

Modélisation du Savoir-Faire des Experts pour l'Indexation de la Base Ornementale

M.-C. Larabi¹

N. Richard¹

O. Colot²

C. Fernandez-Maloigne¹

¹ IRCOM-SIC, UMR CNRS 6615, Université de Poitiers
Blvd M. & P. Curie, BP 30179 - 86962 FUTUROSCOPE Cedex

{larabi,richard,fernandez}@sic.univ-poitiers.fr

² I3D, FRE CNRS 2497, Université des Sciences et Technologies de Lille
59655 VILLENEUVE D'ASCQ Cedex

colot@i3d.univ-lille1.fr

Résumé

Dans cette contribution, nous nous sommes intéressés à la modélisation des connaissances et du savoir-faire des experts dans l'optique d'indexation par le contenu d'une base d'images de roches ornementales. Cette application est très intéressante pour l'aide à la classification des roches afin de former à la sortie des carrières, des lots homogènes ayant des couleurs et une texture similaires. Après une phase de détermination des besoins de l'application en termes d'attributs de bas-niveau, le savoir-faire des experts est modélisé grâce à une phase d'apprentissage spécifique. Celle-ci permet d'extraire pour chaque expert deux types d'informations : les facteurs de pondération permettant de donner plus d'importance à un attribut par rapport aux autres et les degrés de confiance reflétant l'exactitude de la classification d'un expert et par conséquent son expérience. La combinaison des différents facteurs de pondération et des degrés de confiance de tous les experts permet de générer un vecteur de pondération final efficace mêlant à la fois la capacité discriminatoire des attributs et le savoir-faire des experts.

Mots clefs

indexation, savoir-faire, gestion de connaissance, roches ornementales.

1 Introduction

L'émergence de l'Internet et la rapide expansion du marché de l'image ont permis à la recherche sur les méthodes de navigation dans les bases multimédias (images, vidéo,...) de s'inscrire comme l'une des voies les plus prometteuses de ces dernières décennies [1]. L'un des principaux axes est la recherche d'images par le contenu où l'image est représentée par un ensemble d'attributs de bas-niveau. Le but principal des systèmes d'indexation d'images par le contenu (SIIC) est de retrouver des images à partir des bases de données en se basant sur leur similarité par rapport à une image requête. Cette opération consiste en règle générale à comparer les descripteurs de bas-niveau

de l'image requête calculés en-ligne et ceux des images de la base ayant été calculés hors-ligne. Cependant, cette recherche ne permet pas forcément de faire une comparaison sémantique des images.

Le problème de la recherche d'images par le contenu est loin d'être résolu et beaucoup d'efforts de recherche restent à déployer pour atteindre une recherche s'appuyant cette fois-ci sur la sémantique de l'image (concept haut-niveau) et non uniquement sur une représentation générale du contenu par le biais de descripteurs bas-niveau.

Le problème traité dans cette contribution, s'inscrit dans le cadre de l'indexation des bases d'images spécialisées. Il consiste en l'aide à la classification des roches ornementales (marbre, granit,...). Le but étant d'aider les experts, dans un contexte industriel, à classer les roches ornementales grâce à une base de référence. Le schéma d'indexation répond bien aux besoins, car il permet de donner aux experts la possibilité de comparer leurs images avec une base de référence qui est commune à toute la profession. Les images dans ce système sont représentées par des attributs de couleur [2] et de texture formant ainsi leur signature. La mesure de similarité s'effectue en assignant à chaque attribut un poids pertinent permettant d'avoir une mesure cohérente. D'une part, les experts sont incapables de les définir à cause de leur faible connaissance des outils de l'image. D'autre part, nous sommes loin de disposer du savoir-faire de l'expert dans sa spécialité. La solution envisageable est la gestion du savoir-faire de l'expert dans une phase d'apprentissage transparente qui ne fait appel qu'à des notions relevant de son domaine de compétence.

2 Signature de l'image

La signature d'une image représente le vecteur d'attributs permettant de la caractériser. Elle contient toutes les informations de bas niveau décrivant l'image et permet de faciliter la mesure de similarité entre deux images.

2.1 modèle de l'image

Afin de mieux caractériser les images, ces dernières sont représentées par un modèle arborescent où la granularité du contenu augmente au fur et à mesure de la descente dans l'arbre. La représentation que nous avons adoptée est inspirée du modèle donné par *Rui* [3].

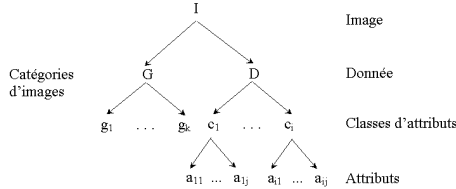


Figure 1 – modélisation des images traitées.

Notre modèle représenté par la figure 1 considère l'image comme un objet I pouvant être modélisé comme une fonction des données de l'image D (information pixel), les classes d'attributs C (couleur, texture) et les attributs A (histogrammes adaptatifs, attributs de la matrice de cooccurrences) mais aussi par sa catégorie G (marbre, corioli, buxy...). Ainsi, I peut être écrit comme :

$$I = O(D, C, A, G), \quad (1)$$

2.2 Représentation de la couleur

Afin de représenter la couleur dans ce système d'indexation dédié aux roches ornementales, nous avons choisi d'utiliser une variante des histogrammes [4] que nous appelons "histogrammes adaptatifs" H_ψ . Ces derniers sont obtenus à partir d'une quantification couleur de l'image par la méthode des palettes locales [5]. Cette quantification permet, à la différence de la quantification uniforme utilisée par la plupart des *SIIC*, de prendre en compte la dynamique de l'image. Elle permet aussi d'obtenir des palettes de faibles tailles sans pour autant altérer l'image ou son contenu sémantique.

Un histogramme adaptatif H_ψ est défini par :

$$H_\psi(\vec{c}) = \frac{1}{MN} \sum_{i \in N} \sum_{j \in M} 1_{(\vec{r}(i,j)=\vec{c})}, \vec{c} \in C, \quad (2)$$

où $H_\psi(\vec{c})$ représente le nombre d'occurrences de la couleur \vec{c} de l'image, N et M sont respectivement le nombre de lignes et le nombre de colonnes et 1_* la fonction indicatrice. C est l'ensemble des couleurs \vec{c} de cardinalité K .

Le vecteur d'attributs associé aux H_ψ est composé d'une part des occurrences et d'autre part des triplets représentant les couleurs de la palette réduite. Dans ce cas, la mise en correspondance n'est plus trivial puisque l'information de position est perdue. Afin de mesurer la similarité entre deux H_ψ , nous avons dérivé à partir des métriques standards, deux distances. L'une utilise la distance couleur pour pondérer les similarités d'occurrences et l'autre exploite les différences d'occurrences pour pondérer les similarités couleur. À cause de la limitation de pages, il est impossible de décrire ces distances. Le lecteur peut se référer à [5] pour de plus amples détails.

2.3 Représentation de la texture

La texture est un attribut très important pour la discrimination des roches ornementales. Pour caractériser cette information nous avons choisi d'utiliser une adaptation des matrices de cooccurrences aux images couleur. Cette adaptation s'effectue grâce à la quantification couleur citée précédemment. De la matrice de cooccurrences adaptative, nous extrayons deux types de descripteurs : l'histogramme des occurrences et un vecteur d'attributs composé de deux attributs d'Haralick [6] (le contraste et le moment des différences inverses) et d'un attribut baptisé critère de cohérence couleur que nous avons défini [5].

Le calcul des matrices de cooccurrences adaptatives dépend de deux paramètres : la distance δ et l'orientation θ . Pour les besoins de l'application, les attributs de texture sont calculés selon les quatre orientations standards (0° , 45° , 90° et 135°) sauf pour l'histogramme des occurrences pour lequel les quatre matrices sont moyennées.

3 Modélisation du savoir-faire

La signature de l'image étant calculée, il est indispensable de penser à la manière avec laquelle seront combinés les pouvoirs discriminants des descripteurs. L'approche que nous avons suivie consiste en la modélisation du savoir-faire grâce à une phase d'apprentissage permettant de faire fonctionner le système en suivant leurs compétences et leurs expériences.

3.1 Protocole d'apprentissage

Il est impossible de demander aux experts (notés E_m) de définir directement des poids ($\alpha_{a_{ij}}^{E_m}$) à attribuer aux différents paramètres utilisés, étant donné leurs faibles connaissances des outils de caractérisation de l'image. De notre côté, il nous est impossible de définir correctement les poids car nous ne disposons pas de l'expérience et du savoir-faire des experts lorsqu'il s'agit de classer une roche ornementale. Le but de cette étape est donc d'effectuer un apprentissage transparent à partir des différents experts en faisant appel à des notions qui relèvent de leurs domaines de compétences. Cette étape permettra de générer des poids par classe d'attributs ou même par attribut afin de combiner leurs pouvoirs discriminants tels qu'ils sont perçus par l'expert.

La démarche suivie consiste à proposer aux experts un ensemble d'images d'apprentissage noté Δ contenant un nombre d'images représentatif à la fois de la taille de la base d'images et des catégories qui la forment. Ensuite, il est demandé à chaque expert, en réponse à une image requête que nous lui proposons, de classer les images de Δ , de la plus similaire à la moins similaire à la requête, en faisant uniquement appel à ses compétences et son savoir-faire. Ce processus est répété plusieurs fois avec des images requêtes différentes et des experts différents.

Afin que les résultats d'apprentissage soient cohérents, il est indispensable de respecter les points suivants :

- L’expert doit être dans des conditions d’évaluation satisfaisantes : un écran correctement calibré et une salle d’évaluation adaptée (normes des évaluations psychovisuelles [5, chap2]);
- À chaque choix de l’expert, l’image sélectionnée est masquée afin de ne pas perturber les choix ultérieurs ;
- Répéter plusieurs fois l’apprentissage avec, à chaque fois, une nouvelle requête afin de pouvoir disposer d’un ensemble de données représentatif du savoir-faire de l’expert.

3.2 Facteurs de pondération

Les facteurs de pondération représentent les degrés d’importance que donne un expert à un attribut de la signature de l’image en fonction de son pouvoir discriminant. Cette importance se traduit par un poids $\alpha_{a_{ij}}^{E_m}$ assigné par chaque expert E_m à chaque représentation de classe d’attributs a_{ij} . Ces poids regroupés dans des vecteurs, comme le donne l’équation 3, sont directement calculés à partir de la phase d’apprentissage décrite précédemment.

$$P^{E_m} = (\alpha_{a_{11}}^{E_m}, \dots, \alpha_{a_{ij}}^{E_m}, \dots, \alpha_{a_{IJ}}^{E_m})^T. \quad (3)$$

Une fois la phase d’apprentissage terminée, en notant I_Q^Δ l’image requête de Δ et I_B^Δ une image de Δ , les facteurs de pondération liés aux savoir-faire des experts peuvent être extraits grâce aux étapes suivantes :

1. Calculer les distances entre l’attribut a_{ij} de I_Q^Δ et I_B^Δ ;
2. Classer les distances en fonction du classement effectué par l’expert E_m ;
3. Effectuer un deuxième classement des distances dans un ordre croissant (représenter la vérité terrain¹) ;
4. Calculer l’erreur entre le classement de l’expert et le classement de la vérité terrain.

$$\begin{aligned} \varepsilon_{a_{ij}}^{E_m} = & \sum_{t=1}^{A/3} (dist_V(t) - dist_{E_m}(t))^2 + \\ & \theta \sum_{t=\frac{A}{3}+1}^{2A/3} (dist_V(t) - dist_{E_m}(t))^2 + \\ & \lambda \sum_{t=\frac{2A}{3}+1}^A (dist_V(t) - dist_{E_m}(t))^2, \end{aligned} \quad (4)$$

où A représente le cardinal de Δ , $dist_V(t)$ est le vecteur des distances avec classement croissant de la vérité terrain, $dist_{E_m}(t)$ est le vecteur des distances issues du classement de l’expert E_m , θ et λ des facteurs pénalisant pour le second et le dernier tiers de la base d’apprentissage. Les facteurs pénalisant permettent de donner beaucoup d’importance aux premières images et moins d’importance aux dernières. En effet, les erreurs sur le dernier tiers du classement sont nettement moins significatives que des erreurs sur le premier tiers. Ce qui voudrait dire, que le système tolère moins les erreurs sur les premières images qui sont très pertinentes. À l’inverse, une erreur commise sur les dernières images n’est jamais significative.

¹La vérité de terrain représente l’information issue des carrières de provenance des roches ornementales

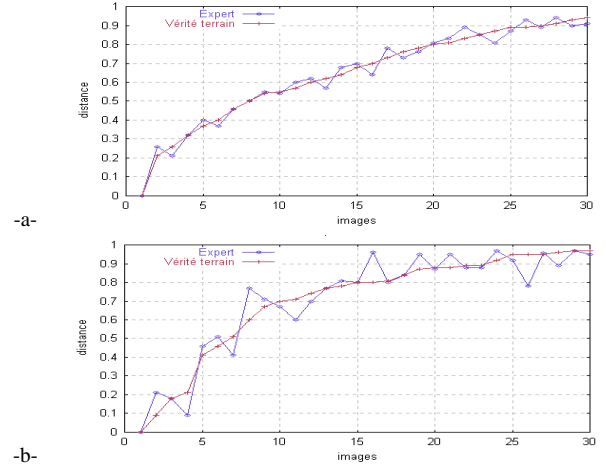


Figure 2 – Exemple d’erreurs de classification faites par un expert pour les attributs a- de couleur et b- de texture.

La figure 2 donne un exemple d’apprentissage à partir du savoir-faire d’un expert sur un ensemble d’images de la base ornementale. Il est possible de remarquer que cet expert favorise la couleur puisque les deux courbes de la figure 2-a sont pratiquement confondues. Quant à la texture la figure 2-b montre que l’expert accorde moins d’importance qu’à la couleur lors de sa classification.

3.3 Degrés de confiance

Les degrés de confiance φ^{E_m} représentent des scores attribués aux experts E_m en fonction de leurs savoir-faire et de leurs expériences. Plus précisément, il s’agit pour un expert donné, d’une note correspondant à l’exactitude de sa classification des images de roches ornementales. Ces scores permettent d’éviter de prendre en compte des experts dont la classification n’est pas fiable, dans la phase finale de mesure de similarité. C’est en quelque sorte un facteur d’amortissement de l’effet de l’expert non-fiable et un facteur de rehaussement de l’effet de l’expert fiable.

Expert/Vérité terrain	Classe a	Classe b
Classe a	BC_1	MC_1
Classe b	MC_2	BC_2

Tableau 1 – les différents cas de classification de roches ornementales

Afin de calculer ces degrés de confiance, une seconde phase d’apprentissage sur Δ est nécessaire. Il s’agit cette fois-ci de proposer aux experts, de faire la classification d’un certain nombre d’images. Une fois cette classification terminée, les résultats sont exploités à l’aide du tableau 1. Ce dernier se résume en deux cas : bonne classification et mauvaise classification. Où BC représente une bonne classification et MC une mauvaise classification. À partir des valeurs obtenues à l’aide du tableau 1 et en posant $BC=BC_1+BC_2$ et $MC=MC_1+MC_2$, les degrés de confiance sont calculés grâce à l’équation suivante :

$$\varphi^{E_m} = \left\{ \begin{array}{ll} \frac{BC-MC}{BC+MC} & \text{si } BC \gg MC, \\ 0 & \text{sinon.} \end{array} \right\} \quad (5)$$

Notons qu'un expert est fiable, avec un degré de fiabilité variable, à partir du moment où MC est inférieur, voire très inférieur, à BC . En cas d'égalité, l'expert est considéré comme non fiable.

3.4 Facteurs de pondération finaux

Le calcul des facteurs de pondération finaux s'articule autour des facteurs de pondération de chaque expert associés à leur degré de confiance. La combinaison est effectuée grâce à l'équation suivante :

$$P_F = \sum_{m=1}^n \varphi^{E_m} * P^{E_m}, \quad (6)$$

où n représente le nombre d'experts.

4 Expérimentation

L'expérimentation a été effectuée sur une base de roches ornementale qui nous a été fournie par le Centre de Recherches sur les Poudres Minérales Colorées de l'École des Mines d'Alès, décentralisé à Pau. L'apprentissage s'est déroulé auprès de 3 différents experts permettant ainsi d'avoir trois différents vecteurs de pondération (P_1 , P_2 et P_3) associés à des degrés de confiance. Les résultats de requêtes donnés par la figure 3 montre que la coopération des différents experts (P_F) permet d'améliorer considérablement les résultats.

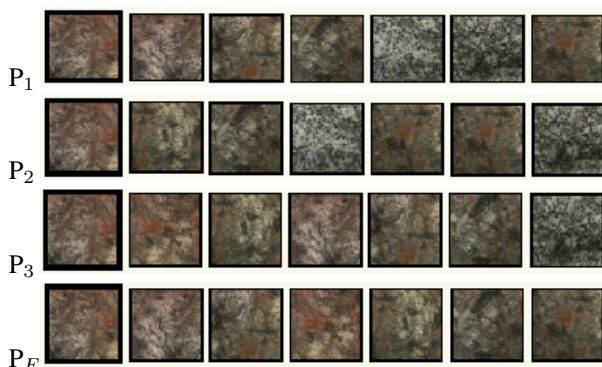


Figure 3 – Exemple de résultats de requête dans la base ornementale

La figure 4 donne les courbes de performances grâce aux critères de précision/rappel [7] permettant de comparer l'utilisation des facteurs de pondération de chaque expert séparément par rapport à la technique de coopération proposée. Les courbes montrent que la coopération permet d'obtenir la meilleure précision sur tout l'intervalle de rappel confirmant par la même occasion l'amélioration apportée par cette technique.

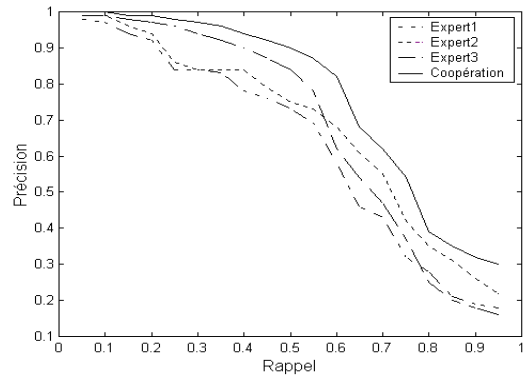


Figure 4 – Courbes d'évaluation par les critères de précision/rappel.

5 Conclusion

Dans cette contribution, nous avons essayé de répondre à la question suivante : pouvons-nous exploiter le savoir-faire des experts dans la boucle d'indexation d'image par le contenu ? D'après les résultats obtenus, il est possible de répondre positivement à cette question. En effet, la modélisation du savoir-faire de l'expert grâce à des facteurs de pondération et un degré de confiance a permis d'éviter la lourde tâche de définition de poids. Les résultats obtenus sont satisfaisants et en continuité de cette gestion de connaissance, au lieu de faire une combinaison, il est possible de mettre les experts en concurrence afin que le vecteur de pondération final puisse d'une part s'adapter à la requête et d'autre part, exploiter les données intrinsèques du système d'une manière non-supervisée [5].

Références

- [1] A. Del Bimbo. *Visual Information Retrieval*. Morgan Kaufmann Publishers, Inc., San Francisco, 1999.
- [2] R. Schettini, G. Ciocca, et S. Zuffi. A survey on methods for colour image indexing and retrieval in image databases. Dans *Color Imaging Science : Exploiting Digital Media*, 2001.
- [3] Y. Rui. *Efficient Indexing, Browsing and Retrieval of Image/Video Content*. Thèse de doctorat, Université de l'Illinois à Urbana-Champaign, USA, 1999.
- [4] M. J. Swain et D. H. Ballard. Indexing via color histograms. Dans *Int. Conference on Color Vision*, pages 390–393, 1990.
- [5] M.-C. Larabi. *Codage et Analyse d'images couleur : Application à l'indexation de base d'images réparties*. Thèse de doctorat, Université de Poitiers, Déc. 2002.
- [6] R. M. Haralick, K. S. Shanmugan, et I. Dunstein. Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics*, 3(6) :610–621, 1973.
- [7] G. Salton et M. J. McGill. *Introduction to Modern Information Retrieval*. Computer Science Series, McGraw-Hill, 1983.