

Indexation d'objets à niveaux de gris à partir d'une mise en correspondance de \mathcal{F} -Signatures

Salvatore TABBONE, Laurent WENDLING

LORIA, Campus Scientifique, BP 239
54 506 Vandœuvre-les-Nancy Cedex, FRANCE
tabbone@loria.fr, wendling@loria.fr

Résumé – Ce papier présente une nouvelle méthode d'indexation d'objets à niveaux de gris à partir d'un histogramme de forces. Le calcul de l'histogramme est de faible complexité et décrit une signature permettant de retrouver certaines transformations géométriques (échelle, translation et rotation). Les résultats expérimentaux attestent de l'aspect promoteur de notre approche.

Abstract – A new method of image indexing based on a histogram of forces is proposed in this paper. The histogram has a low complexity and describes a signature which allows to recover fundamental geometric transformations like the scale factor, the translation and the rotation. Experimental results show the promising start of our approach.

1 Introduction

De plus en plus de systèmes concernant la recherche d'images par le contenu sont proposés. Certains permettent de soumettre des requêtes au système par mots-clés. Cela suppose qu'au préalable chaque image de la base de donnée soit indexée manuellement. D'autres systèmes décrivent le contenu des images par des indices visuels dans l'objectif d'automatiser le processus d'indexation. Ainsi une mesure de similarité fondée sur ces indices visuels permet d'aider l'utilisateur dans sa recherche d'images spécifiques dans la base. Ces approches reposent généralement sur des techniques de traitement du signal. Typiquement les indices les plus utilisés concernent la couleur [5, 15, 19, 23], la texture [2, 9, 10, 18], ou des critères de forme [1, 13, 22]. D'autres approches se focalisent sur des points d'intérêts ou sur des approches multi-résolutions [8, 21, 24].

Dans l'objectif d'améliorer le caractère semi-automatique de l'indexation certains systèmes combinent plusieurs indices visuels [7, 16, 17, 20]. Dans la plupart des cas la complexité des méthodes est élevée et les indices utilisés pour discriminer les objets ne sont pas invariants à certaines transformations géométriques. Ce papier est dédié à l'indexation d'images par le contenu. Notre idée est de représenter les forces dominantes de la photométrie d'un objet et de tenir compte de ces déformations en fonction d'un ensemble d'orientations.

2 Définition d'une 2D F-Signature

Dans ce paragraphe, nous rappelons le principe de calcul de la \mathcal{F} -signature d'un objet de niveaux de gris par rapport à lui même, qui n'est autre qu'un histogramme de forces particulier [11, 12] décrivant les forces d'attractions entre les pixels d'un même objet. La généralisation à l'objet peut se faire en prenant en compte tous les couples de pixels. Néanmoins cela donnerait un histogramme fortement

bruité et dépendant du seuil. Pour pallier ce problème,

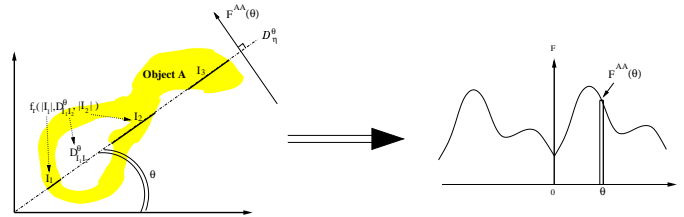


FIG. 1: Définition d'une \mathcal{F} -signature. Angle $\theta \in [-\pi, \pi[$ en abscisse et cumul des forces F en ordonnées.

uniquement des faisceaux de droites suivant un ensemble de directions θ ont été considérés. La méthode est donc fondée sur la manipulation de segments (voir figure 1), c'est-à-dire le calcul de la force d'attraction d'un segment par rapport à un autre. Prenons deux segments I_1 et I_2 portés par une même droite, ayant pour longueurs respectives $|I_1|$ et $|I_2|$ et soit $D_{I_1 I_2}^\theta$ la distance entre I_1 et I_2 . La force d'attraction de I_1 vis-à-vis de I_2 est donnée par :

$$f_r(|I_1|, D_{I_1 I_2}^\theta, |I_2|) = \int_{D_{I_1 I_2}^\theta}^{|I_1| + D_{I_1 I_2}^\theta + |I_2|} \left(\int_0^{|I_2|} \varphi_r(u-v) dv \right) du \quad (1)$$

La fonction φ est définie de \mathbb{R} dans \mathbb{R}_+ , nulle sur \mathbb{R}_- et continue sur \mathbb{R}_+^* telle que : $\forall d \in \mathbb{R}_+^*, \varphi_r(d) = 1/d^r$, où r est caractéristique du type de force que l'on recherche. Par exemple, $r = 0$ correspond à des forces constantes où la distance n'interviendra pas (histogramme d'angles) et $r = 2$ fait intervenir les forces de manière quadratique avec la distance euclidienne. Lorsque l'on manipule un objet A suivant une direction θ , on peut définir un faisceau de droites parallèles, d'angle θ par rapport à l'origine du plan, qui décrivent entièrement A . Prenons-en une, notée \mathcal{D}_η^θ , l'ensemble des segments de A portés par cette droite correspond à : $A_\theta(\eta) = \cup \{I_i\}_{i=1, n}$. L'attraction mutuelle

de ces segments est définie par :

$$F(\theta, A_\theta(\eta), A_\theta(\eta)) = \sum_{i \in 1..n} \sum_{j \in 1..n} f_r(|I_i|, D_{I_i I_j}^\theta, |I_j|) \quad (2)$$

Cela revient à prendre en considération l'ensemble des faisceaux de droites D_θ^n du plan qui décrivent entièrement A :

$$\begin{aligned} F^{AA} | \quad \mathbb{R} &\rightarrow \mathbb{R}_+, \\ \theta &\mapsto \int_{-\infty}^{+\infty} F(\theta, A_\theta(v), A_\theta(v)) dv \end{aligned} \quad (3)$$

Le calcul de $F^{AA}(\theta)$ correspond à une évaluation des forces exercées par l'objet sur lui-même dans la direction θ . Le calcul de F sur l'ensemble d'angles θ_i ($\theta_i \in [-\pi, +\pi]$) définit la signature de l'objet suivant un spectre de directions.

Dans le cadre de notre approche, nous nous intéressons à des objets à niveaux de gris. Soit $\{g_i\}_{i=1..n}$ les niveaux de gris contenus dans une image, la variation entre deux niveaux successifs est donnée par $m_i = g_i - g_{i+1}$. Prenons α_i comme étant la coupe obtenue après une binarisation de l'objet suivant un seuil g_i . Dans ce cas nous ne traitons pas exactement l'objet comme une fonction d'appartenance à l'intensité mais nous nous focalisons sur les variations des niveaux de gris.

Deux schémas de calcul peuvent être pris en compte pour intégrer les niveaux des coupes— le schéma de double somme proposé par Dubois et Jaulent [4]. Celui-ci prend en compte toutes les combinaisons possibles, au niveau des coupes, et engendre un temps de calcul élevé, — le schéma de simple somme proposé par Khrisnapuram et al. [6], qui permet de passer à un calcul linéaire. Ce dernier a été privilégié dans notre étude.

$$F^{AA} = \sum_{i=1..n} m_i F^{A^{\alpha_i} A^{\alpha_i}} \quad (4)$$

Le fait de prendre chaque coupe indépendamment permet de diminuer considérablement les cas. Par exemple le calcul de f_0 correspond à $|I_1| \times |I_2|$ quand I_1 précède I_2 et à $|I_1|^2/2$ si on considère le segment I_1 par rapport à lui-même.

3 Propriétés

3.1 Invariance aux transformations

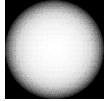
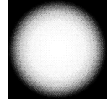
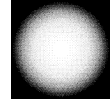
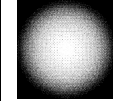
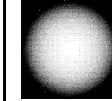
Par définition de la fonction F , nous pouvons énumérer les propriétés d'invariance suivantes [11, 12] :

- Invariance par translation parce que l'objet est traité indépendamment de sa position dans l'image.
- Invariance à la symétrie parce que les forces exercées de part et d'autre de l'objet dans une direction donnée sont les mêmes c'est-à-dire $F_\theta^{AA} = F_{\pi-\theta}^{AA}$.
- Invariance à l'échelle si et seulement si on tient compte de la forme des signatures car ces dernières sont normalisées : $f_r(kd_I, kD_{I J}^\theta, kd_J) = k^{2-r} f_r(d_I, D_{I J}^\theta, d_J)$.
- Invariance par rotation (après recalage) car l'approche est isotropique. Plus précisément, soit θ une direction donnée et soit r la rotation d'angle θ' appliquée à un objet A . Nous avons $F_\theta^{AA} = F_{\theta+\theta'}^{r(A, \theta')r(A, \theta')}$.

Ces notions sont encore vérifiées dans le cas discret avec de faibles variations dues au pas d'échantillonnage de l'histogramme. A partir de la forme de la signature d'un objet binaire, on peut facilement retrouver l'angle de rotation et le facteur d'échelle. Nous avons expérimentalement appliqué plusieurs déformations (échelle, rotation) sur un ensemble d'objets de notre base de données (cf. paragraphe 5). Les erreurs de distance entre les signatures des objets déformés et non déformés étaient en moyenne de l'ordre de 10^{-5} (similarité) avec un pas d'échantillonnage de l'orientation défini à $2\pi/128$.

3.2 Robustesse au bruit

Pour tester la robustesse au bruit de notre approche nous avons défini un ensemble d'images de synthèse d'un même objet (sphère) bruitées avec un bruit blanc de variances différentes. La table 2 indique les différences de variations, en termes de pourcentage, entre l'image originale non bruitée et les images dégradées. Nous pouvons remarquer que les faibles écarts obtenus (maximum d'écart autour de 10^{-4} entre les signatures normalisées pour un bruit de 5db) confirment expérimentalement la robustesse de notre méthode. La robustesse au bruit peut aussi être assimilée à la fonction f qui est décomposable ($f(x_1 + x_2, y, z) = f(x_1, y, z) + f(x_2, y, z)$). Donc de faibles disparités par rapport à la structure globale de l'objet ont peu d'influence sur la signature globale de l'objet. Néanmoins comme plus de segments sont traités, le temps de calcul de la signature est plus élevé quand la sphère est bruitée.

I	I_{20db}	I_{15db}	I_{10db}	I_{5db}
				
max	0.043	0.054	0.073	0.11
$\bar{\mu}$	0.035	0.044	0.060	0.091
RMS	$6 \cdot 10^{-6}$	$8 \cdot 10^{-6}$	$11 \cdot 10^{-6}$	$18 \cdot 10^{-6}$

TAB. 1: Erreurs (maximales, moyennes et quadratiques) obtenues en comparant la signature d'un objet I et des images du même objet dégradées par un bruit additif.

4 Stratégie d'indexation

La recherche d'objets similaires à d'autres se traduit par une mise en correspondance de signatures discrètes. La première étape consiste à définir le facteur d'échelle et de rotation par rapport à l'objet considéré puis calculer un rapport de similarité. Le nombre de permutations circulaires d'un histogramme qui maximise la mise en correspondance définit l'angle de rotation. L'échelle correspond à un ratio moyen entre deux histogrammes.

Pour de très grandes bases de données cette représentation peut s'avérer spatialement coûteuse et parfois imprécise. De plus, mettre en correspondance toutes les signatures stockées dans la base avec la signature d'un nouvel

objet sans aucune stratégie peut s'avérer également très coûteux en temps de calcul. Pour minimiser, la complexité spatiale et temporelle de notre approche nous approximations chaque signature par une série de Fourier.

Soit T le nombre de points d'une \mathcal{F} -Signature.

$$\begin{aligned} x_k(t) &= a_0 + \sum_{k=1}^K (a_k \cos k \frac{2\pi}{T} t + b_k \sin k \frac{2\pi}{T} t) \\ &= a_0 + \sum_{k=1}^K c_k \cos(k\omega_0 t + \varphi_k) \end{aligned} \quad (5)$$

avec

$$\begin{aligned} c_k &= \sqrt{a_k^2 + b_k^2} \\ \varphi_k &= -\tan^{-1}\left(\frac{b_k}{a_k}\right) \end{aligned} \quad (6)$$

C'est-à-dire, dans le cas continu :

$$\begin{aligned} a_k &= \frac{2}{T} \int_0^T x(t) \cos(k\omega_0 t) dt \\ b_k &= \frac{2}{T} \int_0^T x(t) \sin(k\omega_0 t) dt \\ a_0 &= \frac{1}{T} \int_0^T x(t) dt \end{aligned} \quad (7)$$

avec $\omega_0 = \frac{2\pi}{T}$ la fréquence.

Par définition la différence de phase est donnée par:

$$\Delta\varphi = -\tau 2\pi f \quad (8)$$

Pour un même objet ce déphasage représente l'angle de rotation. La distance entre deux signatures est définie par la mesure suivante :

$$\langle x_k(t), x'_k(t) \rangle = \int_0^T (x_k(t) - x'_k(t))^2 dt \quad (9)$$

5 Résultats expérimentaux

La base de données columbia est composée de 360 images représentant plusieurs vues de 5 objets différents. La figure 2 montre une vue de chaque échantillon ainsi que leur \mathcal{F} -signature respective.

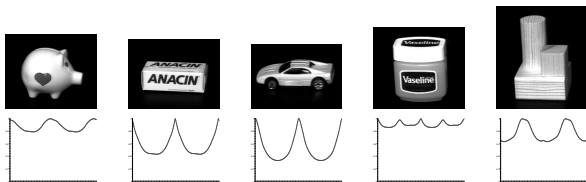


FIG. 2: Exemple de Signatures-2D pour différents objets de la base d'images de Columbia.

Nous pouvons constater que les signatures sont bien différentes sauf si les objets sont proches en forme et en photométrie. Nous avons également remarqué que l'histogramme de forces est robuste aux déformations. Les neuf images les plus proches de la requête sur l'objet du "cochon" sont présentées sur la figure 3.

Nous avons testé une autre base de données, mise à disposition par S. Derode, & al. [3] composée de 103 images de papillons décrivant plusieurs types d'espèce sous différents angles de vue. La figure 4 montre un ensemble de papillons de différents types ainsi que leurs signatures respectives sous forme de représentation continue. La figure 5 présente les neuf plus proches images, trouvées automatiquement dans la base, de la figure p15_1 avec leur représentation par série de Fourier et l'indice de similarité.

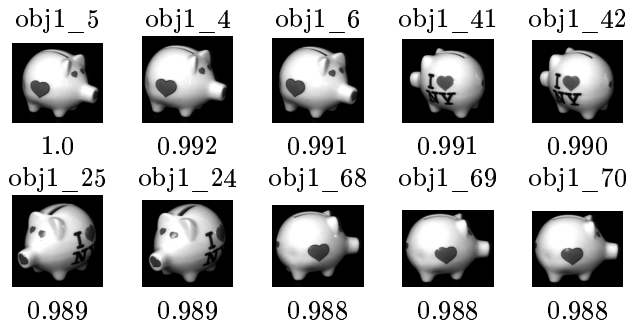


FIG. 3: Ratio de similarité entre les échantillons d'une même classe.

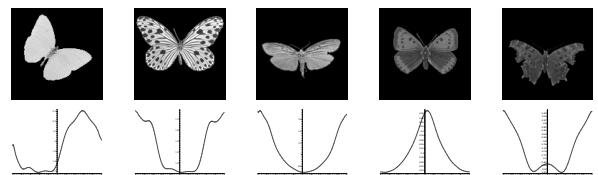


FIG. 4: \mathcal{F} -signatures pour différentes images de la base de papillons.

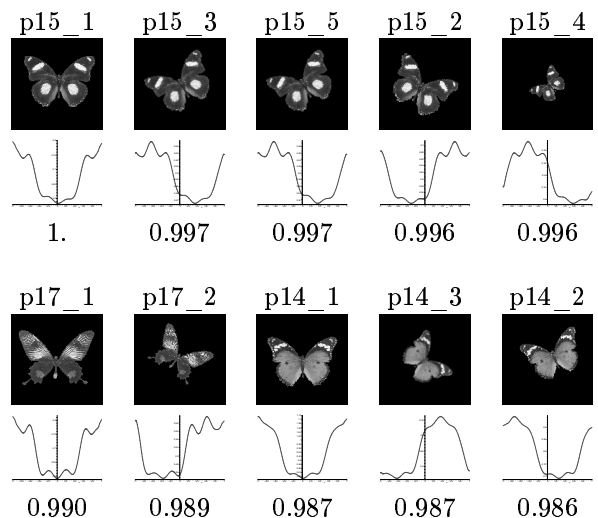


FIG. 5: Application de notre méthode sur p15_1.

6 Conclusion

Nous avons montré que la \mathcal{F} -signature est un indice discriminant de faible complexité pour indexer des images à niveaux de gris. Le rajout d'une dimension (passage à des histogrammes 3D) à la signature des objets afin d'obtenir une décomposition des forces suivant les coupes et augmenter la précision de l'indexation est en cours d'étude (voir la figure 6). De prochains travaux s'orientent également vers l'extension de la notion de \mathcal{F} -signature à des images contenant plusieurs objets.

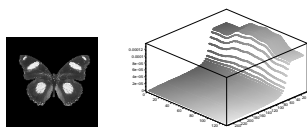


FIG. 6: Extension aux histogrammes 3D.

References

- [1] J. Bigun, S.K. Bhattacharjee et S. Michel. Orientation Radiograms for Image Retrieval: An Alternative to Segmentation, *IEEE International Conference on Pattern Recognition*, vol. C, pages 346–350, Vienne, Autriche, 1996.
- [2] Y.Q. Cen, M.S. Nixon et D. Thomas. Statistical Geometrical Features for Texture Classification, *Pattern Recognition*, 28(4):537–552, 1995.
- [3] S. Derrode, M. Daoudi et F. Ghorbel. Invariant Content-based Image Retrieval Using a Complete Set of Fourier-Mellin Descriptors, *IEEE International Conference on Multimedia Computing and Systems*, vol. 2, 877–881, 1999.
- [4] D. Dubois et M.C. Jaulent. A General Approach to Parameter Evaluation in Fuzzy Digital Pictures, *Pattern Recognition Letters*, vol. 6, 251–259, 1987.
- [5] T. Kato. Database Architecture for Content-based Image Retrieval, *In Image Storage and Retrieval Systems*, volume 1662, pages 112–123, San Jose, CA, 1992, SPIE.
- [6] R. Krishnapuram, J.M. Keller et Y. Ma. Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions, *IEEE Transactions on Fuzzy Systems*, 3(1):222–233, 1993.
- [7] C.P. Lam, J.K. Wu et B. Mehtre. STAR- a system for trademark archival and retrieval, *In Proceedings of the 2nd Asian Conference on Computer Vision*, vol. III, 214–217, Singapour, 1995.
- [8] T. Lonnestad. A New Set of Texture Features Based on the Haar Transform, *Proceedings. 11th IAPR International Conference on Pattern Recognition*, vol.III. La Hague, Pays-Bas, 1992.
- [9] B. Manjunath et W. Ma. Texture Features for Browsing and Retrieval of Image Data, *IEEE Transactions on PAMI*, 18(8):837–842, 1996.
- [10] R. Manmatha, S. Ravela et Y. Chitti. On Computing Local and Global Similarity in Images, *In proceedings of SPIE Conference on Human Vision and Electronic Imaging III*, vol. 3299, 540–551, San Jose, Etats-Unis, 1998.
- [11] P. Matsakis. Relations Spatiales Structurelles et Interprétation d'Images IRIT, thèse de l'Université Toulouse III, 1998.
- [12] P. Matsakis et L. Wendling. A New Way to Represent the Relative Position Between Areal Objects, *IEEE Transactions on PAMI*, 7(21):634–643, 1999.
- [13] R. Mehrotra et J.E. Gary. Similar-Shape Retrieval in Shape Data Management, *IEEE Computer*, 28(9):57–62.
- [14] K. Miyajima et A. Ralescu. Spatial Organization in 2D Segmented Images: Representation and Recognition of Primitive Spatial Relations, *Fuzzy Sets and Systems*, 2(65):225–236, 1994.
- [15] V.E. Ogle et M. Stonebraker. CHABOT: Retrieval from a Relational Database of Images, *IEEE Computer*, 28(9):40–48, 1995.
- [16] Y. Park et F. Ghosiani. ImageRoadMap: a New Content-Based Image Retrieval System, *In Proceedings of the 8th DEXA*, 225–239, Toulouse, 1997.
- [17] A.P. Pentland, R.W. Picard et S. Scaroff. Photobook: Tools for Content-Based Manipulation of Image-Databases, *In Proceedings SPIE Conference on Storage and Retrieval of Image and Video Databases II*, vol. 2158, pages 34–46, San Jose, Etats-Unis, 1994.
- [18] R.W. Picard et T.P. Minka. Vision Texture Annotation, *Journal of Multimedia Systems*, 3:3–14, 1995.
- [19] J.R. Smith et S.F. Chang. Single Color Extraction and Image Query, *In Proceedings of the IEEE International Conference on Image Processing*, Washington, Etats-Unis, 1995.
- [20] J.R. Smith et S.F. Chang. Querying by Color Regions Using the VisualSeek Content-Based Visual Query System, *In Mark T. Maybury, editor, Intelligent Multimedia Information Retrieval*, pp 23–41, AAAI Press, Menlo Park, 1997.
- [21] C. Schmid et R. Mohr. Local Greyvalue Invariants for Image Retrieval, *IEEE Transactions on PAMI*, 19(5):530–535, 1997.
- [22] R.K. Srihari. Use of Multimedia Input in Automated Image Annotation and Content-Based Retrieval, *In Wayne Niblack and R.C. Jain editors, Storage and Retrieval for Image and Video Databases*, pages 249–260, San Jose, Etats-Unis, SPIE, 1995.
- [23] M.J. Swain et D.H. Ballard. Color Indexing, *International Journal of Computer Vision*, 7(1):11–32, 1991.
- [24] T. P. Weldon et W. E. Higgins. Design of Multiple Gabor Filters for Texture Segmentation, *IEEE Int. Conf. Acous., Speech, Sig. Proc.*, Atlanta, vol. IV, pages 2245–2248, 1996.