

Reconnaissance et détection robuste d'objets couleur

C. MAAOUI, H. LAURENT, B. EMILE

Laboratoire Vision et Robotique, UPRES EA 2078,
ENSI de Bourges - Université d'Orléans,
10 boulevard Lahitolle, 18020 Bourges Cedex
tel : 02-48-48-40-68, fax : 02-48-48-40-40
choubeila.maaoui@ensi-bourges.fr

Résumé – Dans cet article, nous présentons une nouvelle approche de reconnaissance d'objets basée sur l'utilisation des moments Zernike appliqués sur des images couleur. La détection d'objets repose sur une approche utilisant une segmentation couleur et l'étiquetage des composantes connexes. Une méthode basée sur les séparateurs à vaste marge est utilisée pour l'étape d'apprentissage et de reconnaissance. Nous illustrons la méthode proposée sur des images couleur incluant quelques objets de la base de données COIL-100.

Abstract – In this paper, we propose an approach of object recognition based on Zernike moments applied on color images. For object segmentation, color segmentation method and the connected component labelling algorithm are used. A Support Vector Machine is used for the learning and recognition steps. Some experimental results on the COIL-100 database are presented.

1. Introduction

Un des problèmes centraux en interprétation d'images concerne le choix d'une représentation pertinente permettant d'accéder à des primitives significatives et fiables traduisant le contenu de l'image. En général, ces descripteurs sont regroupés en 3 classes : les descripteurs liés à la couleur (histogramme) [5], les descripteurs de texture (matrice de cooccurrence, indices de direction principale et de rugosité, filtres de Gabor et ondelettes) [2], [10] et les descripteurs de formes (descripteurs de Fourier et moments invariants, points caractéristiques) [6],[7],[8]. L'extraction de primitives pertinentes est un problème qui ne connaît pas de solution dans le cas général car il semble qu'il n'existe pas d'attributs qui puissent modéliser une base selon tous les points de vues. Les systèmes actuels tentent de combiner des primitives variées pour améliorer la discrimination et la classification des objets.

Nous proposons ici une approche basée sur la combinaison de deux informations : la couleur et la forme pour la détection et la reconnaissance d'objets dans une image. La section 2 présente la méthode proposée qui est composée de plusieurs étapes. La première étape est la détection d'objets basée sur la segmentation couleur et l'étiquetage en composantes connexes. La deuxième étape concerne la reconnaissance d'objets utilisant les moments Zernike appliqués sur des images couleur. Les phases d'apprentissage et de reconnaissance sont réalisées par un séparateur à vaste marge. La section 3 présente quelques résultats expérimentaux obtenus sur la base d'images COIL-100 [12].

2. Méthode développée

La méthode de détection et de reconnaissance d'objets couleur est composée des étapes suivantes :

- Segmentation couleur afin de séparer l'objet du fond complexe,
- Etiquetage de l'image pour distinguer chaque objet,
- Description de chaque objet en utilisant les moments Zernike,
- Apprentissage et reconnaissance des objets à l'aide d'un séparateur à vaste marge.

2.1 Détection d'objets

Cette partie présente l'étape de détection d'objets.

2.1.1 Segmentation couleur

Afin de segmenter l'image couleur, nous avons utilisé un algorithme efficace proposé et décrit par D. Comaniciu dans [3]. Il est basé sur l'analyse d'espace de caractéristiques. Dans ce paradigme, les pixels sont groupés et exprimés dans un espace de couleur. Chaque groupe représente une région homogène dans l'image. L'algorithme utilise la procédure «Mean Shift» ou déplacement de la moyenne pour palier aux problèmes issus des techniques de clustering. L'algorithme de segmentation proposé ici utilise l'espace des couleurs L^*u^*v (au lieu de l'espace RGB) plus robuste pour calculer les distances dans l'hypersphère qui constitue la fenêtre de «Mean shift» [3].

2.1.2 Etiquetage de composantes connexes

L'extraction des composantes connexes est largement utilisée en reconnaissance de formes pour étiqueter les régions détectées par segmentation. La technique consiste à regrouper les pixels voisins dans un ensemble appelé composante connexe. Chaque ensemble est disjoint des autres et peut ensuite être aisément isolé [4].

2.2 Reconnaissance d'objets

Un objet est tout d'abord décrit par un ensemble d'invariants qui sont ensuite utilisés dans les étapes d'apprentissage et de reconnaissance.

2.2.1 Représentation d'objet

La représentation d'objets est une étape importante pour la phase de reconnaissance. Elle doit être invariante aux transformations géométriques (rotation, translation et facteur d'échelle) et robuste face à diverses perturbations (bruit, occultation, etc). La représentation d'objets que nous avons adoptée est basée sur les moments Zernike appliqués sur des images couleur.

Les moments Zernike sont définis comme des polynômes complexes qui forment un ensemble orthogonal complet du disque unité $x^2 + y^2 \leq 1$. Présenté par Zernike (1934), cet ensemble polynomial ZP est défini comme suit :

$$\begin{aligned} ZP &= \{V_{nm}(x, y) | x^2 + y^2 \leq 1\} \\ V_{nm}(x, y) &= V(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \end{aligned} \quad (1)$$

La base polynomiale de Zernike, d'ordre n et de répétition m , est donnée par :

$$V_{nm}(x, y) = R_{nm}(x, y) \exp\left(jm \arctan\left(\frac{y}{x}\right)\right) \quad (2)$$

$n \in \mathbb{N}^+$, $m \in \mathbb{N}^+$ avec $(n-|m|)$ pair et $|m| \leq n$

Le rapport polynomial est défini comme suit :

$$R_{nm}(x, y) = \sum_{s=0}^{\frac{(n-|m|)}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2}-s\right)! \left(\frac{n-|m|}{2}-s\right)!} (x^2 + y^2)^{\frac{(n-2s)}{2}} \quad (3)$$

Soit $\rho = \sqrt{x^2 + y^2}$ la longueur du vecteur à partir de l'origine du pixel (x, y) et $\theta = \arctan(y/x)$ l'angle entre l'axe x et ce vecteur. $R_{nm}(\rho)$, la représentation en coordonnées polaires $(x = \rho \cos \theta, y = \rho \sin \theta)$ de $R_{nm}(x, y)$ est un polynôme de degré n .

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{(n-|m|)}{2}} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+|m|}{2}-s\right)! \left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \quad (4)$$

Les moments Zernike sont les projections de la fonction d'image $f(x, y)$ dans la base orthogonale de fonctions $V_{nm}(x, y)$. Le moment Zernike, d'ordre n et de répétition m , est un nombre complexe défini comme suit :

$$A_{nm} = \frac{n+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) [V_{nm}(x, y)]^* \quad (5)$$

où $x^2 + y^2 \leq 1$ et $*$ définit le complexe conjugué.

Les moments Zernike sont bien connus pour être invariants à la rotation. Une normalisation d'image est nécessaire pour rendre ces moments invariants à la translation et au facteur d'échelle [1].

Zernike sur des images binaires

Théoriquement, les méthodes de reconnaissance d'objets basées sur les moments Zernike sont appliquées sur des images binaires où la fonction $f(x, y)$ est égale à 1 ou à 0.



FIG. 1 : Image binaire de deux objets de la base Coil-100 ayant la même forme mais avec des textures différentes.

La figure FIG. 2 représente les moments Zernike (d'ordre 10 et 11) appliqués sur des images binaires des objets présentés dans la Fig.1. On peut noter que pour les deux objets, nous avons des valeurs semblables des moments Zernike.

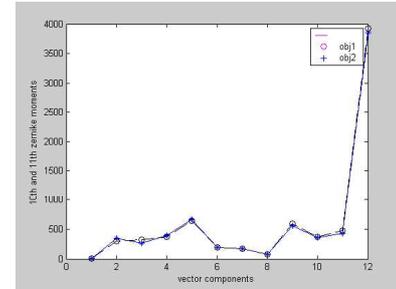


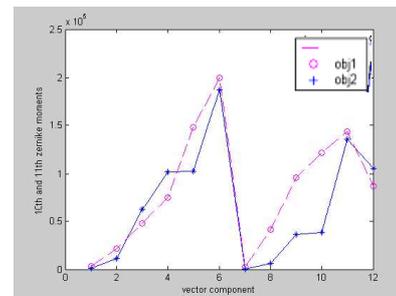
FIG. 2 : Les moments Zernike appliqués sur les images binaires de la FIG. 1.

Zernike sur des images couleur

Pour différencier les objets ayant la même forme, nous avons appliqué les moments Zernike sur des images couleur, où la fonction $f(x, y)$ est le niveau de gris du pixel auquel on ajoute la moyenne des composantes couleur RVB (rouge, vert, bleu).

$$f(x, y) = 0.3R + 0.6V + 0.1B \quad (6)$$

La figure FIG. 3 représente les moments Zernike (d'ordre 10 et 11) appliqués sur les images couleur. La distinction de deux objets de forme semblable peut alors être réalisée.



(c)

FIG. 3 : Les moments Zernike appliqués sur les images couleur de la FIG. 1.

2.2.2 Apprentissage et reconnaissance

Les moments de Zernike calculés précédemment ainsi que la classe de chaque objet forment un ensemble d'apprentissage $\{(x_i, y_i)\}_{i=1..M}$, où chaque $x_i \in \mathcal{R}^d$ et $y_i \in \{1..N\}$ dans le cas où nous cherchons à reconnaître N classes différentes. Notre objectif est de construire une fonction $f(x)$ qui estime les dépendances entre les exemples x_i et les classes y_i et qui minimise le risque d'erreur de classification pour un point x donné n'appartenant pas à la base d'apprentissage. Dans le cas d'un problème à deux classes $y_i \in [-1, 1]$, les séparateurs à vaste marge implémentent l'algorithme suivant. Tout d'abord, les points d'apprentissage x_i sont projetés dans un espace H (de dimension éventuellement infinie) à l'aide d'une fonction $\Phi(\cdot)$. Il est important de noter que pour un même ensemble d'apprentissage, différentes transformations $\Phi(\cdot)$ permettent d'obtenir différentes frontières de décision. En pratique, la transformation est réalisée de manière implicite grâce à un noyau $K(\cdot, \cdot)$ et la fonction de décision s'écrit :

$$f(\mathbf{x}) = \langle w, \Phi(\mathbf{x}) \rangle + b = \sum_{i=1}^{\ell} \alpha_i^* y_i K(\mathbf{x}_i, \mathbf{x}) + b \quad (7)$$

$$\alpha_i^* \in \mathbb{R} \quad \forall i$$

Le critère d'optimalité à maximiser dans le cas des SVM est la marge c'est-à-dire la distance entre l'hyperplan et le point $\Phi(\mathbf{x}_i)$ le plus proche de l'ensemble d'apprentissage dans le cas où $\Phi(\cdot)$ est la fonction identité. Les coefficients permettant d'optimiser ce critère s'obtiennent en résolvant le problème suivant :

$$\begin{cases} \max_{\alpha_i} \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{sous les contraintes,} \\ 0 \leq \alpha_i \leq C, \\ \sum_{i=1}^{\ell} \alpha_i y_i = 0. \end{cases} \quad (8)$$

où C est un coefficient pénalisant les exemples se trouvant dans la marge et permettant d'effectuer un compromis entre le nombre de ces derniers et la largeur de la marge. Originellement, les SVM ont été conçus essentiellement pour les problèmes à 2 classes, cependant plusieurs approches permettant d'étendre cet algorithme aux cas à N classes ont été proposées. La méthode que nous utilisons dans cette communication, est une méthode dite "un contre un" [11].

3. Résultats expérimentaux

Nous avons testé notre approche de détection et de reconnaissance d'objets couleur sur la base de données COIL-100 [12]. Cette base de données contient des images couleur de 100 objets avec 72 vues différentes pour chaque objet.

La FIG. 5 présente quelques exemples objets de la base de données, alors que la FIG. 6 présente un objet de la base pour différentes vues.



FIG. 5 : Quelques objets de la base COIL-100.



FIG. 6 : Exemple d'un objet avec différentes orientations et changements d'échelle.

3.1 Résultats de la détection d'objets

La figure FIG. 6 montre un exemple de détection d'objets de COIL-100 sur un fond faiblement texturé en combinant la segmentation couleur et l'étiquetage des composantes connexes. Le fond est détecté comme la zone la plus dominante dans l'image. Il peut être détecté par calcul de l'histogramme de l'image résultant de la segmentation couleur. Les résultats obtenus sont satisfaisants.

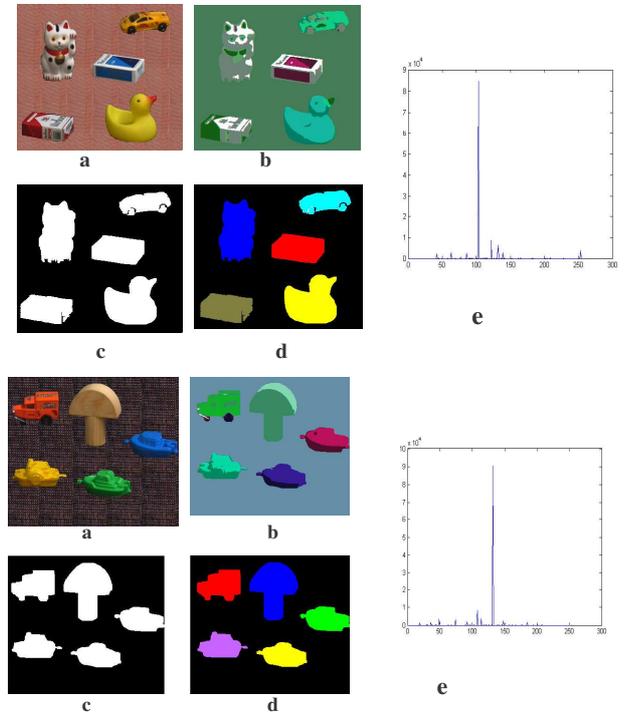


FIG. 6 : Détection d'objets : (a) image originale, (b) segmentation couleur de l'image (a), (c) binarisation de l'image (b), (d) étiquetage de l'image (c), (e) histogramme de l'image (b).

3.2 Résultats de la reconnaissance d'objets

Chaque objet détecté va être identifié en utilisant la description par les moments Zernike et la reconnaissance par SVM. Pour la reconnaissance d'objets, nous avons testé notre approche pour différentes tailles de la base d'apprentissage. Chaque échantillon d'un objet est décrit par un vecteur de dimension 73 (moments Zernike d'ordre 0 jusqu'à l'ordre 15) appliqués sur des images couleur. La figure FIG. 7 représente le taux de reconnaissance des objets de la base de données COIL-100 en fonction de la taille d'apprentissage. Dans ce cas, nous avons utilisé un pourcentage de vues de chaque objet pour l'apprentissage et le reste pour la reconnaissance. On peut noter que le taux de reconnaissance augmente considérablement quand l'information couleur est utilisée.

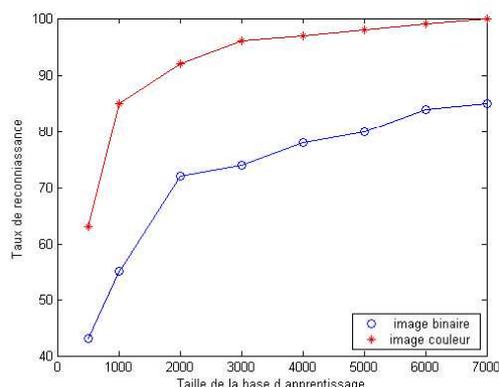


FIG. 7 : Evaluation de taux de reconnaissance en fonction de la taille de la base d'apprentissage.

Pour étudier la robustesse de notre système de reconnaissance, nous avons appliqué notre approche sur une base de données constituée de la base COIL-100 complétée par des images présentant différentes altérations, comme : le fond (noir, en niveau de gris et texturé), l'occultation partielle des objets, la variation de luminosité et la présence de bruit blanc gaussien de variance 5, 10 et 20.

Le tableau TAB. 1 présente le taux de reconnaissance pour une base d'apprentissage constituée des 72 vues de chaque objet et d'une base de test constituée de 50 vues pour chaque altération. La robustesse obtenue est excellente vis à vis de la présence de bruit, bonne dans le cas de l'occultation et de la variation de luminosité et moyenne dans le cas d'un fond bruité ou texturé.

TAB. 1 : Taux de reconnaissance (%) sur des images avec différentes altérations

N	FU	FB	FT	O-NG	O-FN	Lu	Br5	Br10	Br20
50	64	65	60	84	82	74	99.6	99.6	99.8

Avec FU : Fond Uniforme, FB : Fond Bruité, FT :Fond Texturé, O-NG : Occultation avec un fond en Niveau de Gris, O-FN : Occultation avec un Fond Noir, Lu : Luminance, Brx : Bruit blanc gaussien de variance 5, 10 et 20.

4. Conclusion

Nous proposons dans cet article un algorithme de détection et de reconnaissance d'objets basé sur les moments Zernike appliqués sur des images couleur. La combinaison de la méthode de segmentation couleur et l'étiquetage en composantes connexes permettent de détecter des objets dans une scène sur un fond faiblement texturé. Les résultats ont montré que les descripteurs de Zernike proposés utilisant l'information couleur augmentent le taux de reconnaissance d'objets ayant la même forme. En outre, cette approche est robuste face à différentes altérations : occultation, luminosité et bruit.

Dans le futur, nous envisageons d'appliquer la méthode des contours actifs pour améliorer la phase de détection d'objets.

Références

- [1] S. M. Abdallah, E.M. Nebot, et D.C. Rye " Object Recognition and Orientation via Zernike moments". In Chin, and Pong, T .C. editors, Proc. Computer Vision ACCV'98, volume 1 of LNCS 1351, pages 386-393. Springer Verlag, 1998.
- [2] L. Balmelli, A. Mojsilovic, "Wavelet domain features for texture description, classification and replicability analysis". Proceedings of IEEE International Conference on Image Processing (ICIP), vol. 4, October, pp. 440-444. 1999.
- [3] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: color image segmentation", Proceedings of Computer Vision and Pattern Recognition, pp. 750-755, 1997.
- [4] L. Di Stefano, A. Bulgarelli, "A simple and efficient connected components labeling algorithm". Image Analysis and Processing Conference, 27-29 Sept. 1999 pp : 322 – 327.
- [5] J. Hafner, H.S. Sawhney, W. Equitz, M. Flicker, W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 17, No. 7, July 1995.
- [6] M. Hu, "Visual pattern recognition by moment invariants", IRE Trans. On Information Theory, vol. IT-8, pp. 179-167, 1962.
- [7] K.Mikolajczyk and C.Schmid. "Indexing based on scale invariant interest points". In International Conference on Computer Vision, 525-531, 2001.
- [8] E. Person and K. S. Fu, " Shape Discrimination using Fourier Descriptors", IEEE Trans. Syst. Man Cybern. 7, pp, 170-179, 1997.
- [9] J. Platt, "Support Vector Machines", <http://research.microsoft.com/users/jplatt/svm.html>, 2003.
- [10]P. Kruizinga and N. Petkov, "Grating cell operator features for oriented texture segmentation", Proc. of the 14th Int. Conf. on Pattern Recognition, Brisbane, Australia, August 16-20, 1998, pp.1010-1014.
- [11] V. Vapnik, "Support-Vector Network", Machine Learning, vol. 20, issue 3, September 1995, pp. 273-397.
- [12]<http://www1.cs.columbia.edu/CAVE/research/softlib/coil-100.html>