-classifieurs

En entraînant des classifieurs permettant d'associer à des valeurs de paramètres un terme botanique, où plus exactement un classement où chaque terme botanique recensé se voit attribuer une mesure de confiance, il devient possible de convertir une information numérique en concepts sémantiques intelligibles pour l'utilisateur. Mais cet enrichissement des données pourrait tout aussi bien s'avérer utile pour la machine, qui pourrait prendre une décision sur l'appartenance aux espèces en se basant sur une information de plus haut niveau, qu'il serait en particulier possible de rattacher à des connaissances expertes reliant formes et espèces.

On peut alors voir la question de l'identification des espèces comme un problème de classification à deux étages, le premier étant une sous-classification en formes foliaires à partir de chacun des jeux de descripteurs, et le second une phase de décision revenant à fusionner les résultats de ces différents sous-

classifieurs, en faisant éventuellement intervenir une source d'information experte dans ce processus. Plusieurs stratégies sont alors envisageables pour combiner ces résultats de sous-classification, que nous tenterons d'uniformiser dans un cadre aussi général que possible.

Une première étape consiste à interpréter le résultat de chaque sous-classifieur, de façon à ne pas conserver uniquement des mesures de distance (ou de confiance) mais un résultat sous forme de probabilités plus exploitable par la suite. Cette étape que nous qualifierons de **raffinement** vise donc à obtenir pour chaque critère de forme K les différentes probabilités $P(F_{K,f} | \mathbf{p}_K)$, ce à partir non seulement des mesures de confiance $c_K(\mathbf{p}_K, F_{K,f})$ mais également d'informations externes, en lien avec la performance du sous-classifieur. L'idée est en effet d'exploiter les résultats de classification (taux de classification, matrice de confusion) pour mieux savoir ce que signifie vraiment la réponse d'un sous-classifieur. On peut déjà distinguer plusieurs approches à ce stade, donnant différents résultats à l'issue du raffinement :

- ▶ Classification à 1 réponse : en fonction des confiances et de la performance du sous-classifieur, on ne conserve qu'une seule forme F_K^* , que l'on considère comme celle correspondant à la feuille : $P(F_K^* | \mathbf{p}_K) = 1$ et $P(F_{K,f} | \mathbf{p}_K) = 0$ pour toutes les autres formes.
- ▶ Classification à deux réponses : de la même façon, cette méthode ne conserve que les deux formes (en pratique on peut considérer qu'une forme réelle ne se trouve au pire qu'à cheval entre deux formes canoniques) les plus probables F_K^* et F_K^{**} , ordonnées et quantifiées par leurs probabilités telles que $P(F_K^* | \mathbf{p}_K) + P(F_K^{**} | \mathbf{p}_K) = 1$.
- ▶ Description complète : toutes les classes de \mathbb{F}_K sont conservées, et toutes les probabilités $P(F_{K,f} | \mathbf{p}_K)$ calculées à partir des confiances et des performances de classification, avec pour seule condition que $\sum_{F_{K,f} \in \mathbb{F}_K} P(F_{K,f} | \mathbf{p}_K) = 1$.

Dans tous les cas, on peut considérer que le résultat de cette étape de raffinement remplit toutes les probabilités $P(F_{K,f} | \mathbf{p}_K)$, éventuellement avec des valeurs nulles, et peut par conséquent être visualisé comme un histogramme (plus ou moins lacunaire) associant à chaque forme possible une valeur de probabilité.

Disposer de telles probabilités, pour chacun des critères reconnus, va permettre d'envisager une résolution bayésienne du problème d'identification. Il serait également possible de voir ces probabilités comme des masses de croyance, et d'utiliser la théorie des fonctions de croyance pour prendre une décision. Le fait que le cadre de détermination fasse intervenir des formes (voire des combinaisons de formes lorsqu'on fusionne les différents sous-classifieurs) en nombre limité si l'on se borne à une ou deux réponses par classifieur va permettre de réduire l'espace de recherche et de rendre donc cette approche envisageable. Malgré des premiers résultats prometteurs, nous n'avons pas encore finalisé cette fusion crédibiliste qui se présente ici comme une solution réaliste.

La décision probabiliste va alors consister à évaluer la probabilité postérieure de chaque espèce E_e connaissant les observations, à savoir les descripteurs extraits de l'image \mathbf{p} , en se basant sur les probabilités des formes $P(F_{K,f} | \mathbf{p}_K)$ tirées de chaque sous-classifieur. On peut d'ailleurs écrire

$P(F_{K,f} | \mathbf{p}_K) = P(F_{K,f} | \mathbf{p})$ puisque les autres descripteurs de \mathbf{p} n'interviennent pas dans le calcul de cette probabilité. Pour chaque espèce la probabilité que l'on souhaite calculer est donc $P(E_e | \mathbf{p}) = P(E_e | (\mathbf{p}_K)_{K \in \mathcal{K}})$. Pour calculer cette probabilité à partir des résultats des sous-classifieurs, il va être possible de la décomposer en utilisant les partitions de l'espace que constituent les différents critères de forme. Par exemple, pour un premier critère K_1 , on peut écrire :

$$\begin{aligned}
P(E_e | \mathbf{p}) &= \frac{P(E_e, \mathbf{p})}{P(\mathbf{p})} \\
&= \frac{\sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} P(E_e, F_{K_1, f_1}, \mathbf{p})}{P(\mathbf{p})} \\
&= \sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} P(E_e, F_{K_1, f_1} | \mathbf{p}) \\
&= \sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} P(F_{K_1, f_1} | \mathbf{p}) P(E_e | F_{K_1, f_1}, \mathbf{p})
\end{aligned} \tag{6.26}$$

Les partitions formées par les critères sont considérées comme indépendantes, sous l'hypothèse que les formes prises par une feuille selon ces différents critères ne sont pas corrélées. On peut donc répéter cette opération de décomposition suivant un deuxième critère K_2 , et tenir compte de cette propriété d'indépendance qui se traduit par une simplification de toute probabilité conditionnelle impliquant deux critères différents : $P(F_{K_1, f_1} | F_{K_2, f_2}) = P(F_{K_1, f_1})$. Cela donne alors :

$$\begin{aligned}
P(E_e | \mathbf{p}) &= \sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} P(F_{K_1, f_1} | \mathbf{p}) \sum_{F_{K_2, f_2} \in \mathbb{F}_{K_2}} P(E_e, F_{K_2, f_2} | F_{K_1, f_1}, \mathbf{p}) \\
&= \sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} P(F_{K_1, f_1} | \mathbf{p}) \sum_{F_{K_2, f_2} \in \mathbb{F}_{K_2}} P(F_{K_2, f_2} | F_{K_1, f_1}, \mathbf{p}) P(E_e | F_{K_1, f_1}, F_{K_2, f_2}, \mathbf{p}) \\
&= \sum_{F_{K_1, f_1} \in \mathbb{F}_{K_1}} \sum_{F_{K_2, f_2} \in \mathbb{F}_{K_2}} P(F_{K_1, f_1} | \mathbf{p}) P(F_{K_2, f_2} | \mathbf{p}) P(E_e | F_{K_1, f_1}, F_{K_2, f_2}, \mathbf{p})
\end{aligned} \tag{6.27}$$

On peut poursuivre cette décomposition sur chacun des critères de \mathcal{K} jusqu'à avoir pris en compte tous les critères de forme considérés. Au final on obtient une expression décomposée de $P(E_e | \mathbf{p})$ qui s'écrit :

$$P(E_e | \mathbf{p}) = \sum_{f_1=1}^{|\mathbb{F}_{K_1}|} \dots \sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{K_{|\mathcal{K}|}}|} \left(\prod_{k=1}^{|\mathcal{K}|} P(F_{K_k, f_k} | \mathbf{p}) \right) P(E_e | F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}}, \mathbf{p}) \tag{6.28}$$

L'hypothèse à la base de l'utilisation des sous-classifieurs est celle que toute l'information contenue dans les observations peut être résumée par l'appartenance aux classes de formes. En d'autres termes, une fois connues les différentes formes $F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}}$ les observations \mathbf{p} n'apportent plus aucune information supplémentaire, et n'ont plus d'influence sur la probabilité d'appartenance à une espèce, ce qui peut se traduire par l'hypothèse suivante :

$$P(E_e | F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}}, \mathbf{p}) = P(E_e | F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}}) \quad (6.29)$$

En appliquant le Théorème de Bayes, il est possible d'exprimer la probabilité postérieure $P(E_e | F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})$ en fonction des vraisemblances et des probabilités antérieures. Le fait que les différents F_{K_k, f_k} soient indépendants va ici s'avérer particulièrement important, et va permettre de simplifier grandement l'expression :

$$\begin{aligned} P(E_e | F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}}) &= \frac{P(E_e)P(F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e)}{P(F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \\ &= \frac{P(E_e)P(F_{K_1, f_1} | E_e)P(F_{K_2, f_2}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e, F_{K_1, f_1})}{P(F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \\ &= \frac{P(E_e)P(F_{K_1, f_1} | E_e)P(F_{K_2, f_2}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e)}{P(F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \\ &= \frac{P(E_e)P(F_{K_1, f_1} | E_e) \dots P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e)}{P(F_{K_1, f_1}, \dots, F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \\ &= \frac{P(E_e)P(F_{K_1, f_1} | E_e) \dots P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e)}{P(F_{K_1, f_1}) \dots P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \\ &= P(E_e) \prod_{k=1}^{|\mathcal{K}|} \frac{P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \end{aligned} \quad (6.30)$$

Dans cette expression, les $P(E_e)$ et $P(F_{K_k, f_k})$ correspondent à des probabilités *a priori* de rencontrer une espèce ou une forme donnée, et doivent sans doute faire appel à une source de connaissance experte pour être déterminées avec exactitude ; il est également possible de rester plus général et de supposer tous ces évènements équiprobables. En réinjectant cette expression dans la formule de $P(E_e | \mathbf{p})$, puis en factorisant par les termes qui ne dépendent que d'un critère K_k donné, on obtient alors :

$$\begin{aligned}
P(E_e | \mathbf{p}) &= \sum_{f_1=1}^{|\mathbb{F}_1|} \dots \sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{|\mathcal{K}|}|} \left(\left(\prod_{k=1}^{|\mathcal{K}|} P(F_{K_k, f_k} | \mathbf{p}) \right) P(E_e) \prod_{k=1}^{|\mathcal{K}|} \frac{P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \right) \\
&= P(E_e) \sum_{f_1=1}^{|\mathbb{F}_1|} \dots \sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{|\mathcal{K}|}|} \left(\prod_{k=1}^{|\mathcal{K}|} \frac{P(F_{K_k, f_k} | \mathbf{p}) P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \right) \\
&= P(E_e) \sum_{f_1=1}^{|\mathbb{F}_1|} \left(\frac{P(F_{K_1, f_1} | \mathbf{p}) P(F_{K_1, f_1} | E_e)}{P(F_{K_1, f_1})} \left(\sum_{f_2=2}^{|\mathbb{F}_2|} \dots \sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{|\mathcal{K}|}|} \prod_{k=2}^{|\mathcal{K}|} \frac{P(F_{K_k, f_k} | \mathbf{p}) P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \right) \right) \\
&= P(E_e) \left(\sum_{f_1=1}^{|\mathbb{F}_1|} \frac{P(F_{K_1, f_1} | \mathbf{p}) P(F_{K_1, f_1} | E_e)}{P(F_{K_1, f_1})} \right) \left(\sum_{f_2=2}^{|\mathbb{F}_2|} \dots \sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{|\mathcal{K}|}|} \prod_{k=2}^{|\mathcal{K}|} \frac{P(F_{K_k, f_k} | \mathbf{p}) P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \right) \\
&= P(E_e) \left(\sum_{f_1=1}^{|\mathbb{F}_1|} \frac{P(F_{K_1, f_1} | \mathbf{p}) P(F_{K_1, f_1} | E_e)}{P(F_{K_1, f_1})} \right) \dots \left(\sum_{f_{|\mathcal{K}|}=1}^{|\mathbb{F}_{|\mathcal{K}|}|} \frac{P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | \mathbf{p}) P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}} | E_e)}{P(F_{K_{|\mathcal{K}|}, f_{|\mathcal{K}|}})} \right)
\end{aligned} \tag{6.31}$$

Ce qui mène à l'expression finale de la probabilité postérieure d'une espèce $P(E_e | \mathbf{p})$:

$$P(E_e | \mathbf{p}) = P(E_e) \prod_{k=1}^{|\mathcal{K}|} \left(\sum_{f_k=1}^{|\mathbb{F}_k|} \frac{P(F_{K_k, f_k} | \mathbf{p}) P(F_{K_k, f_k} | E_e)}{P(F_{K_k, f_k})} \right) \tag{6.32}$$

Cette expression fait intervenir le résultat des différents sous-classifieurs sous la forme des probabilités $P(F_{K_k, f_k} | \mathbf{p})$ de correspondre à une forme donnée au vu des attributs observés. La Figure 6.17 résume le déroulement du processus menant au calcul de ces probabilités postérieures. Cette formulation laisse de manière assez intéressante une place pour inclure de la connaissance *a priori*, que ce soit sur la présence des espèces dans $P(E_e)$ ou sur l'importance d'une forme donnée dans les $P(F_{K_k, f_k})$, les formes plus rares donnant alors par exemple un poids plus fort aux espèces auxquelles elles correspondent.

Les quantités $P(F_{K, f} | E_e)$ correspondent quant à elles aux probabilités de rencontrer une forme donnée chez une espèce. On peut considérer que c'est cette information que recensent les ouvrages de flore, et donc que ces probabilités peuvent être calculées à partir des formes théoriques des espèces $\mathbf{f}_K(e)$. Faute d'informations plus précises, la solution la plus simple est de considérer que toutes les formes ainsi mentionnées sont équiprobables au sein de l'espèce :

$$P(F_{K, f} | E_e) = \begin{cases} \frac{1}{|\mathbf{f}_K(e)|} & \text{si } F_{K, f} \in \mathbf{f}_K(e) \\ 0 & \text{sinon} \end{cases} \tag{6.33}$$

En simplifiant à l'extrême, dans le cas où l'on suppose l'équiprobabilité des espèces, et des formes de chacun des critères, et si l'on choisit la première stratégie consistant à ne conserver qu'une réponse du classifieur, pour ordonner les espèces suivant leurs probabilités postérieures, la quantité à calculer

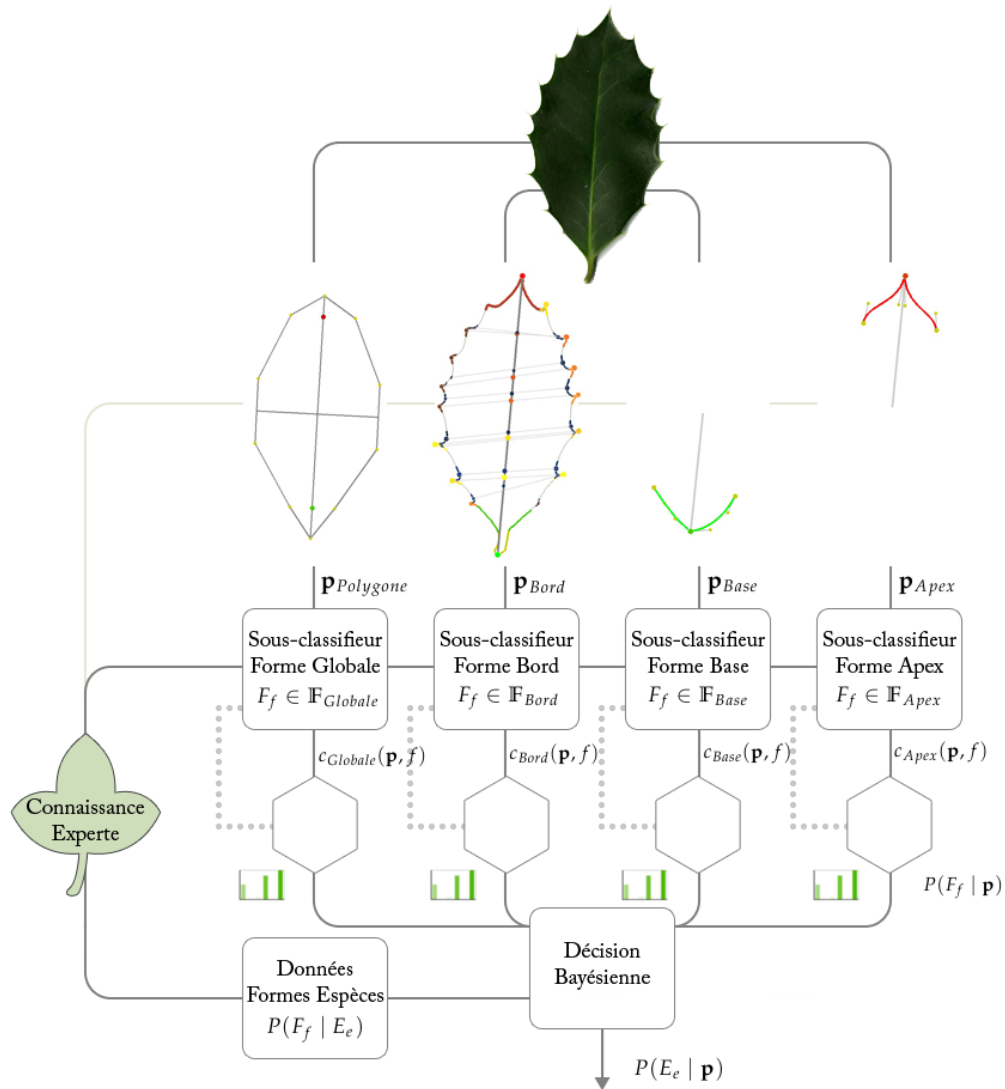


FIGURE 6.17 – Fusion de confiances issues des différents sous-classifieurs de formes pour les feuilles simples

est la suivante :

$$P(E_e | \mathbf{p}) \propto \prod_{k=1}^{|\mathcal{K}|} P(F_{K_k}^* | E_e) \quad (6.34)$$

On se trouve donc simplement dans le cadre d'une classification naïve bayésienne considérant comme attributs les réponses des différents sous-classifieurs, c'est-à-dire les classes de forme attribuées à l'exemple, éventuellement légèrement corrigées suite à l'étape de raffinement. La décision quant à l'appartenance de l'exemple considéré à une espèce se fait alors en choisissant le maximum de la probabilité $P(E_e | \mathbf{p})$. En pratique cela revient à considérer l'ensemble des espèces correspondant à l'intersection de toutes les formes reconnues (les autres se retrouvant nécessairement avec une probabilité nulle) et de les ordonner selon le degré de certitude apporté par chacune des formes.

Cette approche peut toutefois poser problème, par exemple lorsqu'aucune espèce ne présente simultanément les 4 formes reconnues par les sous-classifieurs. Suivant ce modèle de fusion, toutes les espèces se retrouveraient alors avec une probabilité postérieure nulle ce qui pourrait constituer un bon moyen de détecter les situations conflictuelles, mais reste tout de même problématique dans un contexte où une réponse est attendue. Dans cette perspective, la stratégie consistant à garder les deux premières réponses de chaque sous-classifieur, indexées par une mesure de probabilité, apparaît comme un bon compromis, permettant d'éviter les cas d'intersection nulle puisque les ensembles de formes considérés sont nécessairement plus larges.

Pour limiter le risque d'intersection nulle dans ces cas, on peut également considérer la réunion des espèces correspondant à au moins un critère de forme reconnu. C'est une approche moins dangereuse, mais qui s'éloigne alors un peu du formalisme bayésien employé jusque là. La conséquence en termes de calculs est que la quantité à évaluer pour ordonner les espèces (toujours dans un cas d'équiprobabilité *a priori*) peut être approximée de la façon suivante :

$$P(E_e | \mathbf{p}) \propto \sum_{k=1}^{|\mathcal{K}|} \left(\sum_{f_k=1}^{|\mathbb{F}_k|} P(F_{K_k, f_k} | \mathbf{p}) P(F_{K_k, f_k} | E_e) \right) \quad (6.35)$$

Enfin, la dernière stratégie conserve toutes les réponses des sous-classifieurs avec leurs probabilités. Le risque de cette approche est qu'elle ne fait pas de choix très marqués, ce qui risque de biaiser la classification, en laissant à toutes les espèces la chance de bien figurer (aucune n'aura *a priori* de probabilité nulle). Elle donne en particulier une grande importance à la phase de raffinement des probabilités, et la façon de prendre en compte la performance des classifieurs va devenir cruciale. Bien que vraisemblablement la plus fidèle à l'information en entrée, elle se présente comme la plus éloignée du cheminement de l'analyse botanique, en n'excluant jamais définitivement les formes et les espèces, ce qui peut s'avérer bien plus un risque qu'une sécurité.

Ces modèles de fusion de sous-classifieurs permettent ainsi d'utiliser les termes botaniques pour parvenir à l'identification des espèces en manipulant des concepts de haut niveau, simulant ainsi le processus d'analyse qu'emploierait un observateur humain. Ils reposent toutefois sur plusieurs

éléments dont il est difficile d'assurer la véracité, en particulier l'indépendance totale des différents critères de forme, ou encore les connaissances expertes relatives aux formes prises par les espèces qui, si elles sont officialisées par les ouvrages botaniques, n'en portent pas moins la part de d'imprécision et subjectivité inhérente au vocabulaire employé. Il s'agit néanmoins d'une solution intéressante car elle permet un lien direct entre les informations tirées des images et la somme des données collectées par les botanistes sur les espèces, un point particulièrement attrayant dans l'optique d'un processus explicatif de reconnaissance, et d'une transmission de savoirs botaniques.

6.2.3 Performances et intérêts

La pertinence de l'introduction de sous-classifieurs est à mesurer sur deux plans bien différents. La performance de la sous-classification en formes elle-même est importante, afin de savoir si l'étiquetage des descripteurs par les termes botaniques peut être fait de façon fiable. Cela intervient non seulement pour l'identification des espèces, mais aussi pour l'apposition de concepts sémantiques sur les informations extraites de l'image, en vue de les présenter à l'utilisateur. Il faut surtout mesurer l'effet de l'introduction de la sous-classification au niveau de l'identification de l'espèce, pour évaluer l'intérêt qu'il peut y avoir à passer par ce niveau intermédiaire et à tirer profit des connaissances expertes disponibles.

La classification ainsi proposée ne repose donc pas sur un apprentissage des classes qu'elle se propose de reconnaître, puisque l'information liant espèces et formes est purement théorique, et tirée de données externes. Le seul apprentissage automatique a lieu dans la phase d'entraînement des sous-classifieurs. Cela se base donc sur une hypothèse forte, à savoir que les données théoriques et les formes observées sur les individus soient bien en adéquation. Un avantage considérable de cette approche, est que l'ajout de nouvelles espèces doit pouvoir se faire manuellement, par simple définition des formes possibles de l'espèce, en reprenant les critères donnés par les flores.

Pour un maximum de généralité, nous avons d'abord évalué la performance de la classification en utilisant les sous-classifieurs semi-supervisés, qui sont capables de classer toutes les formes rencontrées et donc de permettre l'identification de l'ensemble des espèces considérées. Cette évaluation est faite sur les feuilles simples de la base Pl@ntLeaves II, en utilisant des données théoriques données par la flore de Coste [Cos06], et donc en ne conservant que les 78 espèces généralement natives qui y sont décrites.

La Figure 6.18 donne les résultats de classification en espèces utilisant la sous-classification semi-supervisées. Dans ce cas, les matrices de confusion obtenues pour les différents sous-classifieurs ne sont pas utilisées pour raffiner les réponses. Les différentes stratégies comparées sont alors les suivantes :

- Toutes les réponses des sous-classifieurs conservées, et fusion bayésienne des probabilités
- Toutes les réponses des sous-classifieurs conservées, et fusion additive des probabilités
- Les deux premières réponses conservées, et fusion bayésienne des probabilités

■ Les deux premières réponses conservées, et fusion additive des probabilités

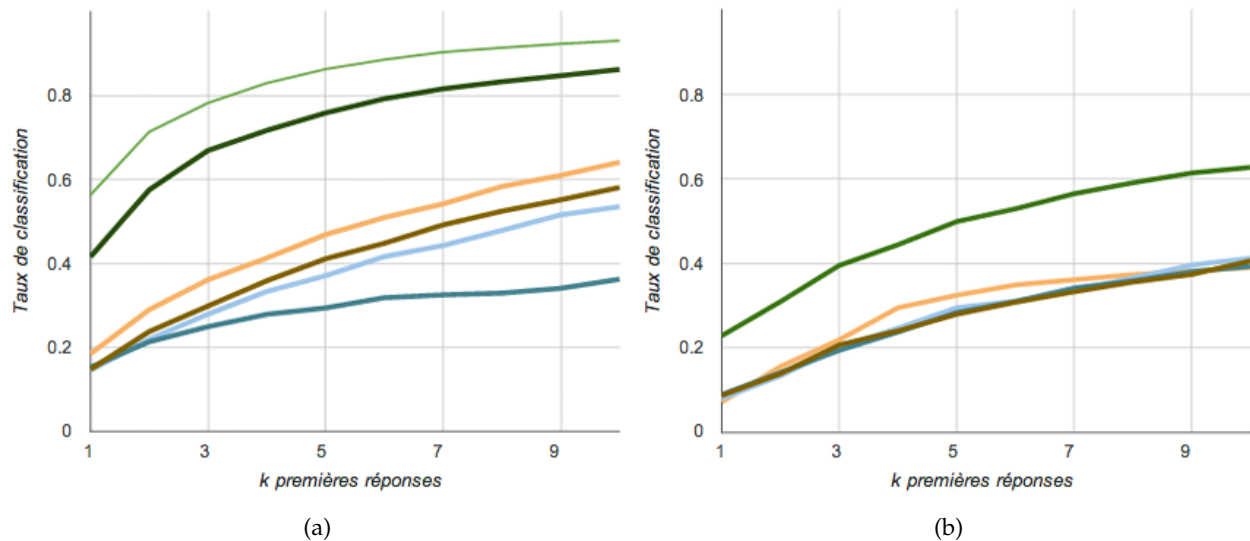


FIGURE 6.18 – Résultats de classification des feuilles simples par fusion de sous-classifieurs de forme, et comparaison avec ■ la classification automatique utilisant directement les paramètres de forme : feuilles sur fond uni (a) et sur fond naturel (b)

De façon intéressante, si la fusion probabiliste des résultats est celle qui fonctionne le mieux lorsque l'on conserve l'ensemble des résultats de sous-classification, elle semble trop risquée lorsqu'un ne garde que les premières réponses. Puisqu'elle revient à considérer une intersection des espèces possibles pour chaque forme, la moindre erreur va en effet être très pénalisante, et il est alors plus souhaitable d'avoir une approche moins restrictive. Il apparaît tout de même que la réponse complète des sous-classifieurs apporte un surplus d'information non-négligeable (dans le cas d'une évaluation fiable des descripteurs en tout cas) et qu'il convient donc de la conserver dans son intégralité.

L'écart de performance avec la classification basée sur un apprentissage automatique à partir des données brutes extraites de l'image est malgré tout criant. Cela montre les limites de l'approche, qui consiste quelque part à introduire un filtre supplémentaire entre l'image et l'espèce. Le niveau d'erreur de ce filtrage vient donc se superposer aux erreurs déjà commises au cours du processus, sans compter que le lien entre forme et espèce ne se base plus réellement sur les données d'apprentissage. Cela explique donc qu'une telle approche, sans raffinement complexe des réponses en termes de formes, échoue à atteindre la même performance qu'une méthode plus directe pour associer les données extraites de l'image et les classes.

Toutefois, dans le cadre de la sous-classification supervisée (par les forêts aléatoires), les résultats peuvent être nettement améliorés par l'introduction de fonctions de croyance dans l'étape de raffinement [LCV13], de façon à prendre en compte au mieux le comportement des sous-classifieurs. Sur un spectre certes plus réduit d'espèces, la performance atteint alors, sur les images sur fond uni, 32 % lorsque seule la première réponse pour chaque forme est prise en compte, et jusqu'à 51 %

lorsque les deux premières réponses sont conservées. Ces résultats tendent donc à prouver qu'avec des performances plus fiables de sous-classification, en l'étendant à l'ensemble des formes possibles, la performance de cette approche de fusion de critères de forme pourrait être tout à fait satisfaisante pour l'identification des espèces.

Dans l'état actuel des choses, les classifieurs de forme offrent tout de même un intérêt non négligeable dans le cadre d'un processus de reconnaissance explicatif. En effet, ils permettent une traduction de l'information extraite de l'image en concepts intelligibles, permettant une meilleure compréhension de ce sur quoi se base le système pour prendre sa décision. Il est ainsi possible d'offrir à l'utilisateur une représentation visuelle, rattachée à des formes canoniques, de ce que le système est en train de reconnaître, comme dans l'exemple donné dans la Figure 6.19.

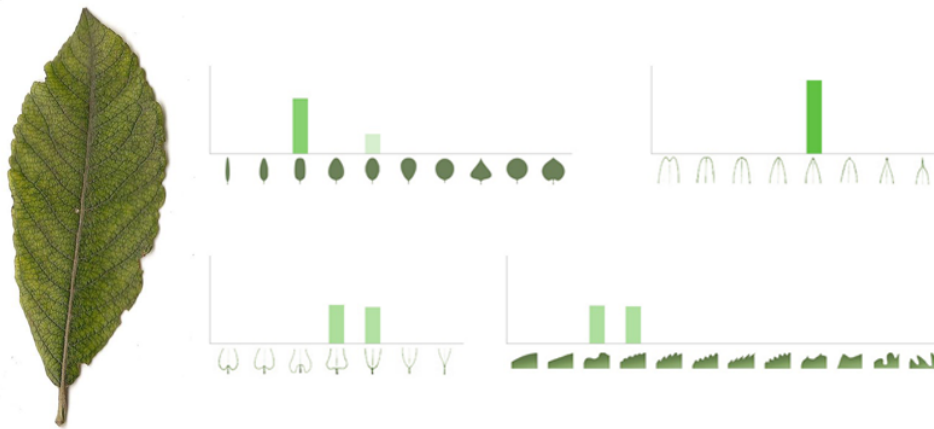


FIGURE 6.19 – Exemple d'étiquetage automatique d'un exemple par les termes de forme utilisant les résultats des sous-classifieurs

Le fait de relier les valeurs des descripteurs à des formes interprétables par l'utilisateur lui donne en particulier l'opportunité de détecter facilement où se situent les erreurs éventuelles du programme (par exemple de voir immédiatement si le système s'est trompé dans la forme du bord). Il pourra ainsi beaucoup mieux comprendre et accepter les identifications ratées qu'avec un système qui agirait plus comme une boîte noire se contentant de livrer les réponses en bout de processus.

Il serait possible d'envisager à cet égard un système plus interactif permettant à l'utilisateur de corriger le déroulement du processus à ce niveau sémantique, en modifiant les formes, ou en pointant sur les critères qui seront vraisemblablement source de confusion. Une telle extension constituerait de plus un moyen d'impliquer d'avantage l'utilisateur dans l'identification et donc de faire en sorte qu'il retire plus de l'utilisation du système de reconnaissance.

Enfin, l'étiquetage en forme a simplement pour vocation de transmettre le vocabulaire associé, en permettant une correspondance visuelle directe entre la forme analysée et les termes employés pour la décrire. Le fait d'apprendre de la sorte à l'utilisateur la signification concrète de ces mots est le

moyen le plus évident de faire le lien entre ce qu'il peut observer et les données qu'il peut par ailleurs rencontrer dans les flores. Les descriptions des espèces par les formes prendront alors plus de sens, un premier pas de franchi vers la compréhension des ouvrages botaniques de référence. La transmission de ce savoir justifie ainsi à elle seule la nécessité d'introduire une part de sémantique, si ce n'est pour l'identification, du moins dans le processus explicatif de la reconnaissance.

6.3 Classification par forêts aléatoires

Les méthodes de classification utilisées jusqu'ici constituent un moyen efficace de valider les choix faits au niveau des descripteurs via un apprentissage léger et rapide, ou de s'assurer de leur cohérence d'un point de vue sémantique, mais les mécanismes de classification à l'œuvre sont quelque part assez naïfs et peinent à rendre compte de façon optimale de la diversité des espèces. Il est donc souhaitable d'appliquer d'autres méthodes de classification plus complexes aux mêmes données de façon à en tirer le maximum d'informations.

6.3.1 Modèles de classification

Les méthodes d'apprentissage automatique les plus souvent utilisées pour la classification reposent sur l'entraînement fin d'un outil de classification spécifique, dont la définition permet généralement bien plus de versatilité que le calcul de distances dans un espace vectoriel. En effet les délimitations entre les classes peuvent alors prendre des formes bien plus complexes, quand l'utilisation de distances limite ces frontières à la définition d'hyperplans souvent trop rigides.

Parmi les modèles les plus communément employés, y compris dans le cadre de l'identification automatique de végétaux, on peut citer les réseaux de neurones artificiels (*artificial neural networks*) [Bis96] ainsi que leur extension probabiliste [Spe90, WBX⁺07]. Ces objets fonctionnent via l'activation (plus ou moins marquée) de neurones organisées en couches permettant une classification rapide des signaux fournis en entrée. Leur structure rend leur optimisation tout comme le déroulement de la classification assez opaque, et il est difficile de savoir "ce qu'il se passe" dans un réseau de neurones entraîné, ce qui réduit les possibilités d'interprétation et de démarche explicative de l'identification.

Les SVM (*support vector machines*) constituent d'autres modèles très utilisés [Vap95, CV95]. Initialement prévus pour les problèmes de classification à deux classes, ils reposent sur la définition d'une frontière séparant les exemples des deux classes en maximisant une marge, distance entre la séparation et les vecteurs d'exemples les plus proches, dits vecteurs supports. Cette méthode permet surtout de séparer des classes non linéairement séparables dans l'espace vectoriel initial par le passage dans une dimension supérieure utilisant des noyaux (*kernels*) [BSS99].

Le passage à plus de classes peut alors se faire par exemple en entraînant autant de classifieurs SVM que de paires de classes. Chacun d'entre eux votera alors pour une classe, et celle ayant reçu le plus de votes sera décrétée comme celle reconnue par l'algorithme. Cette méthode de classification

pose toutefois quelques problèmes dans le cas où l'on souhaite obtenir en sortie une liste d'espèces, puisque les données ne sont jamais considérées dans leur ensemble, et il est possible que certaines classes ne reçoivent que très peu de votes et soient de ce fait pénalisées.

Une méthode prenant mieux en compte la distribution des données prises dans leur intégralité est celle des forêts aléatoires [Bre01]. Elles consistent en un ensemble d'arbres de décision utilisés pour prédire la classe des données fournies en entrée. Les arbres sont formés par une arborescence binaire de nœuds représentant chacun une décision relative à l'un des attributs représentant les données, et les feuilles de l'arbre sont étiquetées par les classes à reconnaître. Un nouvel exemple est ainsi classifié par l'arbre en suivant un chemin induit par des choix binaires à chaque nœud jusqu'à atteindre une feuille déterminant la classe qui va lui être affectée.

Les arbres de décision sont, suivant l'algorithme CART [BFSO84] (*classification and regression trees*), entraînés de façon à maximiser un taux de classification, en établissant à chaque nœud le critère qui permet de mieux séparer les classes des données d'apprentissage, qui sont alors considérées toutes simultanément. Tout l'intérêt des forêts aléatoires vient alors du fait que, non seulement plusieurs arbres (quelques centaines) sont utilisés, mais surtout qu'ils sont tous générés de façon aléatoire, le choix des attributs considérés à chaque nœud de chaque arbre se faisant au hasard. Chaque arbre vote ainsi pour une espèce, et les espèces sont classées selon leur nombre de votes.

Pris indépendamment, la performance de chaque arbre sera relativement faible, en tout cas non-optimale puisque le choix des variables à chaque nœud n'est pas nécessairement le meilleur (même si la position du seuil de décision, elle, est optimale). En revanche cette méthode non-déterministe a pour effet de réduire la corrélation entre les arbres, et par conséquent de produire une classification moins biaisée, capable d'une bonne généralisation même lorsqu'une classe est peu représentée. Les classes hétérogènes peuvent par ailleurs être relativement bien traitées par les forêts aléatoires, puisque qu'une même étiquette peut se retrouver à divers endroits des arbres, et les classes être ainsi scindées si nécessaire.

Cette dimension, ainsi que la prise en compte non seulement des étiquettes des données mais aussi de leur distribution au moment de la détermination des seuils de décision, nous a poussé à choisir cette approche pour la classification des attributs extraits sur les feuilles, un choix partagé par d'autres travaux se confrontant au même problème [AGB⁺12].

6.3.2 Sélection d'attributs et optimisation de paramètres

Nous avons choisi d'entraîner une forêt aléatoire en utilisant l'ensemble des paramètres vectoriels dont nous disposons. Cependant, ceux-ci peuvent différer selon le type de feuille considérée. En effet, il faut distinguer trois cas différents :

- Les feuilles **simples non-palmatilobées** pour lesquelles les paramètres de lobes sont inutiles, 27 attributs seront utilisés : $\mathbf{p}_{Simple} = (\mathbf{p}_{Polygone}; \mathbf{p}_{Base}; \mathbf{p}_{Apex}; \mathbf{p}_{Bord})$ numérotés de 1 à 27

- ▶ Les feuilles **palmatilobées** pour lesquelles il faut rajouter les paramètres de lobes, 34 attributs seront utilisés : $\mathbf{p}_{\text{Palmatilobee}} = \left(\mathbf{p}_{\text{Polygone}}; \mathbf{p}_{\text{Lobes}}; \mathbf{p}_{\text{Base}}; \mathbf{p}_{\text{Apex}}; \mathbf{p}_{\text{Bord}} \right)$ numérotés de 1 à 34
- ▶ Les feuilles **composées** où l'agencement des folioles est important, et pour lesquelles on rajoute donc les paramètres du modèle de feuille composée pour arriver à 35 attributs : $\mathbf{p}_{\text{Composée}} = \left(\mathbf{p}_{\text{Polygone}}; \mathbf{p}_{\text{Base}}; \mathbf{p}_{\text{Apex}}; \mathbf{p}_{\text{Bord}}; \mathbf{p}_{\text{Folioles}} \right)$ numérotés de 1 à 34

De façon à respecter cette distinction, nous entraînons donc trois forêts aléatoires avec des données différentes. Cela suppose donc qu'un nouvel exemple devra être reconnu par l'une d'entre elles seulement, et qu'il doit être possible d'effectuer ce choix au cours du traitement. Les feuilles composées subissant un traitement complètement différent dès la phase de segmentation il n'y a pas de risque de confusion. Pour les feuilles palmatilobées en revanche, il faut s'appuyer sur le paramètre de nombre de lobes du modèle de feuille pour prendre la décision, ce qui accentue encore le poids de son évaluation correcte sur le processus d'identification.

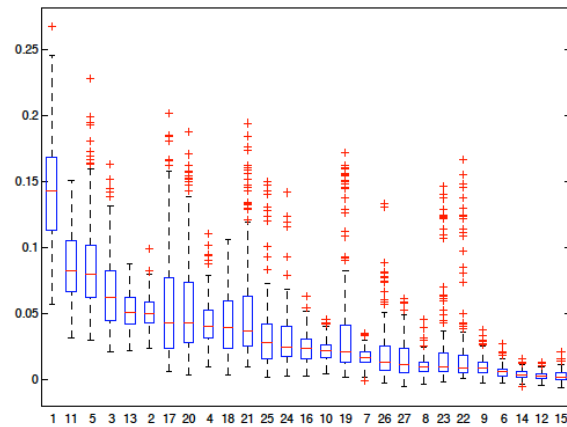
Les forêts aléatoires que nous utilisons comportent 200 arbres, ce qui constitue un nombre suffisant pour obtenir un classifieur stable en dépit de la part d'aléatoire dans l'apprentissage. Le paramètre important à considérer est le nombre d'attributs sélectionnés aléatoirement à chaque nœud. Une valeur communément employée est la racine carrée du nombre d'attributs, soit $\sqrt{|\mathbf{p}|}$, et c'est donc celle que nous emploierons comme valeur de base.

Dans les différentes catégories considérées, tous les attributs n'ont pas la même importance, et leur définition et leur évaluation peuvent faire que certains soient moins robustes et fiables que d'autres, ainsi que la méthode de classification plus naïve l'a mis en évidence. C'est notamment le cas sur les images naturelles où la segmentation est plus difficile et les attributs potentiellement plus bruités. Enfin certains descripteurs sont tout simplement moins représentatifs des espèces qu'ils représentent, et devraient à ce titre être pris en compte dans une moindre proportion.

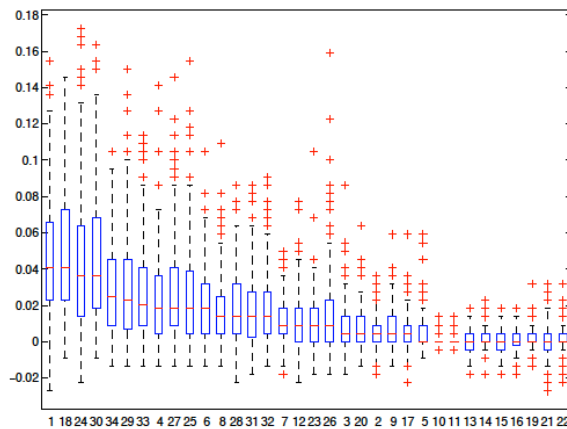
Afin d'évaluer la pertinence relative de chacun des attributs dans les différents contextes, nous avons choisi d'estimer leur mesure d'importance à partir du taux de classification des forêts aléatoires [Bre01]. Cette mesure est calculée comme la variation du taux de bonne classification de la forêt aléatoire dûment entraînée lorsque les valeurs de l'attribut considéré subissent une permutation aléatoire. De façon évidente, plus l'attribut est essentiel à la détermination de l'espèce correcte, plus la diminution sera importante. La Figure 6.20 illustre le résultat de cette procédure sur les différentes catégories d'images.

Il apparaît ainsi que certains attributs revêtent une importance toute particulière dans certains cas, alors que d'autres pourraient tout aussi bien être ignorés sans que cela n'affecte les performances de classification. Il va être ainsi possible de supprimer purement et simplement certains attributs dont la mesure d'importance avoisine 0 (voire se situe en dessous, ce qui tendrait à montrer que l'attribut est source de conflit) de façon à ne conserver que les sources pertinentes d'information. Les graphiques mettent ainsi en évidence les attributs les plus importants qui sont dans l'ordre :

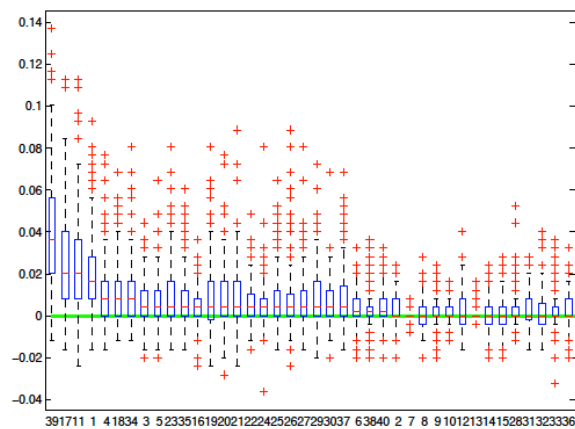
- ▶ Pour les feuilles simples :



(a)



(b)



(c)

FIGURE 6.20 – Importances mesurées pour les différents attributs, triés par ordre décroissant d'importance, pour les exemples de feuilles simples (a), palmatilobées (b) et composées (c) : attributs identifiés par leur numéro en abscisse, et importances en ordonnée, représentées par des diagrammes en boîte

- La largeur du modèle polygonal de feuilles l (1)
 - L'angle du modèle de forme d'apex α_A^* (11)
 - L'angle du modèle de forme de base α_B^* (5)
 - L'angle à la base du modèle polygonal α_B (3)
 - L'angle des tangentes au sommet du modèle d'apex θ_A (13)
 - La position verticale du modèle polygonal p (2)
 - La fraction lisse du bord de la feuille w_0 (17)
 - La courbure moyenne des dents du bord $\bar{\kappa}_+$ (20)
- Pour les feuilles palmatilobées :
- La largeur du modèle polygonal de feuilles l (1)
 - L'angle du modèle de forme d'apex α_A^* (18)
 - La fraction lisse du bord de la feuille w_0 (24)
 - L'échelle moyenne des creux du bord \bar{s}_- (30)
 - La courbure moyenne des dents du bord $\bar{\kappa}_+$ (34)
 - L'échelle moyenne des dents du bord \bar{s}_+ (29)
 - La courbure moyenne des creux du bord $\bar{\kappa}_-$ (33)
 - L'angle à l'apex du modèle polygonal α_A (4)
- Pour les feuilles composées :
- La distance des folioles à l'axe δ (39)
 - La fraction lisse du bord de la feuille w_0 (17)
 - L'angle du modèle de forme d'apex α_A^* (11)
 - La largeur du modèle polygonal de feuilles l (1)
 - L'angle à l'apex du modèle polygonal α_A (4)
 - L'échelle moyenne des dents du bord \bar{s}_+ (18)
 - La position de la dernière paire de folioles p_{n_F} (34)
 - L'angle à la base du modèle polygonal α_B (3)

Pour optimiser de manière un peu plus fine les paramètres des forêts aléatoires, nous avons également mené des tests de façon extensive, notamment pour déterminer le nombre optimal de variables à considérer lors de la sélection aléatoire à chaque nœud. De manière intuitive, lorsque ce nombre vaut 1, les arbres pris indépendamment ne seront pas très fiables, et la performance générale s'en ressentira ; d'un autre côté, plus on se rapproche du nombre total d'attributs, moins la part d'aléatoire est importante, plus les arbres sont corrélés et on perd tout le bénéfice inhérent à la définition de la forêt aléatoire. Il s'agit donc de trouver la valeur offrant le meilleur compromis entre ces deux extrémités. En mesurant les performances de classification avec les différentes valeurs possibles, nous sommes arrivés à des valeurs de 9 pour les feuilles simples, 16 pour les feuilles palmatilobées, et 15 pour les feuilles composées.

6.3.3 Résultats de classification

Les résultats que nous donnons ici sont des taux de bonne classification obtenus sur les différentes catégories. Il s'agit donc d'une information brute de performance qui ne prend pas en compte le rang

de l'espèce correcte, mais simplement si la classification est exacte ou non. La Figure 6.21 donne la distribution des taux de bonne classification suite à des expériences répétées.

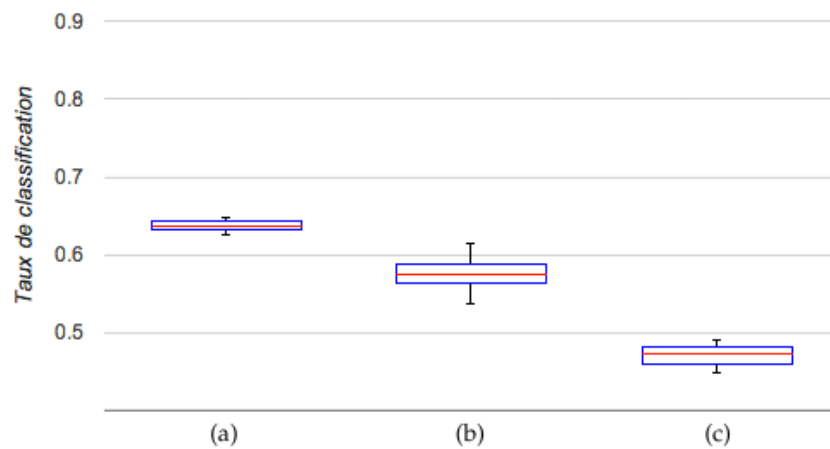


FIGURE 6.21 – Taux de bonne classification mesurés sur les images de feuilles simples (a), palmatilobées (b) et composées (c) représentées par des diagrammes en boîte

Les valeurs obtenues en moyenne (environ 64% pour les feuilles simples, 47% pour les feuilles composées) sont plutôt satisfaisantes et dépassent très nettement celles qui avaient pu être atteintes par une méthode de classification plus naïve. Cela montre bien que la distribution des formes au sein des espèces végétales nécessitent une représentation assez fine, possibilité qui est offerte par les forêts aléatoires, qui améliorent de façon remarquable la précision de la classification.

Cette amélioration peut également être mise en évidence de façon plus objective. La classification par forêts aléatoires ainsi que celle par estimation de modèles gaussiens sur les mêmes données ont toutes deux été utilisées lors de notre participation au challenge d'identification de plantes **ImageCLEF 2012** [GB]⁺[12]. Les résultats obtenus lors de cette compétition, bien que légèrement biaisés par le fait que toutes les images ne rentraient pas dans le cadre pour lequel nous avons développé nos méthodes (photographies de feuillage, espèces non-considérées) offrent un éclairage intéressant sur nos différentes approches. La Figure 6.22 présente les résultats de l'ensemble des participants.

Ces résultats mettent bien en évidence l'apport des forêts aléatoires dans la performance de classification, systématiquement mieux classées, avec un écart moyen de près de 15% du score de classification entre les deux méthodes. A noter que les valeurs présentées ne représentent pas un taux de bonne classification brut mais une mesure harmonisée prenant en compte le rang de l'espèce réelle.

Plus important encore, ils montrent que nos méthodes d'identification se situent parmi les plus compétitives puisque notre participation se classe quatrième équipe dans la catégorie scan, et troisième dans les catégories pseudoscan et photograph. Cette performance constitue une véritable validation de nos méthodes de description et de segmentation des images de feuilles qui permettent d'atteindre des résultats très satisfaisants, au niveau de l'état de l'art en la matière, tout en offrant une démarche de haut niveau originale et adaptée à l'utilisation sur des supports mobiles que nous souhaitons en faire.

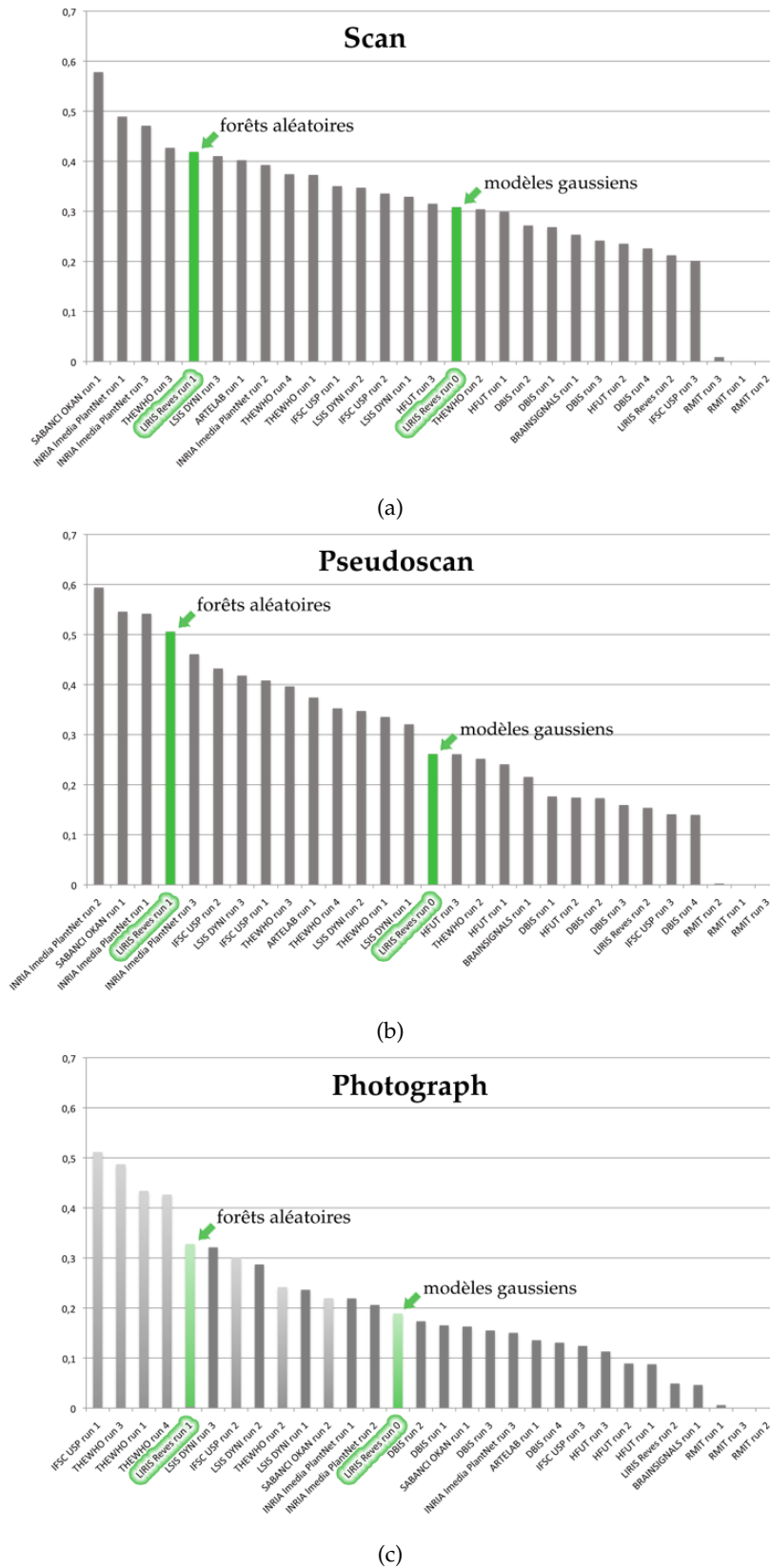


FIGURE 6.22 – Scores obtenus au challenge ImageCLEF 2012 par la classification par forêts aléatoires et par modèles gaussiens, sur les catégories scan (a), pseudoscan (b) et photograph (c)

Bibliographie

- [AGB⁺12] Akhil Arora, Ankit Gupta, Nitesh Bagmar, Shashwat Mishra, and Arnab Bhattacharya. A plant identification system using shape and morphological features on segmented leaflets. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [BCF⁺08] P. Belhumeur, D. Chen, S. Feiner, D. Jacobs, W. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. Sheorey, S. White, and L. Zhang. Searching the world’s herbaria : A system for visual identification of plant species. In *European Conference on Computer Vision*, 2008.
- [BFSO84] Leo Breiman, Jerome Friedman, Charles J. Stone, and R. A. Olshen. *Classification and Regression Trees*. Chapman & Hall/CRC, 1984.
- [Bis96] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, USA, 1996.
- [Bre01] Leo Breiman. Random forests. In *Machine Learning*, pages 5–32, 2001.
- [BSS99] C. Burges, B. Scholkopf, and A. Smola. *Advances in Kernel Methods : Support Vector Learning*. MIT press, 1999.
- [Cos06] H. Coste. *Flore descriptive et illustrée de la France de la Corse et des contrées limitrophes*. Librairie de l’Académie impériale de Médecine, Paris, 1906.
- [CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. In *Machine Learning*, pages 273–297, 1995.
- [Dem68] Arthur P. Dempster. A generalization of bayesian inference. *Journal of the Royal Statistical Society*, 30 :205–247, 1968.
- [DLR77] Arthur P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1) :1–38, 1977.
- [GBJ⁺12] H. Goëau, P. Bonnet, A. Joly, D. Barthelemy, N. Boujemaa, and J.-F. Molino. The ImageCLEF 2012 plant image identification task. In *ImageCLEF 2012 Working Notes*, 2012.

- [Hu62] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2) :179–187, 1962.
- [KBB⁺12] Neeraj Kumar, Peter Belhumeur, Arijit Biswas, D. Jacobs, I. Kress, W. Lopez, and Soares Joao V. B. Leafsnap : A computer vision system for automatic plant species identification. In *European Conference on Computer Vision*, volume 2, pages 502–516, 2012.
- [LCV13] HongHui Liu, Didier Coquin, and Lionel Valet. Reconnaissance de feuilles d’arbres par fusion de décisions partielles. In *GRETSI*, 2013.
- [Mah36] P. C. Mahalanobis. On the generalised distance in statistics. In *Proceedings National Institute of Science, India*, volume 2, pages 49–55, 1936.
- [Sha76] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [Spe90] D. Specht. Probabilistic neural networks. *Neural Networks*, 3(1) :109–118, 1990.
- [Vap95] Vladimir Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., 1995.
- [WBX⁺07] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yuxuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. A leaf recognition algorithm for plant classification using probabilistic neural network. In *IEEE International Symposium Signal Processing and Information Technology*, 2007.

Chapitre 7

Intégration de données biogéographiques

Prendre en compte la plausibilité géographique des espèces



Prunus spinosa L.

Résumé du Chapitre

Dans ce chapitre, nous proposons une approche originale pour améliorer les performances d'identification des espèces en incluant des facteurs géographiques dans le processus de reconnaissance. Les espèces d'arbres possédant chacune leurs préférences en termes de climat et de sol, toutes ne sont pas rencontrées partout en France. L'idée est donc de s'appuyer sur cette connaissance géographique pour estimer la plausibilité de chaque espèce à l'endroit de l'observation. Cette tâche passe par l'établissement de cartographies précises des paramètres qui influent sur la répartition des espèces, et par l'évaluation d'une mesure numérique de leur plausibilité biogéographique. Cette information est ensuite fusionnée avec les résultats de l'image, et les premiers résultats tendent à montrer que cet ajout fournit une amélioration non négligeable. Il ne s'agit pour autant que d'un premier pas, mais les perspectives ouvertes par cette approche sont des plus prometteuses.

Les Sections 7.1 et 7.2 présentent des travaux menés par Céline SACCA, post-doctorante, Pierre-Olivier MAZAGOL, ingénieur d'études, et Thierry JOLIVEAU, professeur des universités, au laboratoire de géographie EVS

DANS l'histoire de la botanique, l'idée a émergé assez rapidement, dès le début du XIX^{ème} siècle, que la géographie avait une influence non négligeable sur la répartition des espèces végétales. En référant à l'échelle d'un territoire les espèces, ou les communautés d'espèces, présentes en différents endroits, et en cherchant à établir les relations avec les caractéristiques géographiques du milieu, la phytogéographie se situe ainsi au croisement de la botanique et de la géographie. Cette branche de la science des plantes est à rapprocher de la biogéographie qui s'intéresse de manière plus générale à l'explication de la répartition du vivant.

Les études phytogéographiques fournissent donc des théories liant certains facteurs environnementaux, climatologiques ou géologiques, à la présence potentielle des espèces végétales à l'état spontané. Par conséquent, la répartition théorique d'une espèce à travers un territoire est une information qu'il semble possible d'obtenir, et d'exploiter dans le cadre du processus d'identification. Les approches pour y parvenir sont multiples, passant par exemple par les travaux existants de cartographie des espèces, des populations d'espèces dans le cadre d'une approche phytosociologique, voire directement de facteurs biogéographiques de bas niveau à mettre en lien avec les caractéristiques écologiques des espèces.

Le fait de disposer d'outils de localisation suffisamment précis, notamment via le GPS aujourd'hui très répandu dans les téléphones, permet d'envisager l'association directe de données géographiques au résultat de l'analyse d'image. La prise en compte de cette source supplémentaire d'informations pour aider la reconnaissance peut elle aussi faire l'objet de différentes stratégies, selon que l'on considère que la géographie permet de limiter le nombre d'espèces potentielles, ou bien qu'elle doit juste venir compléter l'information extraite de l'image. Il est dans tous les cas nécessaire de garder à l'esprit que toute entreprise de cartographie représente une généralisation, laissant de côté des particularités éminemment locales qui permettraient à une espèce de s'adapter à un environnement par ailleurs généralement défavorable. Chercher à établir la plausibilité d'une espèce végétale à un endroit donné est donc un exercice périlleux. Mais si elle est pour ainsi dire impossible à établir avec une certitude totale, elle peut malgré tout permettre avec une approche suffisamment prudente d'avoir une bonne idée de la vraisemblance de l'identification tirée de l'image au regard des informations biogéographiques disponibles.

7.1 Cartographie des facteurs biogéographiques

La question principale posée par la phytogéographie est celle du lien entre facteurs géographiques et répartition des espèces. Reste à définir quels sont ces facteurs qui influent sur la présence d'un arbre, ou plus exactement sur sa capacité à survivre dans un environnement, et comment quantifier les préférences associées à chaque espèce. Ce sont au final ces connaissances, ou du moins leur modélisation théorique, qui doit permettre d'établir une relation entre une position géographique et les espèces que l'on est susceptible d'y rencontrer.

7.1.1 Paramètres écologiques déterminants

Les facteurs qui influent sur la présence des espèces ne sont pas excessivement nombreux, et sont à mettre en lien avec les éléments qui interviennent dans le développement d'une plante. Ces paramètres écologiques sont bien connus, et peuvent être séparés en deux grandes catégories, les facteurs climatiques et les facteurs pédologiques (ceux concernant le sol, à un degré plus fin que la géologie). Dans le cadre du territoire français, les travaux de référence [Jul12] permettent ainsi de distinguer un certain nombre de facteurs qui vont devoir être pris en compte, à commencer par les principales caractéristiques climatologiques :

- ▶ La **température**, certaines espèces étant plus ou moins tolérantes au froid ou au chaud
- ▶ L'**humidité**, certaines espèces pouvant privilégier une atmosphère plus ou moins sèche
- ▶ La **lumière**, certaines espèces pouvant s'épanouir mieux dans un milieu ombragé

Ces critères déterminants ne sont pas nécessairement mesurables directement, et correspondent parfois à des éléments très locaux. Par exemple le niveau de luminosité est un indicateur très variable, selon que l'on se trouve ou non sous couvert végétal la plante ne recevra pas la même quantité de lumière. Il est néanmoins possible de trouver des facteurs géographiques mesurables qui peuvent être rattachés à ces indicateurs climatologiques. Nous avons ainsi choisi, dans notre cadre, de nous intéresser aux grandeurs suivantes :

- ▶ La **température moyenne annuelle** : bien que les températures minimales et maximales soient peut-être plus déterminantes pour la survie d'une espèce, la moyenne fournit une bonne estimation des conditions de température du milieu
- ▶ Les **précipitations moyennes annuelles** : même si elles ne constituent pas une mesure directe de l'humidité atmosphérique, elles sont suffisamment influentes pour être prises comme approximation de ce paramètre
- ▶ L'**altitude** : sans être à proprement parler une donnée climatologique, elle a une influence sur la température, et notamment sur les jours de gel, et constitue donc une grandeur intéressante, en plus d'être facilement accessible

L'intérêt de ces grandeurs est qu'elles peuvent être mesurées et surtout cartographiées à l'échelle d'un territoire, et constituer une référence permettant de décrire les conditions climatologiques d'un endroit donné. Cela est notamment dû au fait que ces indicateurs varient peu localement et représentent donc un milieu de façon fiable. En revanche les facteurs en lien avec la lumière présentent une granularité trop fine pour que des grandeurs mesurables, comme l'**ensoleillement moyen annuel** ou l'**exposition**, soient suffisamment représentatives.

Les trois paramètres ainsi retenus ont été cartographiés à l'échelle de la France, avec une précision de 90 mètres (soit au niveau de la précision atteinte par les systèmes GPS). Différentes sources ont été utilisées, permettant ainsi d'obtenir à partir d'une localisation des informations climatologiques précises. Concernant l'altitude, c'est un modèle numérique de terrain obtenu par SRTM (*Shuttle Radar*

Topography Mission) qui a été utilisé, alors que les mesures climatologiques sont tirées des données AURELHY (Analyse Utilisant le RELief pour l'HYdrométéorologie) [BLB87] établies sur 30 ans de mesures. La Figure 7.1 donne un aperçu de ces cartes sur le territoire français.

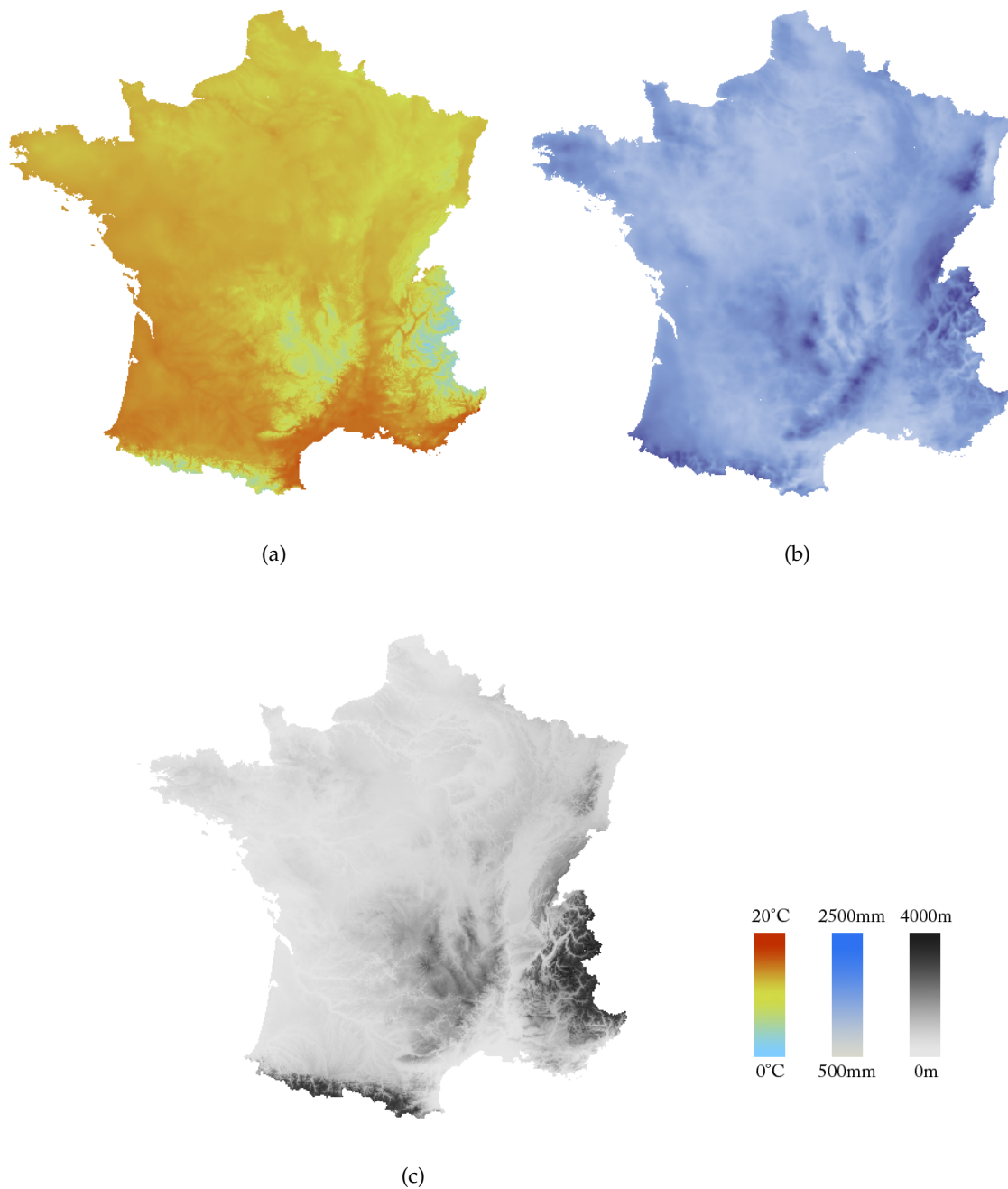


FIGURE 7.1 – Cartographie des paramètres écologiques en lien avec les indicateurs climatologiques sur la France métropolitaine : températures moyennes (a), précipitations moyennes (b) et altitude (c)

En plus des facteurs climatologiques, les caractéristiques du sol jouent un rôle important dans la répartition des espèces, et chacune a ses préférences particulières quant aux conditions pédologiques de son environnement. Les principales caractéristiques du sol qui interviennent dans la capacité d'une espèce à s'implanter dans un milieu donné sont les suivantes :

- ▶ L'**humidité**, certaines espèces affectionnant particulièrement les milieux humides
- ▶ La **salinité**, certaines espèces ne pouvant tolérer le sel, à proximité de la côte par exemple
- ▶ La **texture**, ou les types de sols, certaines espèces préférant les sols rocailloux ou argileux
- ▶ L'**acidité** du sol, mesurable par le pH, certaines espèces privilégiant les sols acides ou basiques

Malgré leur caractère déterminant, ces paramètres sont nettement moins faciles à cartographier de façon fiable. Le problème vient notamment de leur granularité, le fait qu'il s'agisse de caractéristiques de surface les rendant en effet beaucoup moins continus à l'échelle locale que ne peuvent l'être les facteurs climatologiques. Il est aussi difficile de s'assurer qu'avec une résolution de 90m, les unités de terrain soient véritablement uniformes du point de vue de ces paramètres. La présence d'un ruisseau peut par exemple permettre très localement à des espèces de milieux humides de proliférer, et cela ne sera vraisemblablement pas capturé par une carte à la résolution dont nous disposons.

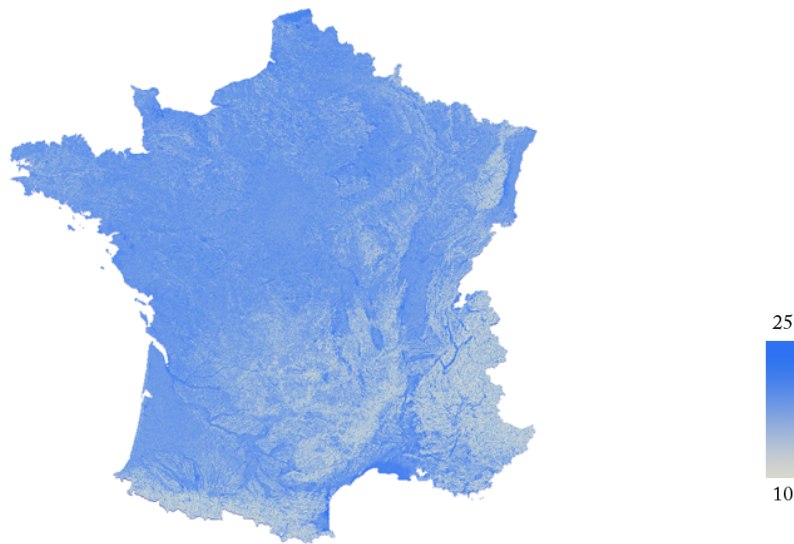


FIGURE 7.2 – Cartographie de l'indice d'humidité des sols sur la France métropolitaine

Il serait malgré tout possible de produire une cartographie approximative, par exemple des zones potentiellement humides, en se basant sur une cartographie des cours d'eaux et des fonds de vallées. Un indice d'humidité des sols ainsi cartographié est présenté dans la Figure 7.2. Ce genre d'approche présente malgré tout trop de risques d'imprécision, et d'occlusion des facteurs locaux, pour être véritablement pris en compte pour l'identification. Par conséquent nous ne considérerons que les cartographies des paramètres climatologiques pour les préférences des espèces. Bien que

la caractérisation ainsi produite soit incomplète, il s'agit des seules mesures locales pouvant être exploitées de façon fiable.

7.1.2 Indices intégrateurs

En parallèle des grandeurs mesurables qui peuvent fournir des informations pertinentes pour juger de la plausibilité des espèces à un endroit donné, il est aussi important de s'intéresser aux travaux déjà menés en biogéographie pour la cartographie de la répartition des espèces. Si la présence des espèces elles-mêmes n'est jamais véritablement cartographiée à l'échelle d'un territoire, l'objectif de nombreux travaux consiste à délimiter des régions homogènes du point de vue des conditions environnementales qu'elles présentent et des espèces qu'elles contiennent.

Pour la France, un exemple d'une telle cartographie est celle des sylvoécorégions [Inv11] qui divise le territoire métropolitain en plusieurs unités homogènes. En particulier, les espèces d'arbres présentes font partie des éléments qui permettent de définir ces régions, et par conséquent la connaissance de la sylvoécorégion dans laquelle on se trouve fournit une idée des essences qui sont susceptibles d'être rencontrées. A une plus large échelle, ces régions sont regroupées en 11 **grandes régions écologiques** (GRECO) elles aussi globalement homogènes, et qui peuvent donc être utilisées pour obtenir une cartographie de la répartition des espèces. La Figure 7.3 donne les limites de ces grandes régions.

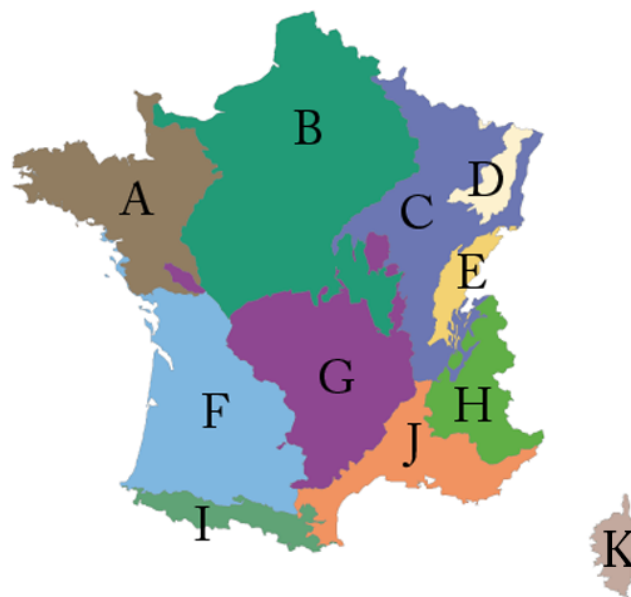


FIGURE 7.3 – Grandes Régions Ecologiques [Inv11] définies par l'inventaire forestier Français, identifiées par leur lettre

Un autre indice de haut niveau est celui des **étages de végétation** qui correspond aux différents niveaux successifs de végétation que l'on peut rencontrer en montant en altitude. Il se caractérise également par une homogénéité des caractéristiques du milieu, qui se répercute sur les espèces que l'on

peut y trouver, chacune pouvant être rencontrée sur une plage d'étages limitée. Leur définition intègre différents critères géographiques, à commencer évidemment par l'altitude, mais pas uniquement puisque l'altitude jusqu'à laquelle monte un étage va par exemple être différente selon le versant, ou selon la température moyenne.

Ces étages ont été cartographiés utilisant une définition phytogéographique [Jul99] spécifiant 9 zones de différentes altitudes, températures et orientations, en se basant sur des cartes de végétation du CNRS au 1 :200000e. La donnée de l'étage correspondant à un endroit donné renseigne alors directement sur les espèces plausibles. Les 9 étages ainsi recensés, et cartographiés à l'échelle de la France (Figure 7.4) sont les suivants :

- | | | |
|---|--|--|
| ▶ ■ Nival | ▶ ■ Montagnard | ▶ ■ Altiméditerranéen |
| ▶ ■ Alpin | ▶ ■ Collinéen | ▶ ■ Mesoméditerranéen |
| ▶ ■ Subalpin | ▶ ■ Planitiaire | ▶ ■ Thermoméditerranéen |

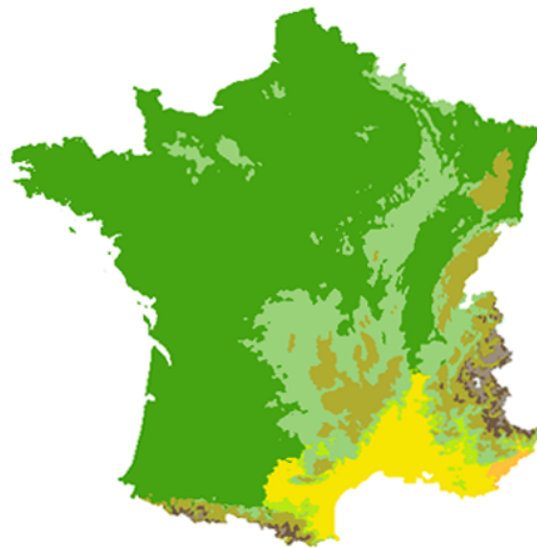


FIGURE 7.4 – Cartographie des étages de végétation sur la France métropolitaine

Ces deux indices, étages de végétation et GRECO, ne sont pas à proprement parler des mesures géographiques, mais plutôt des paramètres intégrateurs, apportant une information de plus haut niveau, puisque résumant en un seul indicateur plusieurs variables climatologiques ou géologiques. La dépendance avec la répartition est alors peut être moins directe, mais la présence d'une espèce peut ainsi être jugée avec un regard plus généralisateur, et probablement moins risqué qu'avec des mesures locales potentiellement variables.

7.2 Mesures de plausibilité des espèces

Les paramètres biogéographiques ainsi cartographiés permettent, en connaissant une localisation précise (par exemple les coordonnées longitude-latitude données par un système GPS), d'avoir une bonne idée de ses conditions environnementales. Reste pour obtenir une information exploitable dans le cadre de la reconnaissance à faire le lien entre ces conditions et les préférences des espèces considérées, de façon à quantifier la plausibilité de chacune d'elles de s'être implantée à une position donnée.

Pour établir cette relation, il est évidemment possible de s'appuyer sur les travaux déjà menés en phytogéographie, dont ces liens constituent le principal objet d'étude. A partir de telles données, indiquant par exemple la préférence de chaque espèce relativement aux différents paramètres géographiques, il est en effet possible de dériver une mesure numérique de plausibilité. L'autre approche consiste à se baser uniquement sur des observations des espèces, précisément géolocalisées, de façon à inférer un modèle de leur répartition relativement aux différents indices, dans une optique qui relève alors grandement de l'apprentissage automatique.

7.2.1 Répartition phytogéographique théorique

La façon la plus simple d'appréhender la répartition d'une espèce est de se baser sur les données présentes dans les flores, de la même façon que nous avons pu le faire pour les données de forme des feuilles. Dans certains cas, les régions dans lesquelles l'espèce se rencontre sont juste nommément listées, mais parfois ce sont les étages de végétation, voire même une carte de répartition, qui sont présentés [RMD⁺89].

FAGACÉES

Quercus petraea (Mattus.) Liebl.
Chêne sessile, Chêne rouvre

DISTRIBUTION GÉOGRAPHIQUE

- commune partout en plaine, sauf dans le Sud-Ouest, où elle est devenue disséminée, et dans la région méditerranéenne, où elle a toujours été rare ;
- jusqu'à 1600 m ; de l'étage collinéen à la base de l'étage montagnard ; étage supraméditerranéen ;
- européenne à tendance subatlantique.

QUERCUS



(a)



(b)

FIGURE 7.5 – Exemples de données biogéographiques théoriques pour le chêne rouvre (*Quercus petraea* Liebl.) : description tirée de la flore de Rameau [RMD⁺89] (a) et correspondance avec la carte des GRECO (b)

Ce genre de données fournit ainsi directement une information en lien avec les cartes dont nous disposons. On remarque en effet dans le cas des cartes de répartition que les frontières apparentes

correspondent assez bien avec celles des grandes régions écologiques, et il va donc être possible à partir de la carte de qualifier la présence d'une espèce dans chaque GRECO. Pour les étages, c'est encore plus simple puisque la correspondance est directe. La Figure 7.5 donne un exemple de l'exploitation des données tirées des flores. Il faut alors simplement définir une façon de quantifier cette information pour en tirer des mesures numériques exploitables.

A partir des données théoriques de répartition, nous construisons ainsi une description pour chaque espèce caractérisant sa présence potentielle au regard des différents indices. Il s'agit ainsi d'une description sur trois niveaux :

- ▶ Pour chaque **étage de végétation**, l'espèce est caractérisée comme **présente** ou **absente** en fonction des données de la flore.
- ▶ Pour chaque **grande région écologique** la carte de répartition et les données géographiques permettent d'associer à l'espèce un niveau de présence parmi : **commune**, **assez commune**, **disséminée**, **assez rare**, **rare** et **absente**.
- ▶ Eventuellement, l'espèce est limitée par une **altitude maximale** ou une **altitude minimale** lorsque celles-ci sont données.

On dispose alors pour l'espèce E_e , de valeurs de présence pr pour chaque étage ϵ , pour chaque GRECO g et également en fonction de l'altitude α :

- ▶ $pr(E_e, \epsilon) \in \{\text{présente, absente}\}$
- ▶ $pr(E_e, g) \in \{\text{commune, assez commune, disseminee, assez rare, rare, absente}\}$
- ▶ $pr(E_e, \alpha) \in \{\text{présente, absente}\}$

A une localisation donnée correspondent une valeur d'**étage**, une **altitude** et une **GRECO** unique, qui vont permettre d'estimer la plausibilité de chaque espèce. Nous utilisons les indicateurs de présence des espèces pour obtenir à partir de ces valeurs une mesure numérique de plausibilité **pl** calculée de manière additive pour garder une approche assez prudente. Les limites d'altitude étant difficiles à appréhender, notamment pour savoir si les valeurs sont à considérer comme des seuils binaires, ou s'il faut y voir une transition, et du fait de leur redondance partielle avec les étages, nous avons choisi de les laisser de côté. La mesure de plausibilité utilisée est alors la suivante :

$$\mathbf{pl}_{Geo}(g, \epsilon, E_e) = \frac{1}{2} \left\{ \begin{array}{l} 1 \text{ si } pr(E_e, g) = \text{commune} \\ 0.95 \text{ si } pr(E_e, g) = \text{assez commune} \\ 0.9 \text{ si } pr(E_e, g) = \text{disseminee} \\ 0.8 \text{ si } pr(E_e, g) = \text{assez rare} \\ 0.7 \text{ si } pr(E_e, g) = \text{rare} \\ 0.5 \text{ si } pr(E_e, g) = \text{absente} \end{array} \right. + \frac{1}{2} \left\{ \begin{array}{l} 1 \text{ si } pr(E_e, \epsilon) = \text{présente} \\ 0.5 \text{ si } pr(E_e, \epsilon) = \text{absente} \end{array} \right. \quad (7.1)$$

A partir des coordonnées GPS, il est possible grâce aux cartographies des paramètres géographiques, de déterminer les valeurs de ϵ et g correspondant à la localisation de l'image, et donc le score de plausibilité défini ci-dessus pour chacune des espèces. Cette valeur n'est jamais nulle, mais elle permet déjà de séparer les espèces correspondant parfaitement au milieu de celles qui y sont moins probables. En associant cette mesure au résultat de l'identification à partir de l'image, il sera alors possible de raffiner les résultats de l'identification.

Afin de mieux visualiser, nous avons calculé cette valeur de plausibilité pour les différentes espèces sur l'ensemble du territoire français, en utilisant les cartes des paramètres biogéographiques. Les résultats peuvent alors être vus comme des cartes de répartition des espèces, ou les endroits où l'espèce est la plus plausible apparaissent en vert foncé, et ceux où elle a peu de chances d'être rencontrée en vert plus clair. Des exemples de telles cartes sont donnés dans la Figure 7.6.

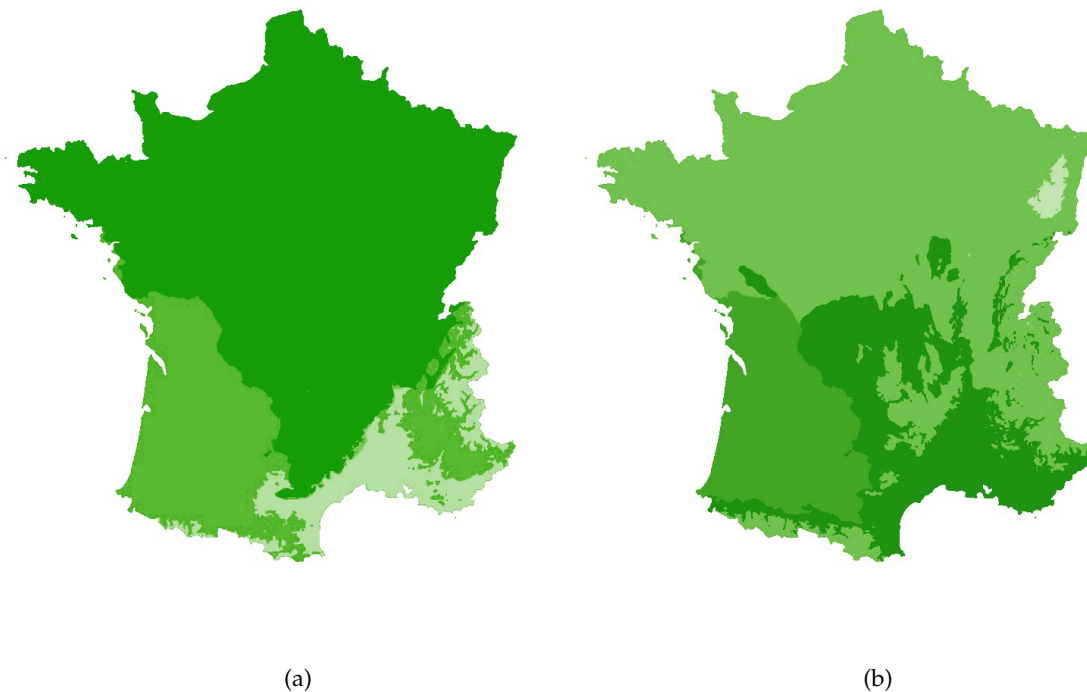


FIGURE 7.6 – Cartes de plausibilité obtenues à partir des présences théoriques des espèces 0.0 — 1.0 : chêne rouvre (*Quercus petraea* Liebl.) (a) et érable de Montpellier (*Acer monspessulanum* L.) (b)

7.2.2 Observations localisées et distribution empirique

L'autre méthode pour déterminer une mesure analogue de plausibilité consiste à voir cette tâche comme un problème de classification automatique. Plus exactement, l'idée est de construire pour chaque espèce un modèle permettant de dire si oui ou non l'espèce est plausible selon les conditions du milieu, à la manière d'un classifieur *one-class*, avec de préférence la possibilité de garder une mesure floue, entre 0 et 1, plutôt qu'une affectation dure.

Cela nécessite de disposer d'une base d'exemples suffisamment fournie pour permettre l'apprentissage des modèles. Nous avons choisi d'utiliser la banque d'observations SOPHY [BDRGH96] construite à partir de relevés phytosociologiques. Ces observations sont géolocalisées de manière précise, ce qui en fait des données particulièrement pertinentes pour notre approche. L'obstacle inhérent à ce type de données est cependant qu'il s'agit uniquement d'informations de présence, et que nous ne disposons pas d'exemples négatifs (des absences géolocalisées) ce qui rend l'entraînement d'un classifieur assez complexe.

La question se pose aussi du choix des attributs à considérer. Dans notre cas, nous avons souhaité évaluer des modèles gaussiens de répartition des espèces, ce qui suppose que les paramètres pris en compte soient des grandeurs numériques d'un espace continu, et non des indices discrets comme peuvent l'être les étages par exemple. Par conséquent nous nous placerons dans un espace tri-dimensionnel prenant en compte les trois attributs climatologiques : la température moyenne, les précipitations moyennes et l'altitude. Chaque observation x de l'espèce E_e est donc caractérisée par un triplet $\mathbf{o}_{e,x} = (T_x, P_x, a_x)$ que l'on peut voir comme un vecteur sur lequel appliquer des outils statistiques. Chaque observation vient en plus dans la banque SOPHY avec une information d'abondance, qui va permettre d'associer un poids $w_{e,x}$ entre 0 et 1 à chacune des observations.

L'idée est alors simplement d'effectuer, à partir de l'ensemble des observations pour une espèce, une modélisation gaussienne de la distribution de ces vecteurs décrivant l'environnement d'une observation, chacun pondéré par l'abondance de l'espèce dans le relevé. Nous utilisons ici encore une représentation par un mélange de gaussiennes, dont le nombre de composantes est fixé à $K = 3$, et qui est estimé par l'algorithme EM à partir des vecteurs centrés-réduits. Chaque espèce se voit ainsi associée à un modèle $\mathbf{GMM}_e = \left(\mathbf{O}_{e,k}; \Sigma_{e,k} \right)_{k=1..3}$ ce qui permet de couvrir des distributions éventuellement peu adaptées à la représentation par une gaussienne unique. La Figure 7.7 montre un exemple de calcul de mélanges de gaussiennes sur les attributs climatologiques.

Pour mettre en évidence l'intérêt de cette représentation, nous avons essayé de mesurer l'adéquation de la modélisation avec la distribution des exemples. C'est un problème assez compliqué, notamment à cause de la multi-dimensionnalité des données. Nous avons donc choisi d'utiliser une méthode proche du test d'adéquation du χ^2 généralement utilisé pour des données à une dimension, en comparant un histogramme suivant les trois dimensions des observations et les probabilités correspondant à chacun des bins de l'histogramme, en utilisant une distance χ^2 .

Le score obtenu correspond à la somme de ces distances sur l'ensemble de l'histogramme, et il doit être le plus proche possible de 0 pour que l'adéquation soit parfaite. Sur l'ensemble des espèces de la base SOPHY, la valeur moyenne de l'adéquation en ne considérant qu'une gaussienne (moyenne-covariance) est de 0.923, et le fait de passer à une modélisation par un mélange de gaussiennes permet de le réduire 0.516, ce qui constitue une amélioration considérable de la modélisation.

Une telle représentation de la distribution des observations permet de calculer une mesure numérique de plausibilité. Il est important de noter que la valeur que l'on souhaite attribuer à l'espèce est bien différente d'une vraisemblance, ou d'une probabilité, puisque qu'elle doit idéalement valoir 1 dès

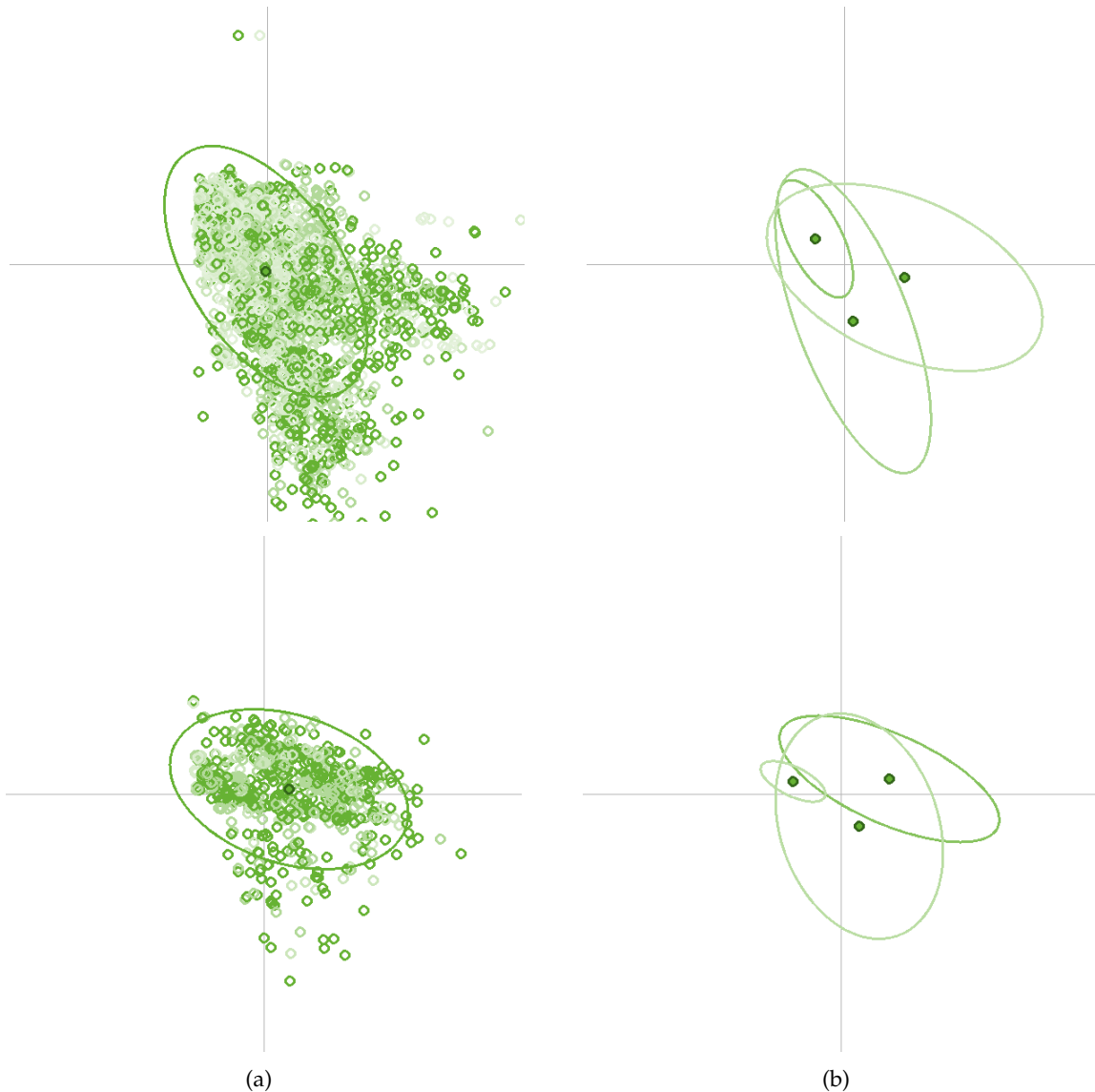


FIGURE 7.7 – Modélisations des distributions des observations suivant l’altitude (axe horizontal) et les précipitations (axe vertical) pour le chêne rouvre (*Quercus petraea* Liebl.) et l’érable de Montpellier (*Acer monspessulanum* L.) : modélisation par une gaussienne (a) et par un mélange de gaussiennes (b)

lors qu’on se trouve dans les conditions typiques de présence de l’espèce, et décroître en dehors. Nous nous basons donc sur la mesure de distance à la surface de l’ellipsoïde défini par le modèle gaussien, introduite au chapitre précédent. Cette distance a en effet la particularité d’être nulle à l’intérieur de l’hyper-ellipsoïde unitaire défini par la matrice de covariance, et de se comporter comme une distance euclidienne en dehors. La distance d’un vecteur \mathbf{o} définissant le milieu d’une observation au modèle \mathbf{GMM}_e d’une espèce est alors la suivante :

$$d_{Obs}(\mathbf{o}, E_e) = \min_k \|\mathbf{o} - \mathbf{O}_{e,k}\|_{\mathcal{E}} \quad (7.2)$$

On peut alors simplement définir une mesure de plausibilité de la même façon que nous avons défini les mesures de confiances pour la classification des espèces, par une exponentielle négative. On a ainsi la garantie que la plausibilité vaudra 1 à l'intérieur des hyper-ellipsoïdes définis par le mélange de gaussiennes :

$$pl_{Obs}(\mathbf{o}, E_e) = e^{-\frac{1}{2}d_{Obs}(\mathbf{o}, E_e)^2} \quad (7.3)$$

En utilisant cette définition de la plausibilité, on peut dans un premier temps calculer la plausibilité moyenne des observations utilisées comme références. Celle-ci s'élève à 0.941, contre 0.851 dans le cas d'une modélisation par une seule gaussienne. Il s'agit toutefois d'une mesure bien incomplète, puisqu'elle ne renseigne absolument pas sur le pouvoir discriminant de cette mesure de plausibilité. Il faudrait pour cela disposer d'exemples négatifs, pour garantir que la mesure de plausibilité soit faible pour ceux-ci. Avec la seule plausibilité moyenne, une mesure qui donnerait toute espèce parfaitement plausible partout atteindrait un score optimal, ce qui n'est évidemment pas désirable. C'est malheureusement un des seuls moyens d'évaluation à notre disposition à ce stade.

On peut néanmoins procéder à une visualisation des résultats sous forme de cartes de répartition, tout comme pour la plausibilité théorique. Les cartes obtenues sont toutefois beaucoup plus nuancées, et présentent donc des variations locales bien plus marquées, ce qui peut être un risque important pour la reconnaissance. La Figure 7.8 donne des exemples de cartes obtenues via cette modélisation de la répartition, et leur superposition avec les observations effectives tirées de la banque SOPHY.

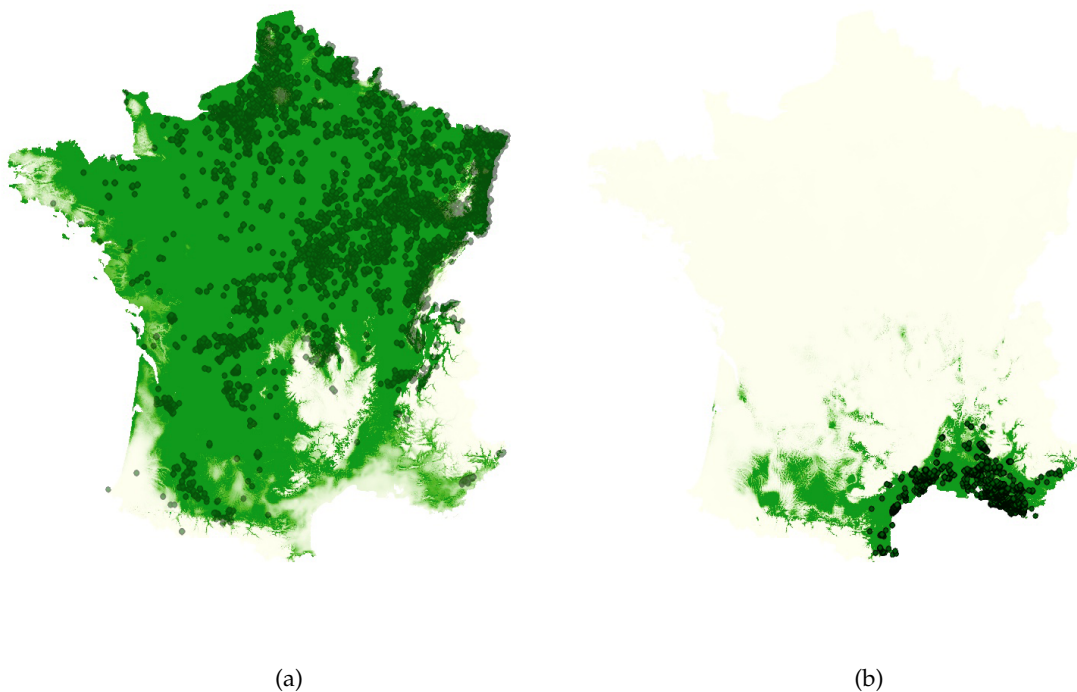


FIGURE 7.8 – Cartes de plausibilité obtenues à partir des observations géolocalisées 0.0 — 1.0 : charme commun (*Carpinus betulus* L.) (a) et chêne kermès (*Quercus coccifera* L.) (b)

Cette approche automatique constitue donc une alternative intéressante à la détermination de la plausibilité basée sur des données théoriques, ou du moins sur des travaux antérieurs spécifiques. Elle pourrait notamment bien se prêter à l'ajout progressif de nouvelles données en utilisant des observations géolocalisées, éventuellement hors du territoire français, pour inférer des modèles de répartition de nouvelles espèces. Elle présente malgré tout certaines limites difficiles à contourner, du fait d'une part de l'absence d'exemples négatifs, et d'autre part de la difficulté d'évaluer exactement la précision de la mesure de plausibilité produite et sa capacité à véritablement faire la décision.

7.3 Fusion de données et résultats

La prise en compte de la plausibilité des espèces considérées en fonction de la position géographique de l'observation soulève un certain nombre de questions. Il faut en effet s'interroger sur le type d'information qui est fourni par la biogéographie. C'est en tout cas une donnée d'un genre bien différent d'une classification, puisque rien ne garantit que l'espèce la plus plausible, étant données les conditions du milieu, soit effectivement celle que l'on va rencontrer. En revanche, une espèce peu plausible sera *a priori* à retirer du champ des possibles, mais que faire si tous les indices tirés de l'image pointent sur celle-ci ?

Pour tenter de répondre au mieux à ces problématiques, nous avons fait le choix de ne considérer l'information géographique qu'en fin de processus, en la fusionnant aux résultats de l'analyse de l'image. La considérer au début du traitement est une option séduisante, pour filtrer les espèces plausibles ou non et réduire le cadre de détermination, mais le risque d'exclure par erreur l'espèce qui correspondrait parfaitement à l'image est trop grand. Considérer d'autre part l'information au même niveau que les différents indices de formes serait par ailleurs minimiser quelque peu l'influence de la géographie, et l'utiliser simplement en bout de chaîne pour, en quelque sorte, réordonner les meilleures réponses du point de vue de l'image apparaît comme la solution la plus pertinente.

On fait ainsi l'hypothèse que l'image de la feuille est associée à un couple de coordonnées géographiques ($lon; lat$) donnant la localisation précise de la plante. Ces coordonnées permettent de trouver les paramètres biogéographiques correspondant à cette localisation, et en se basant sur les cartographies des paramètres, on peut obtenir le vecteur suivant

$$\left(g(lon, lat), \epsilon(lon, lat), T(lon, lat), P(lon, lat), a(lon, lat) \right)$$

C'est ce vecteur qui va permettre pour chaque espèce E_c de calculer une plausibilité géographique $pl(lon, lat, E_c)$, qui peut être calculée de deux façon différente, en s'appuyant soit sur les répartitions théoriques, soit sur les modèles tirés de la base d'observations :

$$pl(lon, lat, E_c) = \begin{cases} pl_{Geo}(g(lon, lat), \epsilon(lon, lat), E_c) \\ pl_{Obs}\left((T(lon, lat), P(lon, lat), a(lon, lat)), E_c \right) \end{cases}$$

La façon dont nous incluons cette plausibilité biogéographique dans le processus de reconnaissance est alors assez simple. Elle va venir pondérer les confiances obtenues à la fin de la classification en espèces, en considérant que sa plausibilité détermine en fait la valeur maximale de confiance que l'on peut accorder à une espèce. A partir des descripteurs \mathbf{p} extraits de l'image et de la position géographique (lon ; lat), on calcule alors la confiance finale accordée à l'espèce E_e de la façon suivante :

$$c_{Geo}(\mathbf{p}, \text{lon}, \text{lat}, E_e) = \mathbf{pl}(\text{lon}, \text{lat}, E_e) \times c_{\oplus}(\mathbf{p}, E_e) \quad (7.4)$$

L'introduction de la plausibilité biogéographique laisse inchangé le score de confiance des espèces qui sont en parfaite adéquation avec l'environnement du point d'observation. Par contre, elle va pénaliser les espèces qui ne le sont pas, les faisant chuter au classement et laissant des espèces plus plausibles s'élever vers les premières places. Il s'agit donc d'un moyen de réarranger les premières réponses de l'identification sans pour autant occulter l'information provenant de la forme. Ce mécanisme de fusion est en tout cas en cohérence avec l'idée que la géographie ne doit pas contredire les indices tirés du contenu de l'image, mais simplement fournir un éclairage sur son analyse qui doit permettre de prendre une meilleure décision finale.

Pour essayer de mesurer l'apport des données biogéographiques sur la performance de la classification, nous avons souhaité comparer les scores obtenus sur une base d'images géolocalisées avec et sans prise en compte de la plausibilité des espèces. La base **Pl@ntLeaves II** [GBJ⁺12] fournit une information de localisation sous forme de coordonnées longitude et latitude, et nous l'avons choisie à nouveau pour appliquer nos méthodes. Toutefois les espèces qu'elle contient ne sont pas toutes des espèces natives en France. Certaines espèces importées ne sont pas référencées dans les ouvrages botaniques dédiés au territoire français, et n'apparaissent naturellement pas dans les relevés d'observations. Nous nous sommes donc limités aux images d'espèces natives, réduisant du même coup le nombre d'espèces considérées.

Il est cependant important de prendre en compte un élément important quand aux informations de géolocalisation fournies avec les images. Il s'agit le plus souvent d'une information de la commune d'acquisition de l'image, traduite en termes de coordonnées, et non de la position exacte de la plante. De plus, l'essentiel des images étant des feuilles arrachées et placées sur fond uni, il est bien difficile de garantir que la position soit véritablement proche de celle de la plante. Par conséquent, l'imprécision est telle que les mesures locales correspondantes, comme l'altitude ou les facteurs climatologiques, pourraient bien être très différents des conditions effectives du milieu dans lequel l'individu a poussé. Il semble donc impossible de considérer la plausibilité tirée des observations comme une source fiable d'informations dans ce contexte de localisation imprécise.

En revanche, dans ce cadre particulier, les indices intégrateurs doivent avoir une granularité suffisamment large pour que cette imprécision ne soit pas trop préjudiciable. C'est donc la première expression de la plausibilité des espèces que nous avons utilisée pour mesurer l'influence de la géographie. La Figure 7.9 montre la comparaison entre une classification où la plausibilité biogéographique

est prise en compte et celle où elle ne l'est pas. Plus que les scores obtenus, ce qui est à retenir est l'apport manifeste de la connaissance géographique, qui donne simplement une meilleure classification, améliorant clairement le classement de l'espèce correcte.

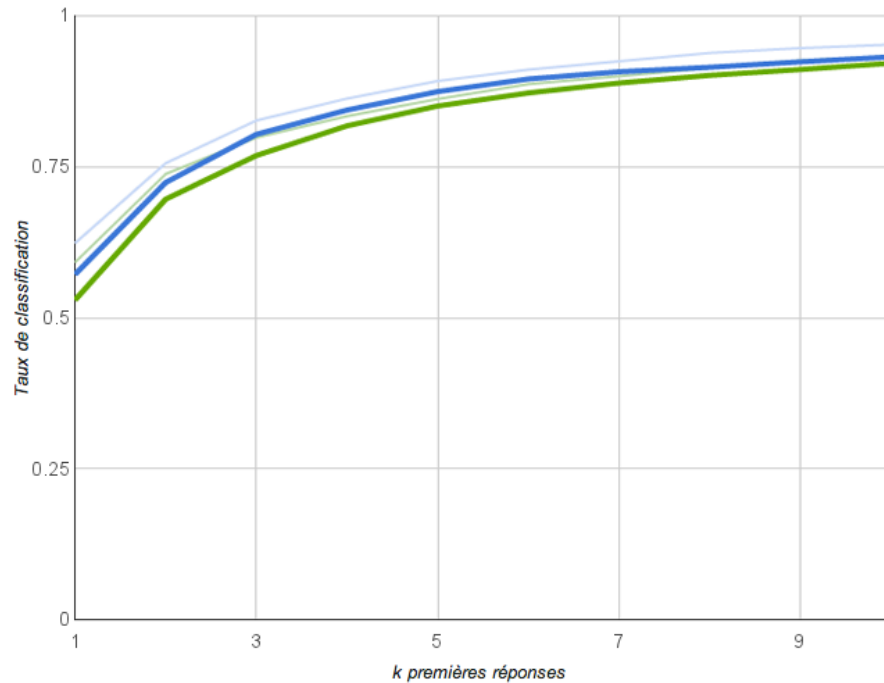


FIGURE 7.9 – Influence de la géographie sur la classification : scores de classification obtenus sur des images sur fond uni de feuilles d'espèces natives : ■ sans et ■ avec prise en compte de la plausibilité biogéographique

Pour mettre en lumière cette amélioration, nous avons également mesuré l'écart en rang de l'espèce correcte entre le classement sans information géographique et celui obtenu lorsqu'elle est prise en compte. La Figure 7.10 donne le résultat de cette comparaison sous forme d'histogramme de la différence de rang. Celle-ci doit idéalement être négative pour que la géographie apporte une amélioration. Le cas le plus représenté est celui d'une différence nulle, dans lequel il faut inclure toutes les feuilles qui étaient initialement bien reconnues (près de 55%) et pour lesquelles la plausibilité géographique ne pouvait pas amener de correction. A côté de cela, il apparaît assez nettement que l'apport de la géographie va dans le sens d'une amélioration pour une grande majorité des exemples, en tout cas bien plus souvent qu'elle ne conduit à des erreurs.

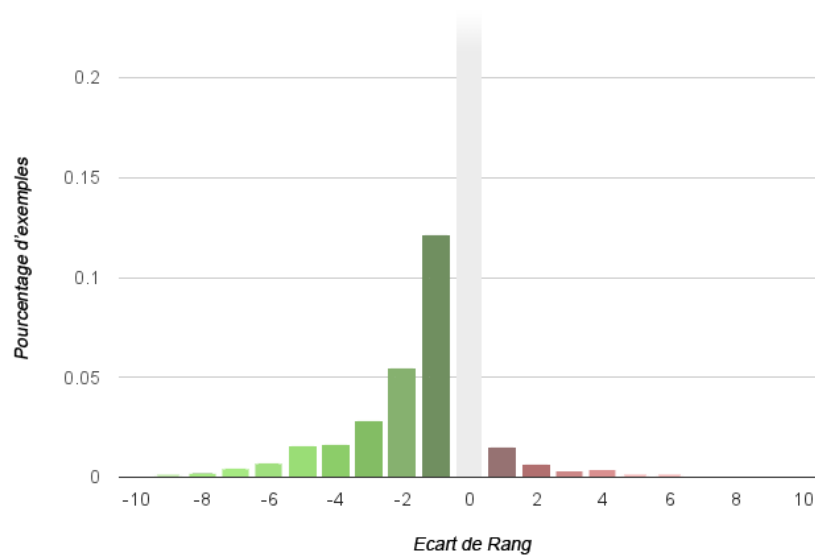


FIGURE 7.10 – Ecart de rang de l'espèce correcte dû à l'ajout de la plausibilité biogéographique dans la classification

Il semble donc que la connaissance de l'étage et de la région écologique permettent déjà de prendre une meilleure décision pour la reconnaissance de l'espèce. Au niveau des espèces considérées, on peut supposer que la prise en compte de cette information géographique de haut niveau permet surtout de séparer les espèces méditerranéennes des autres. En effet, les différences de milieu sont telles que certaines espèces sont inféodées au climat méditerranéen tandis que d'autres ne sauraient y survivre. La connaissance de cette simple distinction constitue déjà un filtre efficace pour limiter les risques de confusion.

En ce qui concerne la mesure de plausibilité prenant en compte des mesures plus locales, il faudrait, pour s'assurer de sa pertinence, disposer d'une base d'exemples dont la localisation soit absolument certaine, de façon à garantir que les coordonnées associées à l'image soient bien celles où la plante a été observée. De telles données ne sont pas faciles à obtenir, mais sans elles il serait impossible de valider véritablement les modèles de répartition. Pour constituer une telle base, l'idéal serait un outil collaboratif de collecte permettant à un nombre conséquent d'utilisateurs de transmettre des images géolocalisées de feuilles sur la plante. Nous avons espoir que l'application **Folia** dans laquelle nous avons implémenté nos méthodes, présentée dans l'Annexe B, permettra à terme de bâtir une banque de données étiquetées de façon fiable, qui constituera une base solide pour évaluer l'apport des informations biogéographiques pour l'identification des espèces.

Bibliographie

- [BDRGH96] H. Brisse, P. De Ruffray, G. Grandjouan, and M. Hoff. La banque de données phytosociologiques "SOPHY". In *International workshop on European vegetation survey*, pages 177–223, 1996.
- [BLB87] P Bénichou and O. Le Breton. AURELHY : une méthode d'analyse utilisant le relief pour les besoins de l'hydrométéorologie. In *Deuxièmes journées hydrologiques de l'ORSTOM, Montpellier*, pages 299–304, 1987.
- [GBJ⁺12] H. Goëau, P. Bonnet, A. Joly, D. Barthelemy, N. Boujema, and J.-F. Molino. The ImageCLEF 2012 plant image identification task. In *ImageCLEF 2012 Working Notes*, 2012.
- [Inv11] Inventaire Forestier National. *Les Sylvéoécorégions (SER) de France métropolitaine*. IGN, 2011.
- [Jul99] Philippe Julve. Carte phytogéographique de la France. *Cahiers de Géographie Physique*, 13, 1999.
- [Jul12] Philippe Julve. Baseflor. Index botanique, écologique et chorologique de la flore de France, version 31, <http://perso.wanadoo.fr/philippe.julve/catminat.htm>, 2012.
- [RMD⁺89] J.C. Rameau, D. Mansion, G. Dumé, J. Timbal, A. Lecointe, P. Dupont, and R. Keller. *Flore forestière française : Guide écologique illustré*. Institut pour le Développement Forestier, 1989.

Conclusions

Derniers regards en arrière, et vers de nouveaux horizons



Quercus petraea Liebl.

NU long du processus de reconnaissance des espèces d'arbres décrit, explicité et évalué dans cette thèse, un certain nombre de problèmes scientifiques ont été soulevés. Nous avons tenté d'y apporter des solutions qui, si elles ont été pensées pour fonctionner dans le cas spécifique de notre objet d'étude, ont une portée générale bien réelle. Les questions posées par un problème aussi difficile concernent à la fois la segmentation d'images et la représentation sous-jacente de son contenu, la description des formes et des contours, et des points de classification particulièrement complexes.

La principale contribution reste malgré tout la réalisation d'un système complet d'identification des espèces d'arbres à partir d'images de feuilles, permettant une démarche explicative, et prenant éventuellement en compte des données géographiques pour affiner la décision. Ce système se voit d'ailleurs concrétisé sous la forme de l'application *Folia* * pour iPhone, qui a fait l'objet de développements spécifiques détaillés dans l'Annexe B. Les méthodes développées dans ce but ont dans un autre contexte été validées par des performances honorables lors de participations successives au challenge ImageCLEF de reconnaissance de végétaux, qui ne se situait pas toujours exactement dans notre contexte de travail [GBJ⁺11, GBJ⁺12, GBJ⁺13].

Dans la suite nous résumerons donc les contributions apportées au niveau des différentes étapes du processus de reconnaissance, et leurs applications possibles dans des contextes différents. Concernant l'identification des espèces d'arbres enfin, les méthodes étant encore et toujours perfectibles, des pistes d'amélioration seront évoquées, pour donner une idée des axes possibles d'extension des travaux de recherche présentés dans cette thèse.

Analyse d'image et segmentation

Une part importante du travail a consisté à mettre au point une méthode satisfaisante pour traiter un problème extrêmement difficile : la segmentation d'objets naturels très variables dans des images acquises sous des conditions peu contrôlées en environnement complexe. Nous avons considéré qu'un processus purement automatique était dans ce cadre pratiquement voué à l'échec, et avons donc fait le choix d'une méthode supervisée, sans pour autant devenir véritablement une segmentation interactive. Ce choix était d'autant plus évident que, le but étant de réaliser une application pour smartphone, l'intervention d'un utilisateur peut s'envisager très facilement.

Nous proposons en premier lieu une modélisation originale des couleurs de l'objet d'intérêt, avec pour but d'estimer la correspondance de chaque pixel de l'image aux couleurs d'une région initiale représentative de l'objet. Cette modélisation, basée sur une combinaison d'une représentation globale de la couleur par régression linéaire dans l'espace $L^*a^*b^*$, et d'une estimation de la couleur moyenne locale de l'objet en tout point de l'image, est un moyen particulièrement efficace de séparer l'objet d'intérêt de son environnement même complexe. Ce n'est pas à proprement parler une segmentation,

*. *folia* : application gratuite pour iPhone et iPad, disponible sur l'App Store <https://itunes.apple.com/app/folia/id547650203>

Conclusions

mais le calcul d'une dissimilarité de couleur, notion plus floue qui peut ensuite servir de base à un algorithme de segmentation.

Cette représentation a été développée pour traiter des images naturelles de feuilles d'arbres, et prend donc en compte certaines spécificités (surface globalement uniforme en couleur, reflets potentiels, objets de même couleur dans le fond, variations locales de la couleur). Aussi peut-elle s'adapter à des objets complètement différents qui partagent quelques unes de ces particularités. Elle fonctionne par ailleurs tout aussi bien sur des images plus simples, ou des méthodes plus basiques seraient certes suffisantes. La Figure 7.11 donne quelques exemples de cartes de dissimilarité calculées suivant notre méthode sur des objets différents.

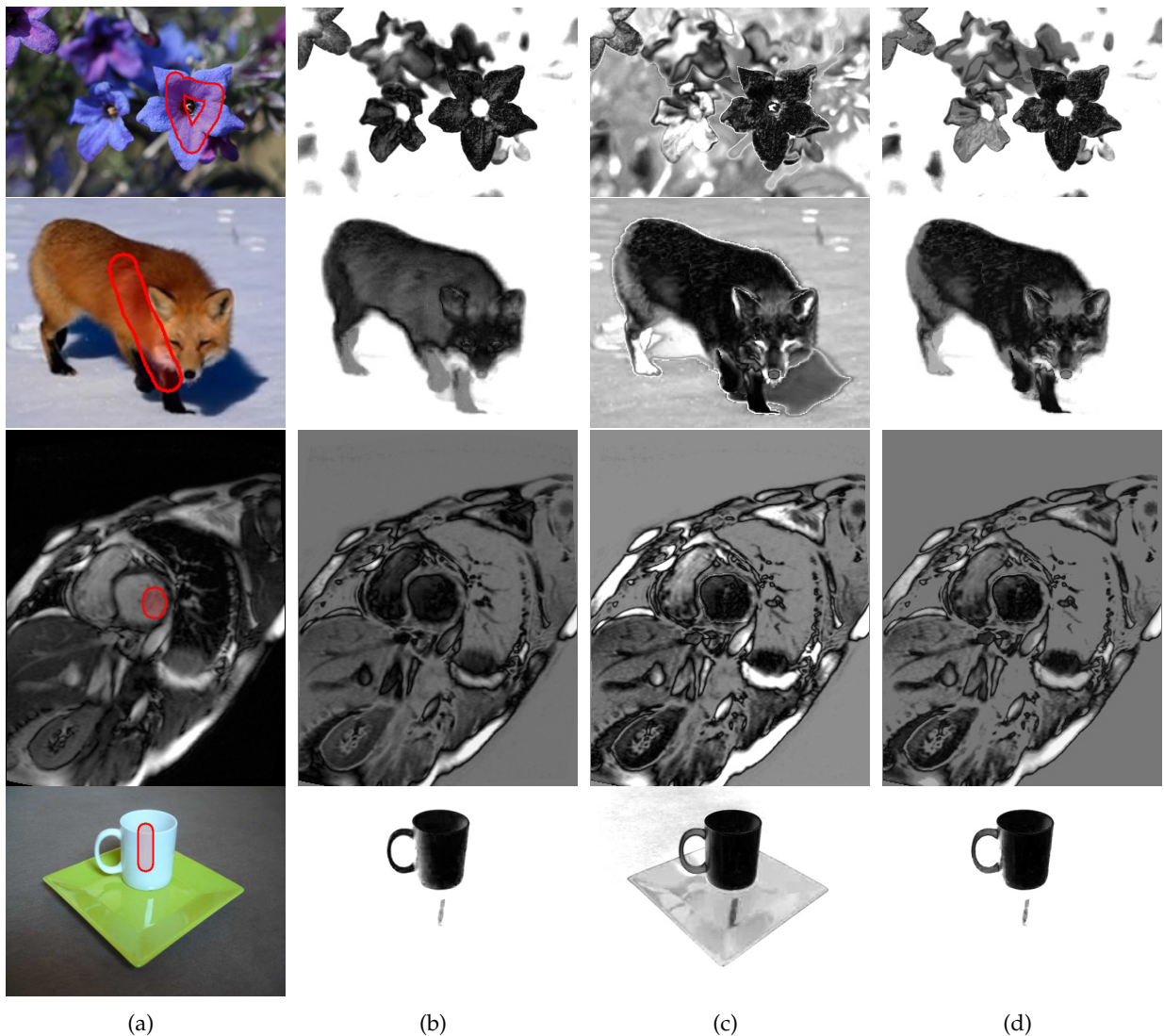


FIGURE 7.11 – Cartes de dissimilarité de couleur obtenues sur d'autres types d'objets : coloriage initial (a), dissimilarité au modèle global de couleur (b), dissimilarité à la moyenne estimée locale (c) et dissimilarité combinée (d)

La dissimilarité de couleur sert de base à l'évolution de modèles paramétriques déformables dédiés, dont le but est de s'ajuster à l'objet en dessinant une région aussi grande que possible contenant essentiellement des pixels similaires à la couleur de l'objet. La modélisation de forme utilisée pour les feuilles simples est un outil particulièrement avantageux pour effectuer une pré-segmentation de l'image, en garantissant que la forme obtenue corresponde à celle d'une feuille. Il s'agit d'une méthode particulièrement bien adaptée aux images naturelles, où le modèle bien que très flexible, agit comme un *a priori* de forme suffisamment robuste pour retrouver la feuille d'intérêt dans une image complexe [CTVC11].

Les modèles de feuilles utilisés ont été développés manuellement, de manière inductive, en se basant sur les enseignements tirés de la botanique. Une piste particulièrement intéressante serait que ces modèles puissent être appris automatiquement, en utilisant des statistiques de forme, à la manière des *active shape models*, mais en travaillant directement sur la forme pour une plus grande flexibilité. Sans doute que les axes principaux de variation correspondraient alors dans une certaine mesure aux paramètres que nous avons choisis comme variables des modèles, mais il serait intéressant d'utiliser une modélisation basée sur des exemples réels de forme.

En ce qui concerne les méthodes d'optimisation des paramètres des modèles, et donc l'ajustement du modèle à l'image, l'approche proposée de descente de gradient et cycles de températures forme un cadre général de minimisation d'énergie, qui s'applique de la même façon aux divers modèles introduits. Elle peut donc tout à fait être transposée à d'autres modèles, représentant d'autres objets, dans des cadres différents d'analyse d'images, où la segmentation pourrait bénéficier de connaissances sur la forme d'objets potentiellement variables.

Enfin, la dernière phase de segmentation basée sur un modèle de contours actifs introduit une contrainte forte de forme tout en laissant suffisamment de latitude au contour pour détecter les structures locale du bord des feuilles. Cet algorithme introduit pour notre première participation au challenge ImageCLEF [CTM⁺11] s'avère particulièrement compétitif, et se comporte mieux que la plupart des méthodes de segmentation de l'état de l'art [GBVC⁺13]. Cela constitue un véritable pas en avant pour la résolution du problème si complexe qu'est la segmentation d'image d'objets naturels.

L'extension de cette méthode aux cas des feuilles composées donne aussi des résultats intéressants [CTM⁺13a], dans un contexte assez particulier d'extraction simultanée de différents objets indépendants, sans en connaître initialement le nombre. L'apport de la modélisation est ici considérable, mais les problèmes abordés vont bien au delà du cas des feuilles d'arbres, et on pourrait envisager leur application sur des objets tout autres.

Les méthodes de segmentation introduites répondent donc de manière convaincante, bien que toujours perfectible, aux nombreux obstacles que l'on rencontre dans les images naturelles. Elles constituent le socle de tout le processus de reconnaissance, et ce sont elles qui permettent d'envisager une utilisation réelle de notre système, non cantonnée à un environnement supervisé mais capable de se confronter aux imperfections que ne manqueront pas de lui soumettre les utilisateurs.

Description de formes

Le fil conducteur de la définition des descripteurs à extraire de l'image, une fois la forme d'intérêt segmentée, a été la référence constante aux critères utilisés en botanique. La description des formes que nous avons proposée permet donc de caractériser indépendamment les différents aspects de la forme des feuilles, capturant ainsi les propriétés qui permettent aux botanistes de reconnaître les espèces. Cette description compréhensive de la feuille donne des résultats tout à fait satisfaisants [CTM⁺13b], comparables à des descripteurs de formes communément employés, en utilisant un nombre d'attributs généralement bien plus réduit.

Ces résultats tendent à confirmer l'intérêt d'inclure une part de connaissance spécifique à l'objet d'étude dans la définition des attributs utilisés pour la décrire. La description fournie sera ainsi bien souvent plus légère, et ne nécessitera pas de réduction de la dimensionnalité pour être pertinente. C'est un choix fort, qui n'est malgré tout pas sans risque, puisque l'extraction des descripteurs repose alors sur une interprétation de la forme, nécessaire pour considérer des informations de haut niveau, et que les erreurs à ce stade peuvent s'avérer extrêmement coûteuses.

Un enrichissement possible de la description que nous utilisons pourrait venir d'estimations de la confiance à accorder aux valeurs extraites, de façon à s'assurer du bon déroulement du processus de reconnaissance. Savoir à quelles informations donner le plus de poids, par exemple en quantifiant leur vraisemblance, ou en mesurant leur cohérence à travers une analyse multi-échelle, constituerait indéniablement un avantage pour prendre une meilleure décision.

Dans le cadre particulier de l'analyse du contour, nous avons proposé une méthode originale basée sur le *curvature-scale space* permettant de détecter les points les plus remarquables du contour et d'utiliser leurs propriétés comme une description de la forme. Cette information s'avère particulièrement utile dans le cas des feuilles, où elle améliore nettement la description [CTCV13], mais il serait possible d'en tirer une description plus générale des contours (sans prendre en compte des propriétés de symétrie, de zones particulières du contour) basée sur les propriétés des points dominants.

Les méthodes développées à ce titre en s'inspirant des chaînes de caractères offrent un moyen efficace d'inclure la dimension spatiale dans la description du contour. C'est un aspect qui manque souvent aux descripteurs génériques de formes, dont un certain nombre sont construits par agrégation, ou qui nécessite le cas échéant avec un coût en calcul non-négligeable. Les outils de traitement des chaînes offrent une complexité moindre, grâce au calcul par programmation dynamique, tout en fournissant une description visuellement assez fidèle de l'information de contour. Les outils développés permettent de manipuler des données vectorielles dans les chaînes, et pourraient par ailleurs tout aussi bien être utilisés avec d'autres descriptions des points dominants (SIFT, SURF, etc.).

Enfin, la description de formes que nous proposons vient avec un contenu visuel fort, qui la rend facile à interpréter pour un utilisateur. Cela s'inscrit dans le choix d'un processus explicatif de reconnaissance, et le haut niveau d'abstraction auquel se place les attributs du système rend plus facile la compréhension des étapes menant à la décision, et par conséquent des erreurs éventuelles.

Classification de données complexes

Afin de répondre aux difficultés posées par un problème de classification des plus ardues (nombre élevé de classes, variabilité intra-classe très importante et inconstante, recouvrement manifeste entre différentes classes) nous avons introduit plusieurs méthodes pour parvenir à l'identification des espèces. La première d'entre elles passe par la définition d'une distance adaptée à ce genre de données, qui permet d'obtenir très rapidement des résultats satisfaisants par l'évaluation de simples modèles gaussiens [CTM⁺13b].

Cette mesure de distance apparaît comme supérieure à la distance euclidienne et à la distance de Mahalanobis pour prendre en compte la variabilité des classes, et offre donc un champ d'application assez vaste, généralisable en dimension quelconque. La définition que nous avons introduite convient ainsi pour juger de l'appartenance de points expérimentaux à un ensemble de modèles gaussiens en permettant d'obtenir des valeurs comparables pour les différents modèles tout en prenant en compte la forme des distributions.

Nous avons ensuite proposé un système original de classification en termes de forme permettant de faire le lien entre les informations extraites de l'image et les données botaniques théoriques utilisées pour la détermination des espèces. Une telle approche fournit des informations sémantiques particulièrement intéressantes dans un contexte pédagogique, et pourrait même être exploitée pour la classification, à l'aide d'outils plus avancés [LCV13].

Les mécanismes de fusion proposés pour prendre une décision dans ces approches ont permis d'améliorer les performances de reconnaissance, et constituent un formalisme pouvant s'adapter à d'autres problèmes de classification. Les résultats obtenus sont meilleurs lorsque la décision s'effectue par fusion des réponses d'une première classification que lorsque les informations sont considérées d'un seul bloc, un résultat qui peut se généraliser à d'autres applications de la fusion d'informations.

Concernant la classification, les forêts aléatoires ont démontré une indéniable supériorité par leur capacité à s'adapter à la distribution des données, à représenter des classes hétérogènes, et à permettre une sélection efficace des attributs les plus représentatifs. La performance ainsi obtenue nous a permis d'être particulièrement compétitifs face aux méthodes de l'état de l'art [CAT⁺12] et impose donc ce classifieur comme le choix prioritaire pour ce type de données complexes.

Pour finir, nous avons tenté d'explorer un axe peu étudié pour la reconnaissance automatique en incluant des informations géographiques dans un processus de fusion de données hétérogènes. Les premiers résultats montrent qu'il s'agit d'une piste très prometteuse [CTS⁺13], mais sa validation effective nécessite des données difficiles à constituer, qui seules pourront permettre de juger du réel impact des données biogéographiques.

Perspectives pour l'identification de végétaux

Les travaux présentés dans cette thèse constituent une base solide pour un système d'identification des espèces d'arbres, mais la difficulté du problème est telle que des améliorations sont envisageables. A commencer par la segmentation, tant la complexité des images naturelles représente un défi pour les méthodes usuelles. L'intégration de plus de connaissance sur les formes à extraire de l'image, connaissance pouvant par ailleurs provenir d'exemples réels plus que de modèles artificiels, constitue une piste à privilégier au vu des résultats obtenus en utilisant des modèles de forme. En particulier, la prise en compte d'une forme de symétrie et de régularité à l'échelle du bord de la feuille dans l'évolution des contours actifs devrait permettre d'accroître la robustesse de la segmentation.

La possibilité d'utiliser plus de supervision de la part de l'utilisateur dans la phase de segmentation constitue également une piste à envisager. Demander à l'utilisateur de corriger les erreurs manifeste dans la forme extraite, et propager les informations qui peuvent être tirées de son intervention, devrait permettre d'améliorer significativement la qualité de la segmentation, présentant des résultats visuellement plus satisfaisants à l'utilisateur, tout en l'impliquant davantage dans un processus de reconnaissance interactif.

Concernant la description des formes foliaires, nous avons mis en lumière la pertinence d'une caractérisation par les éléments discriminants recensés par la botanique, mais tous n'ont pas été inclus dans notre représentation actuelle. C'est notamment la nervation qui manque à l'appel, et une détection et une modélisation explicite des nervures des feuilles pourrait venir s'ajouter à nos descripteurs comme une source supplémentaire d'informations permettant de discriminer un peu plus certaines espèces.

Le caractère hautement sémantique de nos descripteurs, renforcé par la classification possible en vocabulaire botanique, ouvre des perspectives particulièrement intéressantes d'études dans le cadre de l'interaction homme-machine et de l'apprentissage. Il semble pertinent de se pencher sur la façon dont un utilisateur perçoit le lien entre la présentation qui lui est faite des informations extraites de l'image, ce qu'il peut lui même y observer, et la décision prise par le système en termes d'espèce, pour savoir si une compétence d'identification est ainsi transmise. Il est par ailleurs possible qu'une plus grande interactivité contribue à une meilleure transmission des connaissances dont notre application se veut le porteur.

La représentation des espèces par des modèles canoniques permet dans un autre domaine d'établir des relations de proximité entre elles, voire également des liens entre les espèces et les termes de forme, qui peuvent constituer des données propices à une visualisation interactive. La Figure 7.12 montre un exemple de graphe de similarité des espèces, et on peut imaginer une interface permettant à un utilisateur de naviguer à travers les espèces par similarité, où de situer la réponse du système dans une carte des espèces plutôt que sous forme de liste. Il peut également s'agir d'une source complémentaire de connaissance, utile pour différentes méthodes de recherche d'information, dans le contexte des images ou dans un cadre plus général.

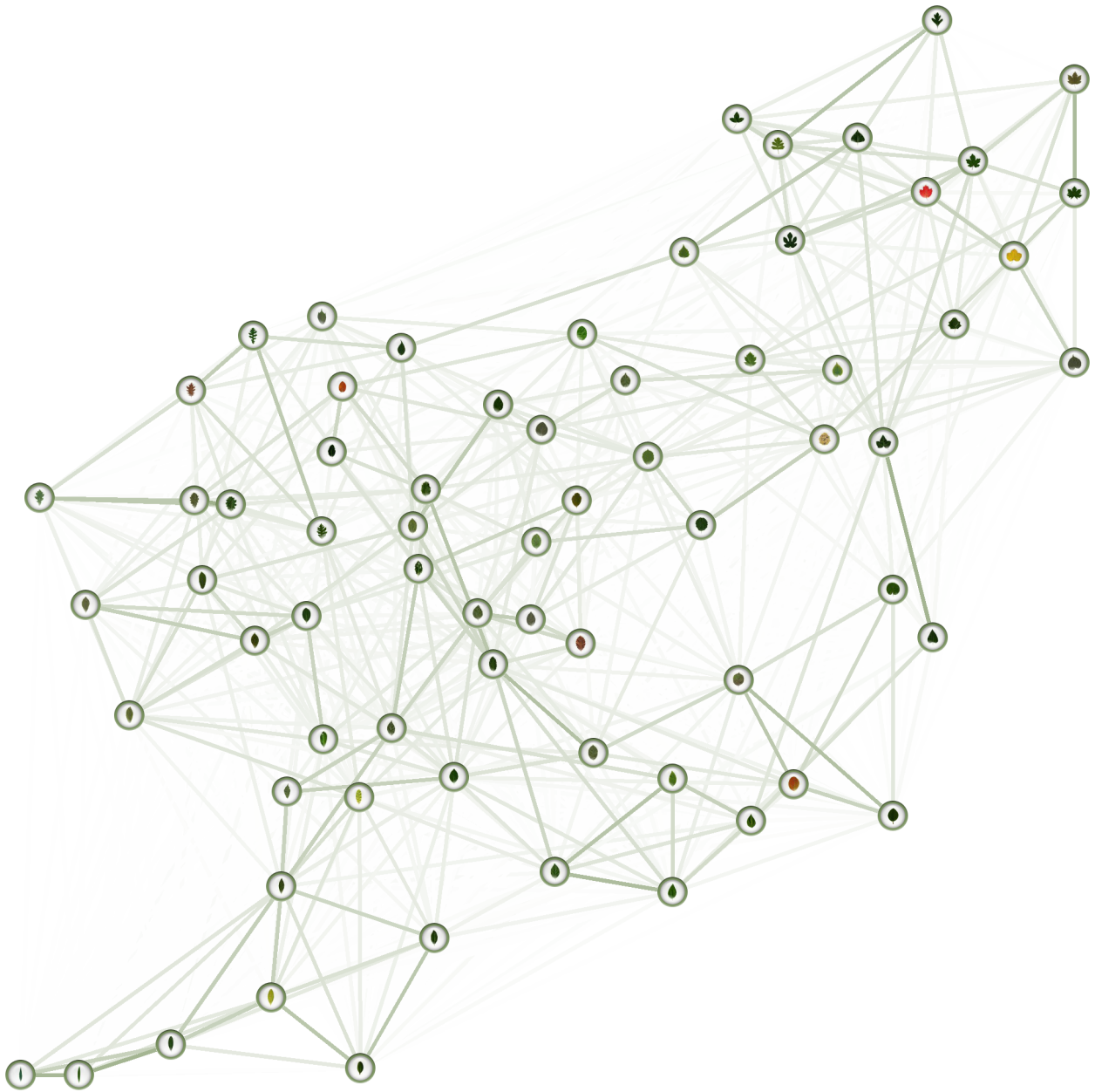


FIGURE 7.12 – Exemple de graphe représentant la proximité des espèces entre elles : plus une arête est sombre plus les espèces correspondent à des modèles similaires

La prise en compte de la géographie est une voie des plus prometteuses, et apparaît comme un domaine où il reste énormément de choses à explorer, tant du point de vue des paramètres influents que de la fusion de données aussi hétérogènes. Elle se heurte pour l'instant à un manque de données fiables, mais la distribution d'un système d'identification permettant de récupérer des observations géolocalisées se présente comme un moyen de remédier à ce problème. On touche ici au domaine de la science citoyenne, où la masse des utilisateurs peut contribuer à faire progresser les connaissances, en répertoriant les espèces, en mesurant l'évolution de leur répartition, ou bien même en enrichissant les données utilisées pour concevoir le système d'identification.

Les méthodes développées dans cette thèse pour la reconnaissance des espèces se sont focalisées sur les feuilles, mais on peut sans problème imaginer leur transposition aux autres organes des plantes (fleurs, fruits, écorces) en utilisant certes des modèles et des descripteurs différents, mais avec le même schéma de fonctionnement et la même utilisation de la botanique comme référence. L'identification serait alors multimodale, avec de nouvelles questions de fusion de sources d'information qui devraient être traitées, chose que nous avons déjà commencé à entreprendre [CTS⁺13]. Les problèmes ainsi posés sont nombreux, tant du point de vue de la segmentation, de la possibilité d'une modélisation, que de la description d'objets bien différents, mais ils constitueraient une voie passionnante d'extension de nos méthodes de reconnaissance.

Par son implémentation concrète, le système présenté tout au long de cette thèse constitue le cœur d'une application originale d'aide à l'identification des espèces d'arbres, destinée à quiconque se sentant un attrait pour le monde végétal sans pour autant maîtriser les arcanes de la botanique. Elle se propose de couvrir l'essentiel de la diversité des essences forestières natives en Europe, et a pour but d'être fonctionnelle dans un cadre d'utilisation réel en environnement naturel. Plus qu'un simple outil d'identification, le système se veut un témoin des mécanismes à l'œuvre pour la détermination des espèces, avec l'idée directrice que la transparence sémantique doit pouvoir bénéficier à la compréhension de l'utilisateur. Le but est en tout cas atteint : entrouvrir la porte vers l'univers de la botanique, comme un moyen intuitif de transmettre une compétence qui se fait aujourd'hui rare, et de satisfaire dans le même temps le désir intemporel qu'a tout un chacun de comprendre la Nature.

Bibliographie

- [CAT⁺12] Guillaume Cerutti, Violaine Antoine, Laure Tougne, Julien Mille, Lionel Valet, Didier Coquin, and Antoine Vacavant. ReVeS Participation - Tree Species Classification Using Random Forests and Botanical Features. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [CTCV13] Guillaume Cerutti, Laure Tougne, Didier Coquin, and Antoine Vacavant. Curvature-scale-based contour understanding for leaf margin shape recognition and species identification. In *VISAPP 2013 - Proceedings of the International Conference on Computer Vision Theory and Applications*, 2013.
- [CTM⁺11] Guillaume Cerutti, Laure Tougne, Julien Mille, Antoine Vacavant, and Didier Coquin. Guiding active contours for tree leaf segmentation and identification. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
- [CTM⁺13a] Guillaume Cerutti, Laure Tougne, Julie Mille, Antoine Vacavant, and Didier Coquin. A model-based approach for compound leaves understanding and identification. In *IEEE International Conference on Image Processing*, 2013.
- [CTM⁺13b] Guillaume Cerutti, Laure Tougne, Julien Mille, Antoine Vacavant, and Didier Coquin. Understanding leaves in natural images - A model-based approach for tree species identification. *Computer Vision and Image Understanding*, 2013.
- [CTS⁺13] Guillaume Cerutti, Laure Tougne, Céline Sacca, Thierry Joliveau, Pierre-Olivier Mazagol, Didier Coquin, and Antoine Vacavant. Late information fusion for multi-modality plant species identification. In *CLEF 2013 Working Notes*, 2013.
- [CTVC11] Guillaume Cerutti, Laure Tougne, Antoine Vacavant, and Didier Coquin. A parametric active polygon for leaf segmentation and shape estimation. In *7th International Symposium on Visual Computing*, 2011.
- [GBJ⁺11] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The CLEF 2011 Plant Images Classification Task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.

- [GBJ⁺12] H. Goëau, P. Bonnet, A. Joly, D. Barthelemy, N. Boujemaa, and J.-F. Molino. The ImageCLEF 2012 Plant Image Identification Task. In *ImageCLEF 2012 Working Notes*, 2012.
- [GBJ⁺13] H. Goëau, P. Bonnet, A. Joly, V. Bakic, D. Barthelemy, N. Boujemaa, and J.-F. Molino. The ImageCLEF 2013 Plant Identification Task. In *CLEF 2013 Working Notes*, 2013.
- [GBVC⁺13] Manuel Grand-Brochier, Antoine Vacavant, Guillaume Cerutti, Kevin Bianchi, and Laure Tougne. Comparative study of segmentation methods for tree leaves extraction. In *Proceedings of the International Workshop on Video and Image Ground Truth in Computer Vision Applications*, page 7. ACM, 2013.
- [LCV13] HongHui Liu, Didier Coquin, and Lionel Valet. Reconnaissance de feuilles d'arbres par fusion de décisions partielles. In *GRETSI*, 2013.

Annexe A

Base de données botaniques

Liste des espèces retenues et formes foliaires











Salix alba L.

Résumé du Chapitre

Cette annexe donne la liste des espèces que nous avons retenues comme devant figurer parmi celles reconnues par une application d'identification des espèces natives d'arbres, et pour lesquelles nous avons pu disposer de suffisamment de données pour les inclure dans notre système de reconnaissance. Chaque espèce y est ainsi décrite par ses noms binomial et vernaculaire, ainsi que par ses différents rangs taxinomiques selon la classification APG III. En se basant sur les données tirées de flores de références, nous avons effectué la liste des propriétés de forme des feuilles de chacune espèce, décrites par le vocabulaire botanique, liste que nous avons utilisée comme référence dans nos travaux.




















A. Base de données botaniques

 <p><i>Acer campestre</i> L.</p>	<p>Erable champêtre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Palmatilobée  Sinuée  Lobée  Subaiguë  Obtuse  En Coeur
 <p><i>Acer monspessulanum</i> L.</p>	<p>Erable de Montpellier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Palmatilobée  Entière  Obtuse  En Coeur
 <p><i>Acer negundo</i> L.</p>	<p>Erable negundo</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Imparipennée




















A. Base de données botaniques

 <p><i>Acer opalus</i> Mill.</p>	<p>Erable à feuilles d'obier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Palmatilobée  Sinuée  Obtuse  En Coeur
 <p><i>Acer platanoides</i> L.</p>	<p>Erable plane</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Palmatilobée  Dentée  Acuminée  Aiguë  En Coeur
 <p><i>Acer pseudoplatanus</i> L.</p>	<p>Erable sycomore</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Sapindaceae</i> Genre <i>Acer</i></p>	 Palmatilobée  Dentée  Courtement Acuminée  En Coeur


















A. Base de données botaniques

 <p><i>Alnus glutinosa</i> (L.) Gaetrn.</p>	<p>Aulne glutineux</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Betulaceae</i> Genre <i>Alnus</i></p>	 Suborbiculaire  Obovale  Sinuée  Dentée  Tronquée  Echancrée  En Coin
 <p><i>Amelanchier ovalis</i> Med.</p>	<p>Amélanchier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Amelanchier</i></p>	 Ovale  Denticulée  Arrondie
 <p><i>Arbutus unedo</i> L.</p>	<p>Arbousier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Ordre <i>Ericales</i> Famille <i>Ericaceae</i> Genre <i>Arbutus</i></p>	 Ovale  Oblongue  Obovale  Dentée en Scie  Obtuse  En Coin

A. Base de données botaniques

 <p><i>Betula pendula</i> Roth</p>	<p>Bouleau verruqueux</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Betulaceae</i> Genre <i>Betula</i></p>	 Triangulaire  Rhomboïdale  Ovale  Doublement Dentée  Acuminée
 <p><i>Betula pubescens</i> Ehrh.</p>	<p>Bouleau pubescent</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Betulaceae</i> Genre <i>Betula</i></p>	 Ovale  Rhomboïdale  Dentée  Aiguë  Courtement Acuminée
 <p><i>Buxus sempervirens</i> L.</p>	<p>Buis</p>
<p>Clade <i>Eudicotylédones</i> Ordre <i>Buxales</i> Famille <i>Buxaceae</i> Genre <i>Buxus</i></p>	 Ovale  Lancéolée  Entière  Arrondie  Echancrée  Tronquée

















A. Base de données botaniques

 <p><i>Capparis spinosa</i> L.</p>	<p>Câprier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Brassicales</i> Famille <i>Capparaceae</i> Genre <i>Capparis</i></p>	 Ovale  Orbiculaire  Entière  Arrondie  Obtuse
 <p><i>Carpinus betulus</i> L.</p>	<p>Charme</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Betulaceae</i> Genre <i>Carpinus</i></p>	 Ovale  Oblongue  Doublement Dentée  Acuminée
 <p><i>Castanea sativa</i> Mill.</p>	<p>Châtaignier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Castanea</i></p>	 Lancéolée  Oblongue  Obovale  Dentée  Cuspidée






















A. Base de données botaniques

 <p><i>Celtis australis L.</i></p>	<p>Micocoulier de Provence</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Cannabaceae</i> Genre <i>Celtis</i></p>	 Ovale  Lancéolée  Dentée en Scie  Dentée  Acuminée  Inégale
 <p><i>Cercis siliquastrum L.</i></p>	<p>Arbre de Judée</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fabales</i> Famille <i>Fabaceae</i> Genre <i>Cercis</i></p>	 Cordée  Orbiculaire  Entière  Arrondie  Echancrée  En Coeur  Echancrée
 <p><i>Colutea arborescens L.</i></p>	<p>Baguenaudier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fabales</i> Famille <i>Fabaceae</i> Genre <i>Colutea</i></p>	 Imparipennée  Entière  Obcordée









A. Base de données botaniques

 <p><i>Cornus mas L.</i></p>	<p>Cornouiller mâle</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Ordre <i>Cornales</i> Famille <i>Cornaceae</i> Genre <i>Cornus</i></p>	 Ovale  Entière  Aiguë  Acuminée  En Coin
 <p><i>Cornus sanguinea L.</i></p>	<p>Cornouiller sanguin</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Ordre <i>Cornales</i> Famille <i>Cornaceae</i> Genre <i>Cornus</i></p>	 Ovale  Elliptique  Entière  Ondulée  Courtement Acuminée
 <p><i>Corylus avellana L.</i></p>	<p>Noisetier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Betulaceae</i> Genre <i>Corylus</i></p>	 Suborbiculaire  Doublement Dentée  Acuminée  En Coeur



A. Base de données botaniques

 <p><i>Cotinus coggygria</i> Scop.</p>	<p>Arbre à perruques</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Anacardiaceae</i> Genre <i>Cotinus</i></p>	 Suborbiculaire  Obovale  Entière  Obtuse  Arrondie  En Coin
 <p><i>Crataegus azarolus</i> L.</p>	<p>Azarolier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Crataegus</i></p>	 Pennatilobée  Obovale  Pennatifide  Dentée  Lobée  Mucronée  En Coin
 <p><i>Crataegus laevigata</i> (Poir.) DC.</p>	<p>Aubépine à deux styles</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Crataegus</i></p>	 Pennatilobée  Obovale  Dentée  Obtuse  En Coin











A. Base de données botaniques

 <p><i>Crataegus monogyna</i> Jacq.</p>	<p>Aubépine monogyne</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Crataegus</i></p>	 Pennatilobée  Ovale  Pennatifide  Dentée  Lobée  En Coin
 <p><i>Euonymus europaeus</i> L.</p>	<p>Fusain d'Europe</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Celastrales</i> Famille <i>Celastraceae</i> Genre <i>Euonymus</i></p>	 Ovale  Lancéolée  Denticulée  Acuminée
 <p><i>Euonymus latifolius</i> (L.) Mill.</p>	<p>Fusain à larges feuilles</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Celastrales</i> Famille <i>Celastraceae</i> Genre <i>Euonymus</i></p>	 Ovale  Oblongue  Denticulée  Acuminée




















A. Base de données botaniques

 <p><i>Fagus sylvatica</i> L.</p>	<p>Hêtre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Fagus</i></p>	 Ovale  Entière  Ondulée  Ciliée  Aiguë  En Coin
 <p><i>Ficus carica</i> L.</p>	<p>Figuier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Moraceae</i> Genre <i>Ficus</i></p>	 Palmatilobée  Triangulaire  Crénelée  Dentée  Obtuse  En Coeur
 <p><i>Frangula alnus</i> Mill.</p>	<p>Bourdaine</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rhamnaceae</i> Genre <i>Frangula</i></p>	 Ovale  Elliptique  Entière  Ondulée  Aiguë









A. Base de données botaniques

 <p><i>Fraxinus angustifolia</i> Vahl</p>	<p>Frêne oxyphylle</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Fraxinus</i></p>	 <p>Imparipennée</p>	
 <p><i>Fraxinus excelsior</i> L.</p>	<p>Frêne commun</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Fraxinus</i></p>	 <p>Imparipennée</p>  <p>Dentée en Scie</p>	 <p>Acuminée</p>
 <p><i>Fraxinus ornus</i> L.</p>	<p>Frêne à fleurs</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Fraxinus</i></p>	 <p>Imparipennée</p>  <p>Dentée</p>	 <p>Acuminée</p>




















A. Base de données botaniques

 <p><i>Hedera helix</i> L.</p>	<p>Lierre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Apiales</i> Famille <i>Araliaceae</i> Genre <i>Hedera</i></p>	 Palmatilobée  Ovale  Entière  Acuminée  Obtuse
 <p><i>Ilex aquifolium</i> L.</p>	<p>Houx</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Aquifoliales</i> Famille <i>Aquifoliaceae</i> Genre <i>Ilex</i></p>	 Ovale  Elliptique  Epineuse  Entière  Dentée  Sinuée  Aiguë  Epineuse
 <p><i>Laburnum anagyroides</i> Medik.</p>	<p>Cytise aubour</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fabales</i> Famille <i>Fabaceae</i> Genre <i>Laburnum</i></p>	 Trifoliée  Entière  Obtuse



A. Base de données botaniques

 <p><i>Laurus nobilis L.</i></p>	<p>Laurier sauce</p>
<p>Clade <i>Magnoliidae</i> Ordre <i>Laurales</i> Famille <i>Lauraceae</i> Genre <i>Laurus</i></p>	 Lancéolée  Elliptique  Entière  Ondulée  Aiguë  Acuminée  Atténuée
 <p><i>Ligustrum vulgare L.</i></p>	<p>Troène</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Ligustrum</i></p>	 Lancéolée  Elliptique  Entière  Obtuse  En Coin
 <p><i>Malus sylvestris Mill.</i></p>	<p>Pommier sauvage</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Malus</i></p>	 Ovale  Suborbiculaire  Denticulée  Crénelée  Acuminée

A. Base de données botaniques

 <p><i>Mespilus germanica L.</i></p>	<p>Néflier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Mespilus</i></p>	 Oblongue  Entière  Denticulée  Aiguë
 <p><i>Nerium oleander L.</i></p>	<p>Laurier-rose</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Gentianales</i> Famille <i>Apocynaceae</i> Genre <i>Nerium</i></p>	 Lancéolée  Entière  Aiguë  En Coin
 <p><i>Olea europaea L.</i></p>	<p>Olivier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Olea</i></p>	 Ovale  Oblongue  Lancéolée  Entière  Aiguë  Obtuse  Mucronée  Atténuée



















A. Base de données botaniques

 <p><i>Paliurus spina-christi</i> Mill.</p>	<p>Paliure</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rhamnaceae</i> Genre <i>Paliurus</i></p>	 Ovale  Obtuse  Entière  Oblique  Denticulée
 <p><i>Pistacia lentiscus</i> L.</p>	<p>Pistachier lentisque</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Anacardiaceae</i> Genre <i>Pistacia</i></p>	 Paripennée  Mucronée  Entière  Arrondie
 <p><i>Pistacia terebinthus</i> L.</p>	<p>Pistachier térébinthe</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Anacardiaceae</i> Genre <i>Pistacia</i></p>	 Imparipennée  Mucronée  Entière  Obtuse


















A. Base de données botaniques

 <p><i>Populus alba</i> L.</p>	<p>Peuplier blanc</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Populus</i></p>	 Suborbiculaire  Palmatilobée  Obtuse  Sinuée  Anguleuse  Lobée
 <p><i>Populus nigra</i> L.</p>	<p>Peuplier noir</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Populus</i></p>	 Triangulaire  Acuminée  Dentée  Crénelée
 <p><i>Populus tremula</i></p>	<p>Tremble</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Populus</i></p>	 Suborbiculaire  Aiguë  Dentée  Sinuée


















A. Base de données botaniques

 <i>Prunus avium</i> L.	Merisier
Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i>	 Elliptique  Obovale  Acuminée  Doublement Dentée
 <i>Prunus laurocerasus</i> L.	Laurier-cerise
Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i>	 Lancéolée  Oblongue  Denticulée  Aiguë  Acuminée  En Coin
 <i>Prunus mahaleb</i> L.	Cerisier de Sainte-Lucie
Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i>	 Ovale  Dentée  Crénelée  Courtement Acuminée  En Coeur























A. Base de données botaniques

 <p><i>Prunus padus</i> L.</p>	<p>Cerisier à grappes</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i></p>	 Obovale  Acuminée  Denticulée
 <p><i>Prunus serotina</i> Ehrh.</p>	<p>Cerisier d'automne</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i></p>	 Oblongue  Elliptique  Dentée en Scie  Acuminée  Aiguë
 <p><i>Prunus spinosa</i> L.</p>	<p>Prunellier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Prunus</i></p>	 Oblongue  Obovale  Lancéolée  Finement Dentée  Acuminée  En Coin





















A. Base de données botaniques

 <p><i>Punica granatum L.</i></p>	<p>Grenadier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Myrtales</i> Famille <i>Lythraceae</i> Genre <i>Punica</i></p>	 Oblongue  Entière  Obtuse  Arrondie  En Coin
 <p><i>Pyrus pyraster (L.) Du Roi</i></p>	<p>Poirier sauvage</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Pyrus</i></p>	 Ovale  Suborbiculaire  Dentée  Arrondie
 <p><i>Quercus cerris L.</i></p>	<p>Chêne chevelu</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Obovale  Oblongue  Sinuée  Pennatifide  Aiguë

A. Base de données botaniques

 <p><i>Quercus coccifera</i> L.</p>	<p>Chêne kermès</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Ovale  Oblongue  Epineuse  Entière  Epineuse  Aiguë  Tronquée
 <p><i>Quercus ilex</i> L.</p>	<p>Chêne vert</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Ovale  Oblongue  Entière  Epineuse  Dentée  Aiguë  Epineuse
 <p><i>Quercus petraea</i> Liebl.</p>	<p>Chêne rouvre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Obovale  Sinuée  Lobée  Arrondie  En Coin



A. Base de données botaniques

 <p><i>Quercus pubescens</i> Willd.</p>	<p>Chêne pubescent</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	<p>  Obovale  Sinuée  Lobée  Pennatifide  Arrondie  Tronquée </p>
 <p><i>Quercus pyrenaica</i> Willd.</p>	<p>Chêne Tauzin</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	<p>  Obovale  Oblongue  Sinuée  Lobée  Pennatifide  Tronquée </p>
 <p><i>Quercus robur</i> L.</p>	<p>Chêne pédonculé</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	<p>  Obovale  Oblongue  Sinuée  Lobée  Auriculée </p>

A. Base de données botaniques

 <p><i>Quercus rubra</i> L.</p>	<p>Chêne rouge</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Obovale  Lobée  Epineuse	
 <p><i>Quercus suber</i> L.</p>	<p>Chêne-liège</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fagales</i> Famille <i>Fagaceae</i> Genre <i>Quercus</i></p>	 Ovale  Oblongue  Entière  Dentée  Epineuse	
 <p><i>Rhamnus alaternus</i> L.</p>	<p>Nerprun alaterne</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rhamnaceae</i> Genre <i>Rhamnus</i></p>	 Ovale  Lancéolée  Dentée  Cuspidée  En Coin  Atténuée	








A. Base de données botaniques

 <p><i>Rhamnus cathartica</i> L.</p>	<p>Nerprun cathartique</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rhamnaceae</i> Genre <i>Rhamnus</i></p>	 Ovale  Elliptique  Denticulée  Acuminée  En Coin
 <p><i>Rhus coriaria</i> L.</p>	<p>Sumac des corroyeurs</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Sapindales</i> Famille <i>Anacardiaceae</i> Genre <i>Rhus</i></p>	 Imparipennée  Dentée
 <p><i>Robinia pseudoacacia</i> L.</p>	<p>Robinier faux-acacia</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Fabales</i> Famille <i>Fabaceae</i> Genre <i>Robinia</i></p>	 Imparipennée  Entière  Arrondie  Echancrée

A. Base de données botaniques

 <p><i>Ruscus aculeatus L.</i></p>	<p>Fragon faux-houx</p>
<p>Clade <i>Monocotylédones</i> Ordre <i>Asparagales</i> Famille <i>Asparagaceae</i> Genre <i>Ruscus</i></p>	 Ovale  Lancéolée  Entière  Epineuse
 <p><i>Salix alba L.</i></p>	<p>Saule blanc</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Salix</i></p>	 Lancéolée  Denticulée  Acuminée
 <p><i>Salix caprea L.</i></p>	<p>Saule marsault</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Salix</i></p>	 Oblongue  Obovale  Ovale  Entière  Crénelée  Obtuse  Recourbée

A. Base de données botaniques

 <p><i>Salix cinerea</i> L.</p>	<p>Saule cendré</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Salix</i></p>	
 <p><i>Salix fragilis</i> L.</p>	<p>Saule cassant</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Salix</i></p>	
 <p><i>Salix viminalis</i> L.</p>	<p>Saule des vanniers</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Malpighiales</i> Famille <i>Salicaceae</i> Genre <i>Salix</i></p>	 Lancéolée  Entière  Ondulée  Acuminée












A. Base de données botaniques

 <p><i>Sambucus nigra</i> L.</p>	<p>Sureau noir</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Dipsacales</i> Famille <i>Adoxaceae</i> Genre <i>Sambucus</i></p>	 Digitée  Acuminée  Dentée
 <p><i>Sorbus aria</i> (L.) Crantz</p>	<p>Alisier blanc</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Sorbus</i></p>	 Ovale  Elliptique  Dentée  Doublement Dentée  Arrondie  Obtuse
 <p><i>Sorbus aucuparia</i> L.</p>	<p>Sorbier des oiseleurs</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Sorbus</i></p>	 Imparipennée  Dentée









A. Base de données botaniques

 <p><i>Sorbus domestica</i> L.</p>	<p>Sorbier domestique</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Sorbus</i></p>	 Imparipennée  Dentée	
 <p><i>Sorbus torminalis</i> (L.) Crantz</p>	<p>Alisier des bois</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Rosaceae</i> Genre <i>Sorbus</i></p>	 Pennatilobée  Ovale  Lobée  Doublement Dentée  Acuminée  Tronquée  En Coeur	
 <p><i>Syringa vulgaris</i> L.</p>	<p>Lilas des jardins</p>	
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Oleaceae</i> Genre <i>Syringa</i></p>	 Cordée  Ovale  Triangulaire  Entière  Acuminée  En Coeur	










A. Base de données botaniques

 <p><i>Tilia cordata</i> Mill.</p>	<p>Tilleul des bois</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Malvales</i> Famille <i>Malvaceae</i> Genre <i>Tilia</i></p>	 Cordée  Suborbiculaire  Dentée en Scie  Acuminée  Obliquement En Coeur
 <p><i>Tilia platyphyllos</i> Scop.</p>	<p>Tilleul à grandes feuilles</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Malvidae</i> Ordre <i>Malvales</i> Famille <i>Malvaceae</i> Genre <i>Tilia</i></p>	 Cordée  Ovale  Dentée en Scie  Acuminée  Obliquement En Coeur
 <p><i>Ulmus glabra</i> Huds.</p>	<p>Orme de montagne</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Ulmaceae</i> Genre <i>Ulmus</i></p>	 Obovale  Doublement Dentée  Acuminée  Inégale

A. Base de données botaniques

 <p><i>Ulmus minor</i> Mill.</p>	<p>Orme champêtre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Rosidae</i> Clade <i>Fabidae</i> Ordre <i>Rosales</i> Famille <i>Ulmaceae</i> Genre <i>Ulmus</i></p>	 Ovale  Dentée  Doublement Dentée  Courtement Acuminée  Aiguë  Inégale
 <p><i>Viburnum lantana</i> L.</p>	<p>Viorne lantane</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Dipsacales</i> Famille <i>Adoxaceae</i> Genre <i>Viburnum</i></p>	 Ovale  Suborbiculaire  Dentée  Arrondie  En Coeur
 <p><i>Viburnum opulus</i> L.</p>	<p>Viorne obier</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Dipsacales</i> Famille <i>Adoxaceae</i> Genre <i>Viburnum</i></p>	 Palmatilobée  Lobée  Dentée  Aiguë

A. Base de données botaniques

 <p><i>Viburnum tinus L.</i></p>	<p>Viorne-tin</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Campanulidae</i> Ordre <i>Dipsacales</i> Famille <i>Adoxaceae</i> Genre <i>Viburnum</i></p>	 Ovale  Oblongue  Entière  Obtuse  Aiguë
 <p><i>Vitex agnus-castus L.</i></p>	<p>Arbre au poivre</p>
<p>Clade <i>Eudicotylédones</i> Clade <i>Asteridae</i> Clade <i>Lamiidae</i> Ordre <i>Lamiales</i> Famille <i>Lamiaceae</i> Genre <i>Vitex</i></p>	 Digitée  Entière

Annexe B

L'application Folia pour iPhone

Réalisations et déploiement sur une interface smartphone



Ulmus minor Mill.

Résumé du Chapitre

Cette annexe présente l'implémentation des méthodes décrites dans cette thèse sous la forme d'une application pour iPhone. Les algorithmes d'analyse d'image et de classification développés au laboratoire ont été repris quasiment tels quels et inclus dans une structure d'application permettant leur mise à jour fréquente. Leur mise en place au sein d'une interface smartphone a demandé quelques adaptations ainsi qu'une définition plus ordonnée du processus de reconnaissance. La réalisation que nous présentons est l'application Folia d'aide à l'identification des végétaux, qui concrétise nos ambitions de processus interactif et explicatif de reconnaissance des espèces d'arbres.

Ce Chapitre présente des réalisations développées en partenariat avec Thomas BARTHOULOT, de l'entreprise Trydea, prestataire du projet.

NOUR véritablement donner un sens aux méthodes développées au cours de cette thèse, il était indispensable de les mettre en pratique dans le cadre d'une application susceptible d'être utilisée par le grand public. Le but est d'essayer à nouveau une forme de connaissance botanique par un processus explicatif de reconnaissance, au travers de représentations visuellement éloquentes et du lien qui est fait avec les espèces végétales. Cela passe par l'implémentation concrète des méthodes proposées spécifiquement dans ce but, au sein d'un outil interactif et intuitif qui devrait être à même de se rapprocher de cet objectif : l'application Folia *

B.1 Modèle de développement de l'application

L'application a été développée pour les plateformes iPhone et iPad dans un premier temps, choix défini dès le démarrage du projet, offrant en prime la possibilité séduisante de pouvoir être implémentée sans recodage des algorithmes de reconnaissance d'images. Cela passe par l'utilisation de bibliothèques précompilées, et facilite grandement le développement de l'application en faisant de la partie traitement une brique interchangeable, à condition de structurer le tout de façon appropriée.

B.1.1 Architecture générale

Le paradigme choisi pour implémenter l'application est le patron **Modèle-Vue-Contrôleur** (MVC) qui a pour but de découpler toute la partie de visualisation et d'interaction avec l'utilisateur (la composante **Vue**) du cœur de l'algorithme de traitement (la composante **Modèle**) par le biais d'un composant faisant l'interface entre les deux (la composante **Contrôleur**) qui transmet les signaux provenant de la Vue sous forme de consignes pour modifier l'état du Modèle. L'organisation générale de ce patron est présentée dans la Figure B.1.

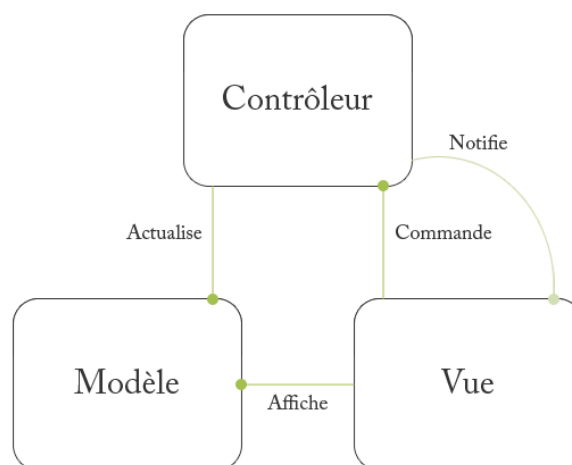


FIGURE B.1 – Représentation du patron Modèle-Vue-Contrôleur dans un cas général

*. *folia* : application gratuite pour iPhone et iPad, disponible sur l'App Store <https://itunes.apple.com/app/folia/id547650203>

Dans notre cadre particulier, les tâches allouées aux différentes composantes respectent cette séparation, de façon à ce que l'implémentation de l'interface graphique et des mécanismes d'interaction puisse se faire indépendamment ou presque de celle du système de reconnaissance. Elles sont donc globalement réparties de la façon suivante :

- ▶ La **Vue** fait intervenir l'utilisateur, et se trouve donc responsable des interactions : transmission de la photographie de la feuille, définition d'une région initiale, affichage des images et des résultats, validation des étapes du processus.
- ▶ Le **Contrôleur** a essentiellement un rôle d'interface, qui communique avec le système de traitement des images, en lui fournissant les données et en lui indiquant quand passer à l'étape suivante, et qui avertit la Vue des changements à visualiser.
- ▶ Le **Modèle** correspond au processus de reconnaissance qui a été développé, sans se préoccuper de la visualisation, et doit simplement être structuré pour pouvoir être appelé par le Contrôleur, ce qui est fait par un découpage du traitement en une séquence d'étapes.

Avec ce schéma, la partie traitement est entièrement découplée du reste de l'application, et peut donc parfaitement être modifiée sans nécessiter des changements coûteux dans le reste des composantes. C'est donc un modèle qui s'adapte particulièrement bien à un contexte de recherche enclin à de constantes améliorations et nouvelles voies à explorer, qui peuvent alors être implémentées très rapidement dans une application concrète.

B.1.2 Déroulement du processus

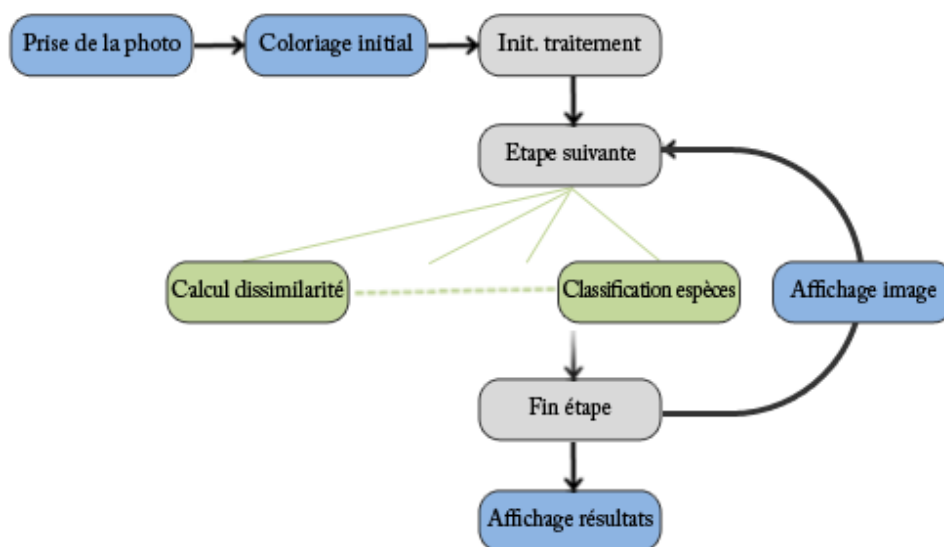


FIGURE B.2 – Déroulement général du processus de traitement d'une image de feuille par l'application : étapes réalisées ■ par la Vue, ■ par le Contrôleur et ■ par le Modèle

Le scénario d'utilisation est des plus simples, il débute par la prise en photo d'une feuille d'arbre, guidée de façon à obtenir une image correspondant au cadre de fonctionnement de nos méthodes, puis par l'acquisition d'une trace laissée par l'utilisateur qui sert de point de départ à l'algorithme de segmentation. A partir de ces données d'entrée, le processus de reconnaissance s'effectue dans l'ordre : segmentation de l'image, extraction des descripteurs et classification en une liste d'espèces. C'est ce classement qui est présenté à l'utilisateur comme résultat de l'identification. La Figure B.2 détaille l'implémentation de ce scénario suivant le schéma MVC, en identifiant les tâches qui vont relever de chacune des composantes.

Le cœur du traitement est donc l'appel successif par le Contrôleur aux différentes étapes du processus de traitement de l'image, qui se fait simplement par l'appel à l'étape suivante déterminée quant à elle par le Modèle. Ce schéma suppose que la reconnaissance soit bien structurée en étapes indépendantes livrant chacune un résultat bien identifié et pouvant être visualisé par la Vue. Cette séquence d'étapes va en fait être différente selon que la feuille considérée est simple ou composée, une distinction qui est effectuée par l'utilisateur, et transmise au Modèle par le Contrôleur. La Figure B.3 donne la séquence des étapes de traitement effectuées dans chacun de ces deux cas.

Pour les étapes correspondant à une évolution, notamment celle des modélisations déformables de formes foliaires évaluées sur l'image, le processus boucle jusqu'à ce que l'évolution soit terminée, permettant dans chaque cas d'afficher l'état courant de l'évolution à l'utilisateur. La visualisation est alors bien plus intéressante que si seul le résultat final de l'étape était affiché, d'autant plus que les temps de calcul sur mobile peuvent être relativement longs. Cet affichage permet de faire comprendre à l'utilisateur qu'il se passe bien quelque chose, tout en lui donnant un aperçu direct du déroulement du programme.

L'ajout de nouvelles étapes dans le traitement est tout à fait possible, et ce sans changements à apporter aux implémentations de la Vue et du Contrôleur. En effet, l'insertion d'une étape supplémentaire sera pour ce dernier transparente, puisque c'est au Modèle de déterminer quelle est l'étape suivante à un point donné du processus. C'est d'ailleurs ce que nous avons mis en pratique pour ajouter dans un second temps la chaîne de traitement des feuilles composées, après avoir implémenté nos algorithmes pour les feuilles simples.

Le schéma de développement ainsi défini nous a permis d'implémenter nos méthodes de reconnaissance de façon quasi-directe, en reprenant le code utilisé pour leur développement avec un simple effort de structuration. Une attention particulière a également dû être portée aux fuites mémoires, dont les conséquences sont bien plus rédhibitoires sur un smartphone, où la mémoire est une denrée limitée. Une fois ce travail fait, et les implémentations des autres composantes réalisées parallèlement, il aura suffi d'assembler le tout pour obtenir une application fonctionnelle, véritable concrétisation de nos travaux.

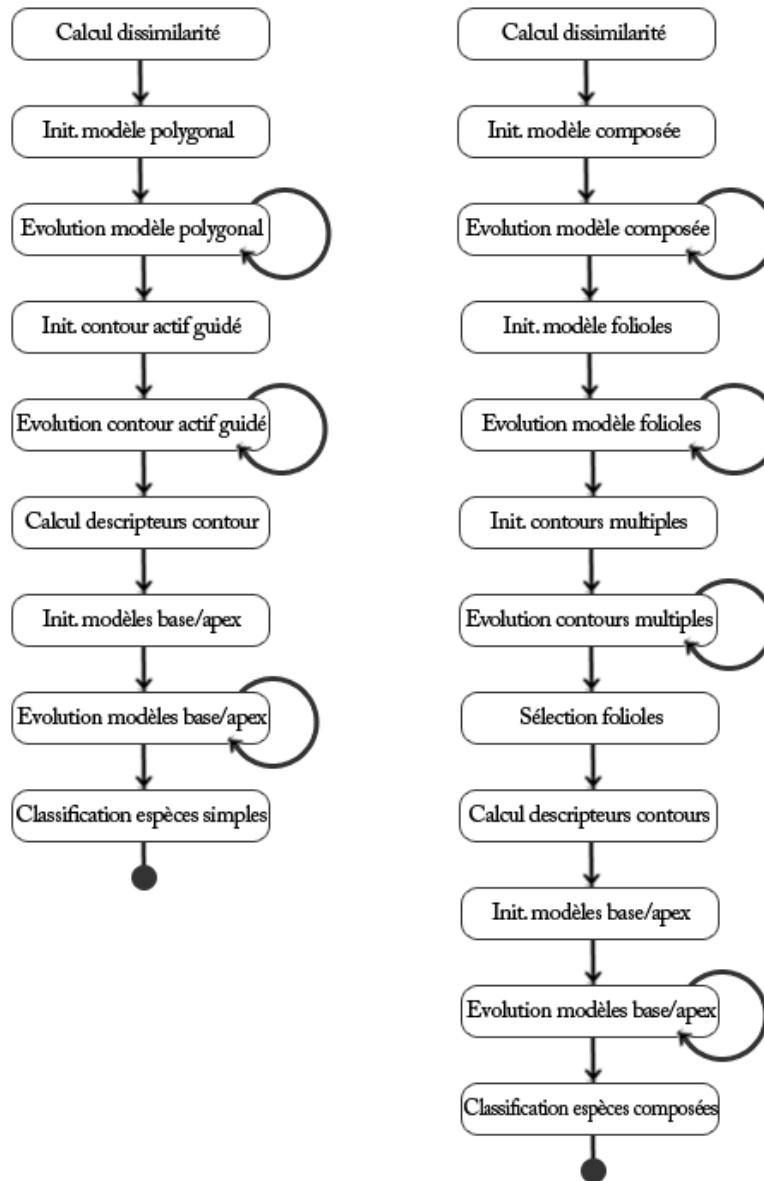


FIGURE B.3 – Séquences des étapes de traitement effectuées par le Modèle dans le cadre du processus de reconnaissance des feuilles simples (à gauche) et de celui des feuilles composées (à droite)

B.2 Présentation de *folia*

L'application Folia a pour but premier de constituer la mise en pratique de notre méthode de reconnaissance des feuilles d'arbres, mais elle vient avec quelques autres fonctionnalités qui rendent son utilisation plus enrichissante. L'écran d'accueil de l'application (Figure B.4 (a)) permet non seulement de lancer une identification, mais aussi d'accéder à la liste des espèces reconnues par l'application et à leurs fiches descriptives (Figure B.4 (b)) ainsi qu'à l'herbier où vont venir s'enregistrer toutes les feuilles rencontrées par l'utilisateur (Figure B.4 (c)).

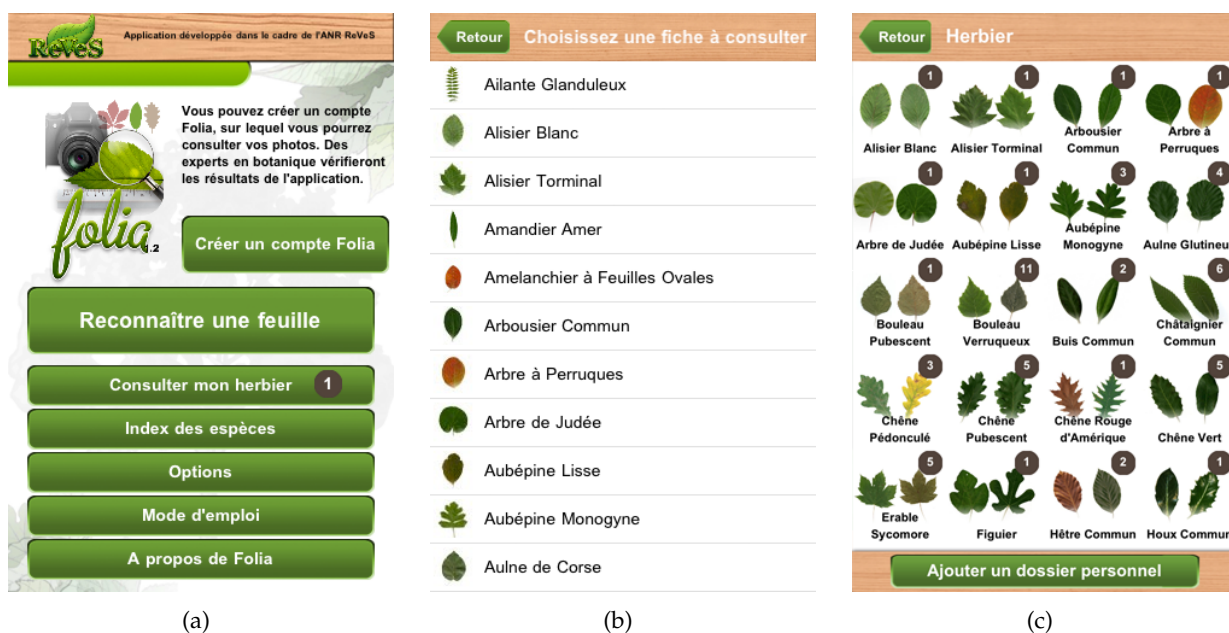


FIGURE B.4 – Visuels de l'application Folia : page d'accueil (a), index des espèces reconnues (b) et herbier de l'utilisateur (c)

Ces éléments viennent donner une dimension de collection et de découverte à l'application, et la font s'inscrire dans la durée plus que comme un gadget à utilisation unique. Ils ne sont toutefois là que pour étoffer ce qui est la raison d'être de l'application : le système d'identification.

B.2.1 Processus d'identification

La reconnaissance se fait de manière relativement intuitive pour l'utilisateur (il suffit de prendre sa photo dans le bon sens, et de trouver le coup de main pour un coloriage efficace) mais un mode d'emploi est à disposition pour le néophyte. Le point de départ est la prise de la photographie, qui se voit guidée par un masque figurant une forme de feuille (Figure B.5 (a)), de façon à ce que celle-ci soit correctement orientée dans l'image pour le bon déroulement de nos algorithmes. Vient ensuite


un écran permettant le choix entre feuille simple et composée (Figure B.5 (b)), une distinction peu évidente mais que nous tentons d'expliquer du mieux possible, et un écran permettant le coloriage (optionnel pour les feuilles simples) d'une région initiale (Figure B.5 (c)). Ces trois étapes sont les principales interactions demandées à l'utilisateur.



FIGURE B.5 – Interactions de l'utilisateur avec l'application Folia : prise de photo (a), choix du type de feuille (b) et coloriage de la région initiale (c)

En fonction des options choisies, et à partir des données en entrée, le processus de reconnaissance se déroule, et les images sont présentées à l'utilisateur de façon à ce qu'il puisse suivre l'évolution du traitement. Celles-ci correspondent exactement aux données manipulées par le système, et sont représentées sans altération. Par exemple sont affichées en temps réel les évolutions du modèle polygonal, du contour actif, ou des modèles de formes basales et apicales. Une fois atteinte la fin du processus un écran récapitulatif propose une visualisation des descripteurs extraits et qui seront utilisés pour la classification (Figure B.6 (a)). Cet écran permet une interprétation directe des valeurs des attributs et constitue un premier pas pour faire le lien entre les formes extraites par le système et les références botaniques.

Les descripteurs ainsi présentés sont ensuite utilisés pour la classification, et le résultat est présenté à l'utilisateur sous la forme d'une liste des espèces ordonnées par leur mesure de confiance (Figure B.6 (b)). Chacune de ces espèces donne accès à une fiche descriptive (Figure B.6 (c)), contenant d'une part des images en plus grande résolution de feuilles de l'espèce (images tirées des illustrations présentes sur le site de TelaBotanica[†]) et d'autres part des données théoriques permettant la description de la forme de ses feuilles.

†.  Tela Botanica : <http://www.tela-botanica.org>, réseau francophone de référence de la communauté botanique.

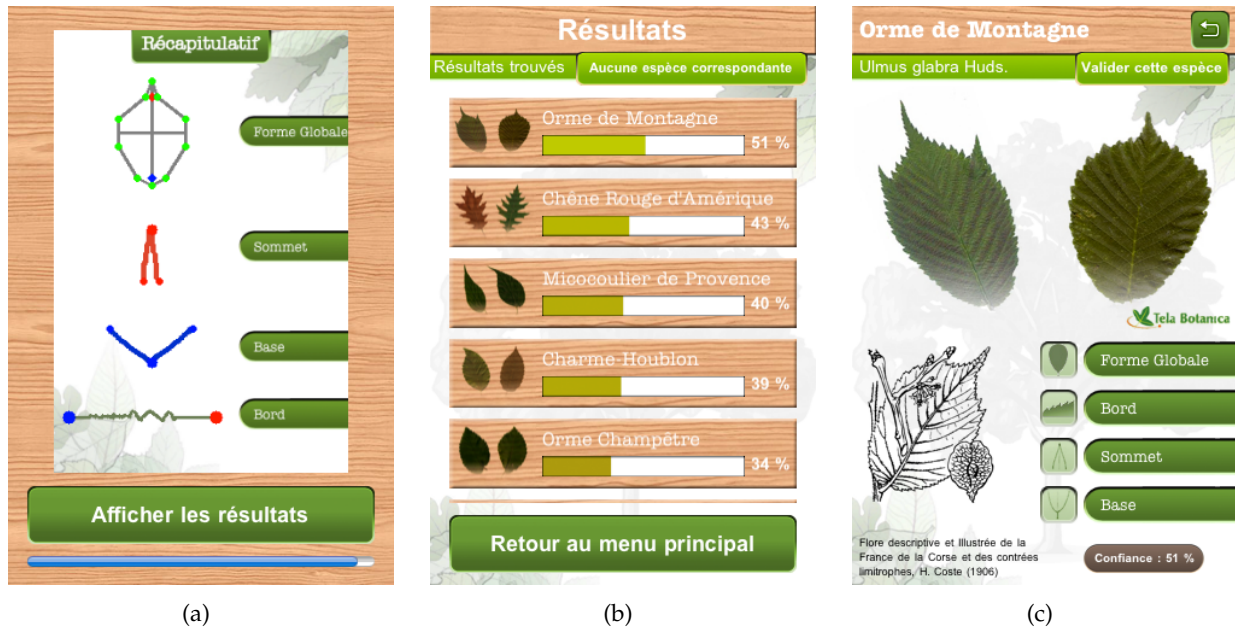


FIGURE B.6 – Identification par l'application Folia : récapitulatif des formes détectées (a), liste ordonnée d'espèces (b) et fiche descriptive d'une espèce (c)

L'objectif est de permettre, à ce stade, à l'utilisateur de finaliser l'identification en sélectionnant l'espèce qu'il considère la plus semblable. Les mesures de confiances renseignent théoriquement sur la fiabilité de la décision fournie par le système, et dans le cas idéal, c'est l'espèce correcte qui figure en tête du classement avec une confiance élevée. Mais dans le cas contraire, le fait que l'utilisateur recherche par lui-même la meilleure espèce candidate, en observant les similitudes selon différents critères, fait en sorte qu'il s'implique davantage dans le processus et intègre plus facilement les compétences d'identification que viserait à transmettre notre outil.

B.2.2 Fonctionnalités complémentaires

Après l'identification, un dernier écran est proposé à l'utilisateur, qui lui propose plusieurs options (Figure B.7), à commencer par la possibilité de conserver l'exemple dûment reconnu dans son herbier personnel. Les autres options sont tout autres, et sont en lien avec la dimension sociale et participative de l'application. Elles figurent ici pour permettre à l'utilisateur de mettre ses identifications en ligne, et, tout aussi important, de rendre ces données accessibles dans un contexte de recherche.

Folia se voit associée à un site web spécifique[‡], où chaque utilisateur peut ouvrir un compte via l'application. Les images dont la diffusion est autorisée sont transmises à ce site, où elles apparaissent sur le compte de leur photographe, et peuvent être rendues publiques. Cela ne se fait que si l'utilisateur y a donné son consentement, mais aussi que l'identification qu'il a effectuée a été validée par un expert.

‡. <http://liris.univ-lyon2.fr/rees/fofia>



FIGURE B.7 – Ecran final de l'identification permettant en particulier la diffusion des données

Les images ainsi diffusées sont accompagnées des diverses informations quant au fonctionnement du système, et en particulier de la position géographique donnée par le GPS du téléphone au moment de la prise de photo.

La diffusion de ces images permet donc petit à petit la constitution d'une base de données de photographies de feuilles, étiquetées par leur espèce une fois celle-ci validée, et géolocalisée par un couple longitude-latitude. La collecte de telles données prend du temps, d'autant que toutes les images ne sont pas d'une qualité égale, mais les informations réunies pourraient constituer une banque d'observations particulièrement intéressante pour de nombreuses activités de recherche, non seulement dans notre contexte de reconnaissance automatique d'images, mais aussi en botanique ou en biogéographie.

B.2.3 Perspectives d'extension

A l'heure actuelle, Folia constitue seulement la base de ce que pourrait être une application d'aide à l'identification des végétaux, puisqu'elle ne concerne qu'un nombre très limité d'espèces, présentes sur un territoire restreint et reconnaissables en utilisant un unique organe. Le fait d'utiliser d'autres organes simultanément pousse à revoir de façon trop complète le scénario d'utilisation de l'application (sans compter le changement de nom qui s'imposerait !) et nous nous contenterons donc de possibilités à court ou moyen terme.

Dans un premier temps, avec le développement de l'interface web, et la possibilité pour chaque utilisateur d'accéder à ses propres photos, il serait intéressant d'avoir une version en ligne du système

d'identification. Cela permettrait par exemple de relancer une reconnaissance sur une feuille qui a été mal identifiée, ou simplement mal orientée, ou encore d'expérimenter avec l'algorithme sur une plateforme moins contraignante.

L'apport de données image provenant des utilisateurs permet par ailleurs d'envisager un auto-enrichissement de la base d'exemples utilisés pour l'apprentissage des espèces. En incluant les exemples validés par les experts, et en entraînant à intervalle régulier les classifieurs utilisés par le système d'identification, il serait en effet possible d'obtenir des modèles plus complets, voire peut-être de reconnaître de nouvelles espèces, si des individus d'espèces pour lesquelles nous ne disposons pas encore de données image sont incidemment transmis et détectés par un expert.

Du point de vue de l'application, l'amélioration la plus notable serait l'inclusion de l'information géographique dans le traitement effectué sur le téléphone, à la manière de ce qui a été décrit au Chapitre 7. Cela pose des contraintes matérielles évidentes, tant il sera difficile d'embarquer la totalité des cartographies des différents paramètres sur un support mobile. En revanche, on peut espérer transmettre une carte locale, voire une valeur ponctuelle, de façon à pouvoir évaluer la plausibilité des espèces et affiner l'identification.

La représentation de l'information extraite de l'image par des formes récapitulatives est pour le moment assez sommaire, mais il est possible d'envisager une meilleure visualisation par une reconstruction plus explicite de la forme traitée. Le fait d'utiliser des modélisations génératives, qui offrent une traduction directe de leurs paramètres en forme, constitue un avantage dans ce contexte, et devrait permettre de recréer en fin de processus une représentation visuelle fidèle et de la présenter à l'utilisateur.

Une dernière extension serait l'ajout de l'étiquetage des formes reconnues par les concepts botaniques, comme introduit à la Section 6.2. En reportant également ces mêmes termes dans les fiches descriptives des espèces, l'établissement du lien entre les objets traités par le système de reconnaissance et les données théoriques serait cette fois-ci complet. La correspondance visuelle entre les formes et le vocabulaire pourrait alors véritablement se faire pour l'utilisateur, qui serait dès lors plus à même de se frotter aux ouvrages botaniques pour pousser plus loin son exploration.

