

# Reconnaissance de formes sur des images comprimées

José Paumard<sup>1</sup>

<sup>1</sup>CMLA ENS de Cachan

61, avenue du Président Wilson

94235 Cachan Cedex

Jose.Paumard@cmla.ens-cachan.fr

## RÉSUMÉ

Le but de ce travail est de montrer qu'il est possible de mettre en œuvre des algorithmes de reconnaissance des formes sur des images comprimées, sans les reconstruire. La méthode de compression choisie procède par extraction des contours multiéchelles quinconce des images. Le problème de reconnaissance traité consiste à localiser des bâtiments sur des images aériennes de grande taille. Les appariements se fondent sur une distance de Hausdorff censurée entre contours, et la recherche d'un bâtiment se fait de façon hiérarchique, en commençant aux échelles grossières, pour se conclure aux échelles fines. Cette méthode est rapide sur station de travail classique, et permet d'atteindre le temps réel sur ordinateur parallèle, dans un cadre opérationnel réaliste. La robustesse de la reconnaissance a été soigneusement étudiée.

## ABSTRACT

The aim of this work is to prove that it is possible to recognize objects in compressed images, without reconstructing them. The compression is based on the extraction of the quincunx multiscale edges of the images. The edges are matched together, scale by scale, using a censored Hausdorff distance. Using this distance, we have been able to recognize and localize buildings in large aerial images. The search is conducted using a coarse to fine approach. This method is fast on standard workstations. Real time has been reached on parallel computer, on a realistic application. Robustness has been carefully studied.

## Introduction

La quantité d'images satellitaires ou aériennes à traiter est en croissance rapide. Deux questions peuvent alors se poser : peut-on compresser ces images, et avec quelles méthodes ? Est-il possible mettre en œuvre des algorithmes rapides de reconnaissance opérant sur la forme comprimée des images ? Nous présentons ici une méthode de reconnaissance et de localisation de bâtiments sur des images aériennes comprimées, qui ne nécessitent pas la reconstruction de ces images.

L'algorithme de compression qui a été choisi se fonde sur l'extraction des contours multiéchelles quinconce des images, nous le présentons section 1. Afin de comparer efficacement ces contours, nous avons choisi d'utiliser la distance de Hausdorff (section 2). L'algorithme multiéchelle de localisation fait l'objet de la section 3. Les performances de cet algorithme, ainsi que sa robustesse sont également exposés dans cette partie.

## 1 Compression d'images par analyse multirésolution

### 1.1 Introduction

Parmi les méthodes de compression existantes, les algorithmes fondés sur la transformée en ondelettes nous intéressent plus particulièrement. Effectivement, l'analyse multirésolution fournit un cadre théorique à l'idée commune qu'une

image contient de l'information à toutes les échelles [8]. De plus, l'utilisation de tels algorithmes va nous permettre d'utiliser les images aux échelles grossières pour guider la reconnaissance et accélérer le processus global.

Deux méthodes peuvent être adjointes à l'analyse multirésolution : la quantification vectorielle des images de détails, décrite dans [3], et le codage des contours multiéchelles, introduit dans [9]. Les contours jouent un rôle reconnu dans les mécanismes psycho-visuels de la reconnaissance des formes, aussi avons nous retenu cette seconde méthode, même si les taux de compression annoncés sont moins bons que pour la quantification vectorielle.

### 1.2 Analyse multirésolution quinconce

L'analyse multirésolution quinconce a été introduite dans [5]. C'est une analyse orthogonale avec un facteur d'échelle  $\sqrt{2}$ , qui utilise des filtres isotropes pour le lissage et le calcul des images de détails. À partir d'une image  $I_j$ , considérée à une échelle  $j$ , on obtient une image  $I_{j+1/2}$  lissée dite *de tendance*, et une image  $J_{j+1/2}$  dite *de détails*, qui contient les détails présents dans  $I_j$  et absents de  $I_{j+1/2}$ . Les deux images  $I_{j+1/2}$  et  $J_{j+1/2}$  forment une représentation complète de  $I_j$ . Le processus peut être réitéré sur  $I_{j+1/2}$  pour obtenir  $I_{j+1}$  et  $J_{j+1}$ .

Nous travaillerons ici sur six échelles. Partant d'une image originale  $I_0$ , nous travaillerons donc sur les images :  $I_{1/2}$ ,  $I_1$ ,  $I_{3/2}$ ,  $I_2$ ,  $I_{5/2}$  et  $I_3$ .

### 1.3 Extraction des contours multiéchelles quinconce

L'extraction des contours multiéchelles quinconce se fait ensuite par application d'un extracteur de contours classique, en l'occurrence Shen-Castan [11], sur chaque image de tendance. La figure 1 présente le résultat d'une telle extraction sur un des bâtiments de notre base.

À ce stade, les images sont donc disponibles sous la forme d'une pyramide d'images binaires, chacune représentant les contours de l'image à une échelle donnée.

## 2 Distance de Hausdorff Censurée

### 2.1 Définition

Parmi les algorithmes qui permettent de comparer les images de contours, la distance de Hausdorff nous a paru adaptée au problème de la localisation. Soit  $d$ , une distance du plan, cette distance a été définie dans [6] par :

$$D(a, B) = \min \{d(a, b), b \in B\} \quad (1)$$

$$h(A, B) = \max \{D(a, B), a \in A\} \quad (2)$$

$$H(A, B) = \max (h(A, B), h(B, A)) \quad (3)$$

On sait que cette distance manque de robustesse, à un point qui la rend inutilisable sur des images bruitées [4, 7, 2, 1]. Or les images que nous allons comparer sont issues d'un extracteur de contours. On ne peut donc pas garantir leur qualité. Une modification de cette distance a été introduite dans [2, 1, 10], qui permet de prendre en compte cette contrainte. Soit  $|A|$  le cardinal de  $A$ , et  $Q_p(A)$  le  $p$ ème élément de  $A$  trié dans l'ordre croissant. La modification introduite consiste à censurer *a priori* les points non significatifs des images :

$$D_p(a, B) = Q_p \{d(a, b), b \in B\} \quad (4)$$

$$h_{p,q}(A, B) = Q_{|A|-q} \{D_p(a, B), a \in A\} \quad (5)$$

$$H_{p,q}(A, B) = \max (h_{p,q}(A, B), h_{p,q}(B, A)) \quad (6)$$

Si  $h_{p,q}(A, B) = R$ , alors tous les points de  $A$  sauf  $q$  ont au moins  $p$  points de  $B$  dans un voisinage de rayon  $R$ . En particulier, si  $R$  a une faible valeur, on peut déduire que  $A$  est "bien inclus" dans  $B$ . C'est la raison pour laquelle  $h_{p,q}(A, B)$  est appelé *mesure d'inclusion* de  $A$  dans  $B$ . Si  $B$  est une image de grande taille, alors on peut définir  $h_{p,q}(A(M), B)$ , inclusion du motif  $A$  centré en un point  $M$  de  $B$ .

### 2.2 Propriétés

Cette modification fait que cette distance de Hausdorff censurée n'est plus à proprement parler une *distance*. La distance d'une image binaire à elle-même n'est pas nulle, l'inégalité triangulaire n'est plus vérifiée. On peut toutefois montrer l'inégalité suivante [10], soit  $P$  une image binaire quelconque :

$$H_{p,q}(A, B) \leq H_{p,0}(A, P) + H_{p,0}(P, B) + \min(H_{p,q}(A, A), H_{p,q}(B, B)) \quad (7)$$

Cela signifie que deux images qui ressemblent fortement à une troisième, se ressemblent entre elles dans une certaine mesure.

Les algorithmes classiques de transformée en distance [12, 13, 14] ne peuvent s'étendre simplement au calcul du  $p$ ème plus proche voisin. Un algorithme rapide de calcul de champ de distance censuré, ainsi que l'étude du gain apporté par l'introduction de cette censure ont été présentés dans [10].

## 3 Recherche et localisation de bâtiments

### 3.1 Algorithme de recherche

L'algorithme de recherche que nous avons écrit se compose de trois étapes. La scène et l'objet recherché sont connus par leur contours multiéchelles. On utilise la mesure d'inclusion de l'objet en différents points de la scène pour faire les appariements.

- Prélocalisation à l'échelle 3, la plus grossière : on forme des paquets de points en lesquels la valeur de l'inclusion est inférieure à un certain seuil. Ces points sont transmis à l'échelle plus fine suivante pour être validés.
- De l'échelle 5/2 à 1 : validation aux échelles intermédiaires. On reçoit de l'échelle précédente des paquets de points pour lesquels les valeurs d'inclusion à toute les échelles plus grossières sont connues. L'inclusion de l'objet dans la scène est alors calculée en chacun de ces points à cette échelle.
- Validation à l'échelle 1/2, la plus fine. Arrivent à cette échelle des paquets de points dans lesquels connaît les valeurs d'inclusion à toute les échelles plus grossières. On calcule enfin les mesures d'inclusion de ces points à l'échelle la plus fine.

On peut définir un seuil à chaque échelle, qui nous permet d'éliminer les points qui ont une mesure d'inclusion trop forte au fur à mesure de la recherche. Les points finalement validés à l'échelle la plus fine sont ensuite classés à l'aide de leur mesure d'inclusion, et les meilleurs sont considérés comme solution de notre problème.

### 3.2 Rapidité, robustesse

Cet algorithme a été testé sur un ensemble de 44 bâtiments extraits d'une scène aérienne de taille 3500×3500 pixels. Cette scène représente une agglomération vue sous incidence oblique à une résolution d'environ 30 cm par pixel. Les bâtiments extraits sont de taille 256×256, et ont été placés sur fond noir.

La localisation d'un bâtiment par cette méthode est réalisée en moins de 30 s sur station de travail SUN SPARC 10, et en quelques dixièmes de secondes sur SYMPHONIE, calculateur SIMD de 1024 processeurs organisés en anneau, développé par le CEA/LETI.

La robustesse de la méthode, notamment vis à vis de changements non linéaires de luminance appliqués à l'image de la

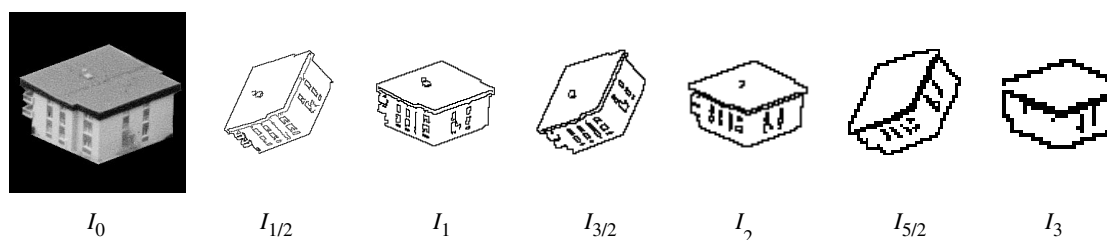


FIG. 1 — Extraction de contours multiéchelles sur un bâtiment de notre base.

scène, a été soigneusement testée. Le taux de reconnaissance ne chute que pour de fortes perturbations.

## 4 Conclusion, perspectives

Nous avons montré dans ce travail, qu'il est possible de reconnaître des bâtiments sur des scènes aériennes comprimées, sans reconstruire les images. La méthode que nous avons développée s'appuie sur des appariements de contours par distance de Hausdorff censurée. Cette distance s'est révélée robuste et rapide à calculer. Sa mise en œuvre multiéchelle nous a permis d'atteindre le temps réel dans un contexte opérationnel réaliste.

Notre travail se tourne à présent vers le prérecalage rapide d'images astronomiques de grandes tailles. Les premiers résultats obtenus dans ce domaine sont particulièrement encourageants.

## Références

- [1] Robert Azencott, François Durbin, and José Paumard. Multiscale identification of buildings in compressed aerial scenes. In *Computer Vision and Pattern Recognition, Vienna*, 1996.
- [2] Robert Azencott, François Durbin, and José Paumard. Robust recognition of buildings in compressed aerial scenes. In *International Conference on Image Processing, Lausanne*, 1996.
- [3] Michel Barlaud, Patrick Solé, Thierry Gaidon, Marc Antonini, and Pierre Mathieu. Pyramid lattice vector quantization for multiscale image coding. *IEEE Transactions on Image Processing*, 3(4), july 1994.
- [4] Marie-Pierre Dubuisson and Anil K. Jain. A modified hausdorff distance for object matching. In *International Conference on Pattern Recognition, Jerusalem*, 1994.
- [5] Jean-Christophe Feauveau. *Analyse multirésolution par ondelettes non orthogonales et bancs de filtres numériques*. PhD thesis, Université de Paris Sud Orsay, 1990. in french.
- [6] Daniel P. Huttenlocher, Gregory A. Klanderman, and Rucklidge William J. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9), september 1993.
- [7] Oh-Kyu Kwon, Dong-Gyu Sim, and Rae-Hong Park. New hausdorff distances based on robust statistics for comparing images. In *International Conference on Image Processing, Lausanne*, 1996.
- [8] Stéphane Mallat. A theory for multiresolution signal decomposition : The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), july 1989.
- [9] Stéphane Mallat and Siphon Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7), january 1992.
- [10] José Paumard. *Reconnaissance multiéchelle d'objets dans des scènes*. PhD thesis, École Normale Supérieure de Cachan, 1996. in french.
- [11] Jun Shen and Serge Castan. An optimal linear operator for step edge detection. *Computer Graphics, Vision and Image Processing*, 54, 1992.
- [12] Gunilla Borgefors. An improved version of the chamfer matching algorithm. In *7th International Conference on Pattern Recognition, Montréal*, 1984.
- [13] Gunilla Borgefors. Distance transformations in digital images. *Computer Vision, Graphics and Image Processing*, 34(86), 1984.
- [14] Gunilla Borgefors. Hierarchical chamfer matching : a parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6), november 1988.