

EXTRACTION DES CARACTERISTIQUES PICTURALES DE BOIS GRAVES POUR UNE RECHERCHE DE FORMES

Victor Chen-yuan, Su Ruan

SIC-CReSTIC, groupe Image

IUT-TROYES

9, rue du Québec

10026 Troyes – France

{Prénom.Nom}@univ-reims.fr

Concours Jeune Chercheur: Non

Résumé

Dans cet article, nous proposons de segmenter des images de tampons de bois gravés anciens à des fins d'illustrations de documents patrimoniaux ou de productions de nouveaux tampons à l'aide de l'analyse multi-échelle. L'enjeu de cet archivage est de fournir aussi bien aux utilisateurs professionnels qu'au grand public les contenus des images de tampons sous forme de logos. Notre contribution est de mettre en œuvre une plate-forme qui permette, en partant d'une image de tampon, d'obtenir l'impression qu'il aurait produite sur papier. Nous montrons dans cette communication que la plupart des méthodes classiques de détection de contours échouent sur ces types d'images mêmes si a priori les contenus semblent composés de traits. Les aplats constitués d'empreintes laissées par des outils sont difficilement éliminés et le dimensionnement des contours est rarement pris en compte. Afin de réduire ces fausses descriptions, nous procédons à l'utilisation de traitements basés sur une analyse multi-échelle afin de différencier au mieux les variétés de structures. Les résultats d'extraction obtenus sont présentés et sont perceptuellement satisfaisants tout en montrant une bonne impression de l'image sans avoir à effectuer trop de retouches.

Mots Clefs

Segmentation, détection de contours, décomposition multi-échelle,

1 Introduction

L'archivage d'images numériques est devenu cette dernière décennie un axe de recherche très actif en tant que données visuelles avec l'émergence d'Internet. Avec ce moyen de communication, il devient possible de mettre des documents anciens rares et fragiles, en particulier des tampons de bois gravés, à la disposition du public sur cédéroms ou sur les réseaux ; la communication de ces documents était jusqu'alors

interdite en raison de leur fragilité et de leur unicité. Afin de faciliter l'accès au patrimoine archivistique, le projet que nous menons a pour but, outre l'archivage des images originales, la production indexée des images bitonales de leur homologue pour la conservation et la diffusion des tampons de bois gravés. Cette activité a été créée pour répondre à des besoins des consortiums comprenant des organismes en charge du patrimoine, des institutions de recherche et des utilisateurs grand public pour une illustration de documents ou pour une fabrication de nouveaux tampons. La tâche importante à réaliser concerne la segmentation des images qui doit fournir un rendu similaire à l'impression qu'il aurait produit sur papier. Au vu des échantillons d'images (figure 1) qui semblent a priori constituer essentiellement d'images de traits, nous avons initié nos procédures de traitement à l'aide des opérateurs classiques de détection de contours. Malheureusement, ces approches n'ont pas permis de réaliser de segmentation attendue. Il semble que ces méthodes standard de traitement ne sont pas très adaptées aux contenus structurelles des images de tampons. Elles échouent sur le dimensionnement des contours et sur la suppression difficile des empreintes laissées par les outils. En fait, la procédure de segmentation doit, outre l'extraction des contours, requérir une analyse apte à discriminer les aspects relatifs aux textures qui peuvent représenter le fond et les objets picturaux de l'image. Pour y remédier, nous orientons nos travaux vers des représentations multi-échelle. Cette théorie de l'espace-échelle développée par Lindberg [1] est fondée sur l'extraction des régions d'intérêt par la recherche des extrema d'une fonction lissée. L'association d'une région entière à un simple extremum est attirante par la simplicité de représentation. Nous proposons dans ce travail de caractériser les images de tampons en mettant en œuvre l'approche par espace-échelle Gaussien [2]. Le modèle présente non seulement de bonnes propriétés sur le plan de la généralité mais aussi de la bonne fiabilité et de robustesse [3]. L'implantation du filtre gaussien est très pratique, il est stable et séparable. Il peut être implanté récursivement si on cherche un gain de temps de calculs. La suite de l'article est articulé comme suit. Nous introduisons dans la section 2 un rappel du cadre

multi-échelle utilisé pour la caractérisation des images de bois gravés. Dans la section 3, nous décrivons l'opérateur DoG (Differential of Gaussian) utilisé initialement pour la caractérisation des points saillants stables [4] et argumentons de ce choix dans le cadre de cette application. Nous présentons dans la section 4 les différents résultats obtenus et nous concluons l'article par de brèves discussions sur l'automatisation et la qualité de la segmentation des images de bois gravés.



Figure 1 - Présentation d'un échantillon d'image de bois gravé ancien

2 Analyse multi-échelle

L'idée principale de l'approche multi-échelle est de mettre en évidence les différentes tailles de structures de l'image par l'application d'un banc de filtres lissants passe bas. Les filtres de faible suppression les plus fins tandis que les filtres de plus grand rayon conservent les grandes structures tout en éliminant les structures plus petites. La description de cette procédure peut être illustrée sur la figure 2. Dans la théorie espace-échelle, l'analyse s'appuie de façon implicite sur un noyau constituée d'une fonction Gaussienne de paramètre σ qui sert de facteur d'échelle. L'expression de la fonction Gaussienne bidimensionnelle est définie par :

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (1)$$

Les atouts de la fonction Gaussienne sont multiples. Elle permet d'obtenir une description invariante de l'image et de constituer aussi un vecteur de caractérisations d'un point donné de l'image moyennant uniquement des dérivées consécutives de la fonction [7]. L'association de ce vecteur aux différentes composantes produites est exploitée dans de nombreux travaux à des fins d'extraction de points d'intérêts stables et répétables en vue d'appariement ou d'indexation d'images [8][9].

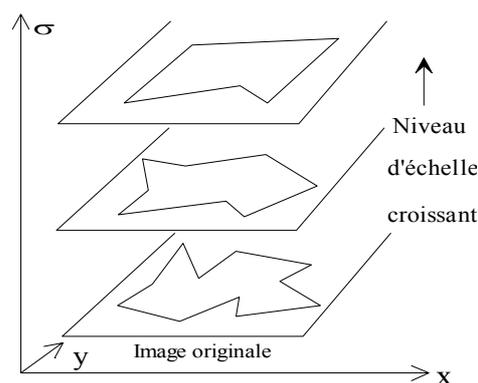


Figure 2 - Illustration de la représentation multi-échelle d'une image

3 Les méthodes de décomposition

Dans cette section, nous introduisons la méthode de décomposition utilisée pour segmenter les images de tampons. Appelons $L(x,y,\sigma)$ la fonction de lissage résultant du produit de convolution entre l'image $I(x,y)$ et la fonction Gaussienne $G(x,y,\sigma)$, soit:

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (2)$$

où \otimes représente l'opérateur de convolution.

Habituellement, les caractéristiques locales sont calculées à partir des dérivées Gaussiennes moyennant une échelle σ appropriée. A partir de cette approche, un certain nombre de descripteurs multi-échelle ont été développés pour des applicatifs tels que l'appariement ou l'indexation des images. Ici, nous nous intéressons à la caractérisation de tampons de bois gravés en exploitant le potentiel de filtrage et de détection de contours pour la décomposition multi-échelle. A ce jour, les résultats de segmentation sont jugés subjectivement par une analyse visuelle de l'expert. En vue de s'affranchir du contrôle humain et rendre la binarisation automatique, nous nous recourons à quelques opérateurs optimaux de détection de contours qui nous informent sur l'efficacité de segmentation. Pour cela, nous testons deux opérateurs de contours optimaux choisis suivant une implémentation simple et un coût de traitement équivalent à notre approche.

3.1 Opérateur de Canny

Afin de juger des performances de chacune des approches, nous introduisons l'opérateur de Canny [10] dont l'approche originale a été développée pour permettre la bonne compréhension des conditions d'une bonne détection de contours moyennant les trois critères : bonne détection, bonne localisation et réponse unique. En outre, pour mettre en œuvre son opérateur

en 2D, Canny a montré que la dérivée première d'une fonction Gaussienne est une bonne approximation de son filtre avec des dégradations minimales telle que:

$$h(x, \sigma) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad (3)$$

Le passage à un espace à deux dimensions se fait simplement en raison de la séparabilité de la gaussienne, on a :

$$h(x, y, \sigma) = h(x, \sigma) \cdot h(y, \sigma) \quad (4)$$

3.2 Opérateur de Shen-Castan

L'opérateur de Shen [11] fait partie de la classe de détecteurs dont la réponse impulsionnelle est discontinue en zéro dans le domaine spatial continu. Ce détecteur de contours proposé a la forme d'une fonction exponentielle symétrique. Le filtre de lissage obtenu par Shen s'écrit :

$$h(x) = ce^{-\alpha|x|} \quad (5)$$

et le filtre de dérivation correspondant :

$$h'(x) = \begin{cases} de^{-\alpha x} & \text{pour } x \geq 0 \\ de^{\alpha x} & \text{pour } x < 0 \end{cases} \quad (6)$$

$$\text{avec } c = \frac{1 - e^{-\alpha}}{1 + e^{-\alpha}}$$

$$\text{et } d = 1 - e^{-\alpha}$$

où c et d représentent le facteur de normalisation du filtre de lissage et de dérivation, α est le paramètre d'étalement des filtres.

3.3 Opérateur Laplacien de la gaussienne

Pour décrire les descripteurs multi-échelle d'une image, le représentant le plus connu d'entre eux est certainement l'opérateur LoG (Laplacian of Gaussian) en raison de sa performance dans la caractérisation des blobs [3]. En général, on utilise le LoG normalisé afin de rendre possible la comparaison des fiabilités de blobs issus d'échelles différentes:

$$\nabla_{\text{norm}}^2 \text{LoG}(x, y, \sigma) = \sigma^2 (L_{xx}(x, y, \sigma) + L_{yy}(x, y, \sigma)) \quad (7)$$

où $L_{uv}(\cdot)$ représente le résultat du filtrage issu des dérivées spatiales partielles d'ordre 2 de la gaussienne suivant la direction u et v respectivement.

3.4 Opérateur Différence de la Gaussienne

Initialement introduit par Marr et Hildreth [12] pour modéliser les systèmes visuels biologiques pour la segmentation d'images, l'opérateur DoG (Differential of Gaussian) connaît un regain d'intérêt en matière de caractérisation d'images pour la simplicité et la souplesse de paramétrages. A la place de l'opérateur LoG(x,y, σ), nous optons pour le modèle de décomposition l'opérateur DoG(x,y, σ) qui présente des propriétés plus intéressantes. Les raisons de cette préférence sont: 1) l'existence des techniques rapides pour générer des fonctions Gaussiennes modélisant une structure pyramidale; 2) l'opérateur DoG(x,y, σ) est quasi-invariant aux changements d'échelles évitant la normalisation des résultats par σ^2 ; 3) la répétabilité du DoG(x,y, σ) dans le cadre de la détection de coins [5] et la fiabilité de la localisation de ces points d'intérêt [6]. Nous proposons donc une alternative intéressante basée sur l'opérateur DoG(x,y, σ) pour extraire outre les contours les informations basées sur les caractéristiques des régions. L'opérateur proposé est obtenu en différenciant simplement deux fonctions gaussiennes séparées par un facteur multiplicateur k:

$$\begin{aligned} \text{DOG}(x, y, k\sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) \otimes I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (8)$$

Nous voyons que le résultat de l'équation (8) est primordial dans la description des caractéristiques qui se composent uniquement de la soustraction de deux fonctions lissées à des échelles différentes. Ce modèle évite ainsi d'introduire la constante de normalisation comme dans l'utilisation du Laplacien $\sigma^2 \nabla^2 \text{LoG}(x, y, \sigma)$ étudié par Lindberg [1] afin de s'assurer de l'invariance des caractéristiques à différentes échelles. Nous illustrons sur la figure 3a et 3b une représentation en 2D de l'opérateur LoG et de l'opérateur DoG utilisés pour la détection des contours qui se caractérisent par la recherche du passage par zéro pour le premier et la pyramide des erreurs pour le deuxième. Nous observons que les deux graphes sont quasi-similaires à une constante près donnée par [13]:

$$\text{DoG}(x, y, \sigma) = 0.4875\sigma^2 \text{LoG}(x, y, \sigma) \quad (9)$$

Malgré leur similitude, la différence principale entre les deux opérateurs réside dans la pratique. Outre les qualités citées précédemment, le DoG permet une scrutation rapide, les objets suivis sont plus rapidement appréhendés. Ainsi, on trouve le DoG comme moteur de recherche dans la procédure de mean-shift [13] connue pour traquer les blobs contenus dans une image.

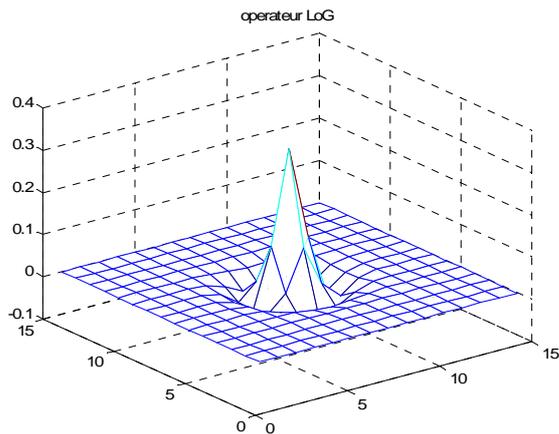


Figure 3a - représentation en 2D de l'opérateur LoG

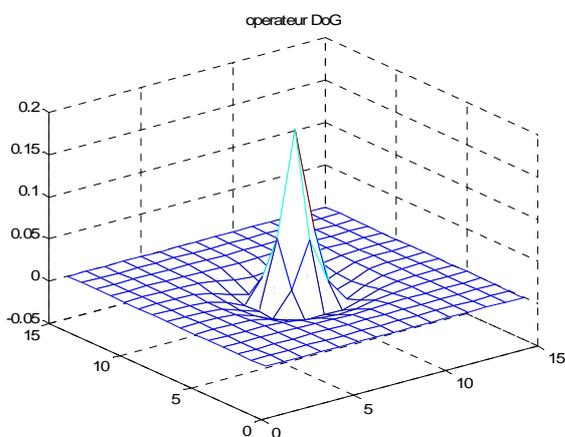


Figure 3b - Représentation en 2D de l'opérateur DoG

4 Expérimentations

Dans ce paragraphe, nous présentons des résultats expérimentaux montrant l'attrait de l'utilisation de la fonction Gaussienne pour la caractérisation et la segmentation des images de tampons de bois gravés. Les expérimentations sont initialement menées sur la base des opérateurs différentiels classiques (gradient, Sobel). Sur l'application visée ensuite, dans le but de s'affranchir de l'analyse visuelle de l'expert et réaliser une binarisation automatique éventuelle, nous avons recouru à quelques filtres optimaux de détection de contours (Canny, Shen) pour une évaluation prédictive des résultats extraits de la carte de saillance ou de contours. L'objectif ici est d'obtenir une évaluation de la qualité des traits de silhouettes et des frontières moyennant ces dispositifs de calculs. Le mode de sélection est-il en adéquation avec le rendu du traitement donné par l'application de la décomposition multi-échelle. Dans son implantation, le filtre de Canny utilisé est constitué du noyau issu de la dérivée première de la Gaussienne, le filtre de Shen est implémenté sous forme de noyau et non pas de séquences récursives, l'opérateur Laplacien LoG(x,y,σ)

est réalisé à partir de la dérivée seconde de la Gaussienne. Nous comparons enfin les résultats obtenus à ceux fournis par le DoG(x,y,σ). Le protocole de segmentation utilisé est très classique : 1) Si la caractérisation est opérée suivant deux directions privilégiées en ligne et en colonne, la binarisation est effectuée selon la technique de recherche des maxima locaux et la construction des contours par hystérésis ; 2) Lorsque la caractérisation s'opère sans direction particulière, l'extraction des contours est réalisée à partir d'un seuillage global. Nous illustrons sur la figure 4 un banc de résultats réalisés par les différents opérateurs. Du point de vue de l'analyse de scènes, les résultats obtenus (Figure 4a) par l'application des opérateurs classiques sont inappropriés, ils ne respectent pas les formes picturales du tampon. Les opérateurs utilisés apparaissent mal adaptés aux tâches de sélection des caractéristiques dominantes (évolution des traits, analyse de texture du fond, discrimination des objets). Nous observons aussi que les opérateurs à double composante directionnelle ne conviennent pas aux types d'images traités. Les traits sont altérés ou entrecoupés. Cette inadéquation de structures pose alors de problème dans la représentation ou la restitution d'objets cartographiques. Les méthodes d'inspirations analytiques sont également mal adaptées à l'interprétation des images. Les images segmentées (Figure 4b) par des filtres optimaux montrent au premier regard une uniformité de contours et un difficile problème de stylisation de rendu. Leur conjonction n'a pas non plus permis d'aboutir à une évaluation de la qualité des caractéristiques extraites. Nous présentons sur la figure 4c les résultats provenant de l'analyse gaussienne multi-échelle. A ce jour, les meilleurs résultats de segmentation sont fournis par les deux approches avec une restitution des caractéristiques plus fine pour le DoG. Les ornements de la tête sont visiblement mieux ressortis avec cet opérateur. Pour cette image caractéristique, la valeur de résolution σ permettant de réaliser une cartographie de contours convenable avec une perte minimale se situe entre 1 et 2. L'extraction des traits caractéristiques est réalisée à l'aide d'un seuillage global. Une expérimentation plus systématique est en préparation notamment l'utilisation de l'approche sur d'autres séries de tampons pour juger de sa stabilité et de sa reproductibilité.

5 Conclusion

Nous avons présenté une approche de segmentation de tampons de bois gravés en s'appuyant sur une technique de décomposition multi-échelle de la Gaussienne. Par l'intermédiaire de cette approche, nous obtenons actuellement une bonne impression de l'image sans avoir à effectuer trop de retouches. Les résultats rencontrés pour une échelle σ compris entre 1 et 2 sont peceptuellement satisfaisants. Dans le cadre de cette étude, nous nous sommes également

confrontés sur la position des traits à la surface des objets. Ainsi, la représentation de certains traits par les opérateurs de contours est incohérente à la perception visuelle, ce sont des traits laissés par les outils et sont considérés comme étant l'image de fond. Globalement, les résultats donnés par l'analyse multi-échelle sur la segmentation d'images de tampons médiévaux sont encourageants. Certes, des points durs subsistent, l'automatisation de la segmentation est somme toute réhibitoire en raison de la diversité de matériaux utilisés et de la variété de tampons à traiter. La qualité de segmentation est également très tributaire des phénomènes liés aux problèmes d'encrage et aux accidents subits par le bois au cours de son existence. L'intervention de l'expert s'avère nécessaire pour juger de la bonne binarisation du tampon. Des améliorations ultérieures peuvent être envisagées en mettant en place une procédure de sélection pertinente de crêtes et de vallées à base de système d'attention visuelle moyennant des modèles de représentations de connaissances adaptées.

Références

- [1] T. Lindberg, *Scale-space theory in computer vision*, Kluwer Academic Publishers, Pays-bas, 1994.
- [2] T. Lindberg, Feature detection with automatic scale selection, *International Journal of Computer Vision*, vol. 30, n°2, pages 79-116, 1998.
- [3] R. Megret, J.M. Jolion, Suivi de blobs de niveaux de gris pour la représentation du contenu dynamique d'un vidéo, *Rapport de Recherche RR-2001-05, RFV, INSA de Lyon*, Septembre 2001.
- [4] D.G. Lowe, Object recognition from local scale invariant features, *International Conference on Computer Vision (ICCV)*, pages 1150-1157, 1999.
- [5] K. Mikolajczyk, C. Schmid, Indexing based on scale invariant interest point, *International Conference on Computer Vision*, pages 525-531, 2001.
- [6] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, vol. 60, n°2, pages 91-110, 2004.
- [7] J.J. Koederink, A.J. Van Doorn, Representation of local geometry in the visual system, *Biological Cybernetics*, 55, pages 367-375, 1987.
- [8] C. Schmid, R. Mohr, Local greyvalue invariants for image retrieval, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, n°5, pages 530-535, 1997.
- [9] Y. Dufournaud, C. Schmid, R. Horaud, Appariement d'images à des échelles différentes, *Actes du 12^{ème} Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, France, vol. 2, pages 327-336, 2000.
- [10] J. Canny, A computational approach to edge detection, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 8, n°6, pages 679-698, 1986.
- [11] J. Shen, S. Castan, An optimal linear operator for edge detection, *IEEE Conference Vision Pattern Recognition*, pages 109-114, 1986.
- [12] D. Marr, E. Hildreth, Theory of edge detection, *Proc. Roy. Soc. London.*, vol. B 207, pages 187-217, 1980.
- [13] R. T. Collins, Mean-shift blob tracking through scale space, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2003)*, pages 234-240, 2003.

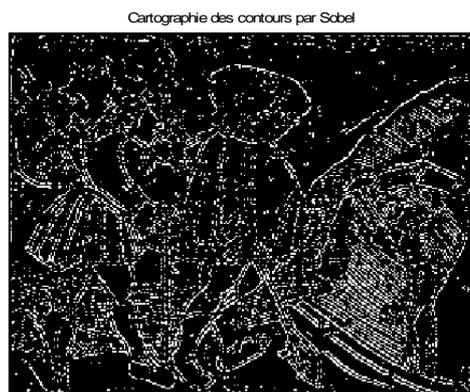
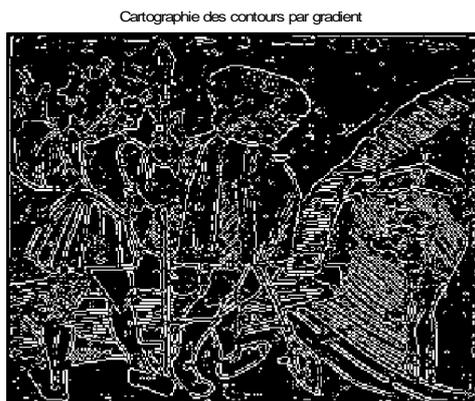


Figure 4a - Cartographie des contours par des opérateurs de contours classiques



Figure 4b - Cartographie des contours par des opérateurs optimaux des contours



Figure 4c - Cartographie des contours par l'analyse gaussienne multi-échelle