

Recherche d'images iconique utilisant les moments de Zernike

V. Chen¹ A. Szabo² M. Roussel¹

¹ LAM, équipe Image

IUT - Troyes
9, rue du Québec
10026 Troyes Cedex – France

² University of Oradea

Street Armatei Romane n°5
Oradea 3700 – Romania

{v.chen, m. roussel}@iut-troyes.univ-reims.fr

Résumé

Nous traitons dans ce papier l'identification de bois gravés anciens au moyen de la technique de mise en correspondance des images. Elle s'appuie sur l'extraction des points d'intérêts et sur l'exploitation pertinente des primitives de formes locales dont les caractéristiques sont générées par les moments de Zernike. Actuellement, la mesure de ressemblance entre deux points homologues est réalisée par une fonction de similarité utilisant la distance Euclidienne. Les résultats de comparaison d'images sont prometteurs. Pour être compatible avec la recherche d'images en ligne, nous devons améliorer le temps de traitements par l'utilisation d'autres opérateurs tels les inégalités triangulaires. L'autre perspective vise à comparer nos résultats avec d'autres approches telles les invariants différentiels ou combiner les deux approches pour obtenir un système robuste.

Mots clés

Moments de Zernike, mise en correspondance, classification, descripteurs de formes.

1 Introduction

Avec l'avènement de l'internet, les échanges de données, en particulier les images, sont intenses et tendent à continuer dans le futur à cause de la profusion de nouvelles applications (bureautique, galerie d'art ou de peinture...) destinées au grand public. C'est dans ce cadre, au travers de l'archivage des images de bois gravés, que se positionne notre approche qui consiste mettre en œuvre des outils d'aide permettant à l'utilisateur de trouver les images qui lui conviennent. Dans ce travail, nous partons de plusieurs reproductions de gravure pour lesquelles nous

devrions établir une classification des bois semblables. Pour contourner les problèmes inhérents aux différentes conditions de fabrication (images biaisées ou décalées, similitude des figurines) et d'utilisation (dispersion de luminance importante d'une image à l'autre), la stratégie que nous avons développée s'appuie sur la recherche par le contenu de l'image et l'élaboration d'un système de reconnaissance qui se fonde sur la mise en correspondance, technique inspirée de la stéréovision. Nous caractérisons les images par leurs formes locales à partir d'une carte de points d'intérêts. Certes les caractéristiques locales sont plus sensibles aux bruits mais associées sous forme de composantes, elles s'avèrent plus robustes et elles permettent de mieux appréhender les formes des imageries analysées. Dans le but de contribuer à l'avancement du domaine, la piste de recherche que nous avons adoptée porte sur la comparaison d'images en appariant les caractéristiques locales de formes qui seront extraites par le biais des moments de Zernike. Elle opère suivant un mécanisme de similarité pour évaluer les images les plus ressemblantes à l'image requête. Pour illustrer notre approche, les expériences sont menées sur des échantillons d'images de tampons gravés anciens et montrent des résultats tout à fait satisfaisants.

2 Stratégie de mise en correspondance

La stratégie de mise en correspondance des images nécessite la construction d'une carte de points d'intérêts des images à apparier. Les points d'intérêts sont recherchés pour être robustes et fiables aux transformations géométriques [1], permettant d'obtenir ainsi une description du contenu de l'image fidèle. Le schéma bloc général du système de reconnaissance et de recherche d'images est présenté sur la figure 1.

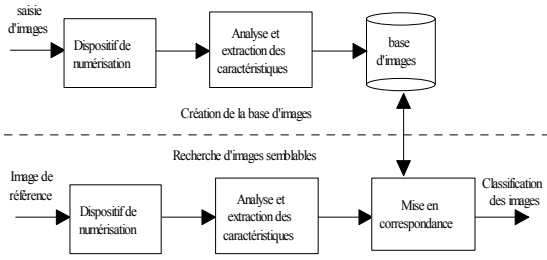


Figure 1: Système de recherche et de comparaison d'images

2.1 Détection des points d'intérêts

Utilisés initialement en stéréoscopie pour la reconstruction 2D/3D et en reconnaissance de formes pour la description des objets, les points d'intérêts se caractérisent par un changement bidimensionnel du signal. La pertinence de ces points repose sur la caractéristique locale et la richesse de l'information contenue dans le signal en ces points, ce qui leur confère une certaine robustesse aux problèmes d'occultation d'une part et de chaînage dans le cas de segmentation en contours ou en régions d'autre part. Les détecteurs de points d'intérêts foisonnent dans la littérature [1, 2]. Pour détecter les points, nous avons choisi le détecteur de Harris amélioré qui utilise un filtre récursif pour les dérivées [3] et un voisinage de cinq pixels pour les maxima locaux. Ce détecteur a été évalué comme étant le meilleur détecteur de points [4] et présente de surcroît une bonne répétabilité pour extraire automatiquement les points. Dans nos expériences, le détecteur de Harris développé est similaire à l'opérateur de Baudet qui caractérise les rendus du filtrage sous la forme d'une matrice telle que :

$$M(\vec{x}, \sigma) = \begin{bmatrix} \langle L_x(\vec{x}, \sigma)^2 \rangle & \langle L_x(\vec{x}, \sigma)L_y(\vec{x}, \sigma) \rangle \\ \langle L_x(\vec{x}, \sigma)L_y(\vec{x}, \sigma) \rangle & \langle L_y(\vec{x}, \sigma)^2 \rangle \end{bmatrix}$$

$$\text{avec } L_x(\vec{x}, \sigma) = \frac{\delta I(\vec{x})}{\delta x} \oplus G_w(\vec{x})$$

$$\text{et } L_y(\vec{x}, \sigma) = \frac{\delta I(\vec{x})}{\delta y} \oplus G_w(\vec{x})$$

où $\vec{x} = (x, y)$ représente le point de l'image et $\langle L \rangle$ le résultat de l'opération de lissage sur l'image I par un masque de dérivées d'une Gaussienne $G_w(\vec{x})$. L'évaluation du point d'intérêt IP proposé par Harris et Stephen [4] est donnée par l'expression suivante :

$$IP_{HS}(\vec{x}, \sigma) = \text{Det}(M(\vec{x}, \sigma)) - k \text{Trace}^2(M(\vec{x}, \sigma))$$

avec $S_b \leq IP_{HS}(\vec{x}, \sigma) \leq S_h$ où σ est le gabarit de la gaussienne et k un paramètre arbitraire.

Les points d'intérêts de l'image seront extraits suivant un seuillage par hystérésis délimité par deux valeurs de seuils S_b et S_h .

2.2 Les moments orthogonaux

On assiste cette dernière décennie dans le domaine du traitement d'images à un regain d'intérêt pour l'utilisation des moments [5]. Ce surcroît d'intérêt peut s'expliquer par la conjonction de plusieurs facteurs liés à l'émergence des calculateurs rapides, aux champs d'applications nouvelles telles la reconnaissance de formes dont l'image est le support et à l'apparition des méthodes de résolution basées sur des invariants algébriques, outils très prisés en reconnaissance de formes. En analyse d'image, les plus intéressants sont les moments orthogonaux qui peuvent être obtenus par le biais des polynômes de Legendre ou de Zernike. Leur utilisation en tant qu'attributs discriminants permet de mettre en évidence les données structurelles pertinentes. En plus, ces deux familles sont tout à fait aptes à représenter ou à reconstruire des objets avec un minimum de redondance de données. Dans ce travail, notre choix s'est porté sur les moments de Zernike pour sa justesse de calculs et pour ses caractéristiques invariantes en rotation.

2.2.1 Les moments de Zernike

Pour une image, les moments de Zernike sont définis par la relation suivante :

$$\lambda_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(r, \theta)$$

sachant que $V_{nm}(r, \theta) = R_{nm}(r) \exp\{jm\theta\}$ où $R_{nm}(r)$ représente le polynôme radial de Zernike. Ce dernier peut être décrit par :

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} r^{n-2s} \text{ où } n$$

et m sont des nombres entiers tels que $n \geq 0$, $|m| \leq n$ et les valeurs de $(n - |m|)$ sont des entiers paires.

2.2.2 Approche pratique

Le calcul du polynôme radial de Zernike $R_{nm}(r)$ est effectué suivant un algorithme récursif [8][9] qui transforme l'écriture originale en modifiant l'indice de r^{n-2s} par r^s et en maintenant $(n-s)$ paire, soit

$$R_{nm}(r) = \sum_{s=m}^n B_{nm} r^s$$

$$\text{avec } B_{nms} = \frac{(-1)^{(n-m)/2} ((n+s)/2)!}{((n-s)/2)! ((m+s)/2)! ((s-m)/2)!}$$

A partir de cette expression, nous déduisons les autres relations par une approche récursive :

$$B_{nms} = 1,$$

$$B_{n(m-2)n} = B_{nmn} \frac{(n+m)}{(n-m+2)},$$

$$B_{nm(s-2)} = -B_{nms} \frac{(s+m)(s-m)}{(s+n)(n-s+2)}$$

2.2.3 Extraction des primitives de formes

Dans ce paragraphe, nous décrivons l'approche utilisée pour extraire les primitives de formes. La méthode d'extraction des formes est inspirée des travaux de Khotanzad [6] qui sélectionnait une série de moments pour une reconstruction d'image et une illustration de la propriété d'invariance en rotation. Une liste des moments est donnée sur la table 1. Ainsi, par une approche de focalisation moyennant un nombre de moments choisis judicieusement, il est possible d'effectuer des traitements adaptés à l'extraction des détails de l'image et à la discrimination efficace des formes. L'ensemble de ces moments constitue un vecteur caractéristique X de composantes λ_{ij} .

Ordre	Numéro des moments	Moments
0	λ_{00}	1
1	λ_{11}	1
2	$\lambda_{20} \lambda_{22}$	2
3	$\lambda_{31} \lambda_{33}$	2
4	$\lambda_{40} \lambda_{42} \lambda_{44}$	3
5	$\lambda_{51} \lambda_{53} \lambda_{55}$	3

Tableau 1: Liste des moments de Zernike

2.2.4 Points de correspondance

Le processus de mise en correspondance consiste à appairer un vecteur de caractéristiques X^r d'un point d'intérêt trouvé dans l'image de référence avec un autre vecteur X^b dans l'image provenant de la base d'images. Cette comparaison locale est effectuée à l'aide de la fonction de similarité $S(\cdot)$ décrite ci-après :

$$S(X^r, X^b) = \sum_{h=1}^p W_h D_h(X_h^r, X_h^b)$$

où $D_h(\cdot)$ représente la distance Euclidienne, W_h l'opérateur de pondération [7] et h le nombre de descripteurs choisis.

Afin de rester à des temps de traitements raisonnables et de contourner le problème de comparaison qui est hautement combinatoire, le champs de recherche effectué sur l'image cible de la base est délimité sur un intervalle de ± 5 pixels dont la ligne centrale correspond aux points d'intérêts détectés dans l'image de référence. A l'issue de cette étape, nous procédons à la classification des points possédant des valeurs supérieures à un seuil T et ne retenons que les points les plus probables moyennant le critère du maximum de vraisemblance effectué sur une fonction décroissante définie par: $\rho = e^{-S(\cdot)/2}$. Pour quantifier le classement des images similaires, nous effectuons une évaluation complète, sur tous les points

d'intérêts, du score correspondant au nombre d'appariements correct par rapport au nombre d'appariements détectés dans les deux images.

3 Résultats expérimentaux

L'approche est appliquée sur des images numérisées en 256 niveaux de gris et de dimension 256x256. Le nombre de points d'intérêts retenu est fixé autour de 300. Cette valeur nous fournit un bon compromis entre une bonne représentation des motifs de l'image et un temps de calculs non pénalisant. Nous présentons dans le texte final les résultats de la recherche d'images par leur contenu sur une base de 50 modèles fournie par la bibliothèque de Troyes ainsi que leur degré de similarité respectif. Sur cette série d'images, ne sont présentées que les 8 images (fig. 2) classifiées les plus proches de l'images de référence repérée en première ligne. Dans l'ensemble, les résultats obtenus sont très encourageants au vu des images traitées qui se caractérisent par leur grande hétérogénéité.



Figure 2: Résultat de comparaison d'images

4 Conclusions et perspectives

Dans ce papier, nous avons présenté une solution au problème d'appariement des images de bois gravés basés sur un exemple particulier qui servait à illustrer l'histoire des quatre fils Aymon au 17^e siècle. Les résultats de recherche d'images par contenu et de classification d'images proposés par le système sont cohérents avec ce qui est donné dans la base d'images. Dans la prochaine étape, nous visons à rechercher une solution pour diminuer le temps de traitements, incompatible actuellement avec la recherche d'images en ligne, en utilisant par exemple le modèle des inégalités triangulaires, procédé très prometteur pour réduire les comparaisons d'objets. L'autre perspective vise à comparer nos résultats avec d'autres approches telles les invariants différentiels ou combiner les deux approches pour réaliser un système robuste.

Références

- [1] C. Schmid, R. Mohr, Local gray level value invariant for image retrieval, *IEEE Trans. Pattern Anal. Match. Intell.*, 19 (5), 530-535, 1997.
- [2] R. Laganière, A morphological operator for corner detection, *Pattern Recognition*, 31 (11), 1643-1652, 1998.
- [3] R. Deriche, Recursively implementing the Gaussian and its derivatives, *Rapports de Recherche n° 1893*, Avril 1993.
- [4] C. Schmid, R. Mohr, C. Bauckhage, Comparing and evaluating interest points, 6th Int. Conference on Computer Vision, Bombay, India, January 1998.
- [5] S. X. Liao, M. Pawlak, On image analysis by moments, *IEEE Trans. Pattern Anal. Match. Intell.*, 18 (3), 254-266, 1996.
- [6] A. Khothanzad, Y. H. Hong, Invariant image recognition by Zernike moments, *IEEE Trans. Pattern Anal. Match. Intell.*, 12 (5), 489-497, 1990.
- [7] M. K. Mandal, T. Aboulnasr, S. Phanchanathan, Image indexing using moments and wavelets, *IEEE Transaction on Consumer Electronics*, vol. 42, n° 3, 557-565, 1996.
- [8] R. Mukundan, K. R. Ramakrishnan, Fast computation of Legendre and Zernike moments, *Pattern Recognition*, vol. 28, n° 9, 1433-1442, 1995.
- [9] R. J. Prokop, A. P. Reeves, A survey of moment based techniques for unoccluded object representation, *Graphical Models and Image Processing*, 54 (5), 438-460, 1992.