

Application des courbes de Peano en recherche d'image par le contenu

Adel Hafiane, Bertrand Zavidovique

Institut d'Electronique Fondamentale, Université de Paris-Sud XI, 91405 Orsay, France
{prénom.nom}@ief.u-psud.fr

Résumé. Dans cet article nous présentons une application des courbes de Peano pour la caractérisation de régions par leur texture et l'établissement d'inter- relations spatiales à des fins de "CBIR". Les résultats obtenus sont comparables à ceux d'un humain sur une base de 330 images aériennes.

1 Introduction

La segmentation est un facteur clef de la recherche d'images par le contenu (CBIR). La couleur, la texture et la forme sont les principaux attributs visuels en CBIR (Flickner et al. 1995, Carson et al. 1999). Ils ne garantissent pas des réponses assez sûres. On les complète donc par des relations spatiales entre objets, supposant que l'image est organisée (Chang et al. 1987, Gudivada et Raghavan. 1995, El-Kwae et Kabuka. 1999). L'abstraction et la sémantique limitées sont ici fondées sur les régions. Compte tenu de l'application (Système d'Information Géographique) nous considérons que les régions significatives correspondent à un effort minimal d'analyse de texture, tel que peut le produire un non spécialiste. Elles sont extraites par balayages de Peano localement optimaux et co-occurrence de gradients constante, liés à la couleur (moyenne et variance R,V,B presque constantes). Les relations spatiales exploitent la droite inter-centres. Le tout est représenté par un graphe (ARG) (Hafiane et al. 2004). Le calcul des liens entre régions exploite le même type de représentation, renforçant le caractère sémantique limité de la distance entre images. Notre étude recouvre 1) caractérisation des régions extraites par l'humain et vérifiées efficaces pour la requête, 2) conception d'une segmentation automatique en régions basée sur les mêmes variables 3) vérification expérimentale de "rappel" et "précision" comparables pour l'automate et l'humain. Le caractère naturellement multi-résolution de notre segmentation/codage d'images fournit un degré de liberté supplémentaire pour adapter la méthode à des classes d'images plus spécifiques en cas de besoin. Nous rappelons brièvement dans la suite ce qu'est un balayage de Peano et la structure d'arbre associée puis le calcul de distance dans l'arbre. Nous expliquons la segmentation par "Fuzzy C-means" (FCM) modifié avant de conclure sur quelques résultats comparant l'homme et la machine.

2 Les Courbes de Peano

Les courbes de Peano appliquent récursivement les points d'un espace (e.g. image) sur une courbe (e.g. balayage) en les ordonnant en séquence de motifs (figure 1 a). Etant donné un entier $x \in \mathbf{Z}$, son code de Peano "en \mathbf{Z} " dans le plan est :

$$p_2(x) = 2^{2[\log_2 x]} + p_2(x - 2^{[\log_2 x]}) \quad \text{avec} \quad p_2(0) := 0 \quad (1)$$

pour un point de l'image $(x, y) \in \mathbf{Z}^2$, son indice dans le parcours de Peano "en \mathbf{Z} " est :

$$p(x, y) = p_2(x) + 2p_2(y) \tag{2}$$

La séquence de Peano "en \mathbf{Z} " sous tend naturellement un arbre binaire, le \mathbf{Z} -arbre, dont les noeuds pointent alternativement des blocs rectangulaires et carrés de l'image (figure 1 b). Cette structure se généralise en $*$ -arbre en variant les parcours, et toutes formules et transformations à l'avenant (Seetharaman et Zavidovitch. 1997).

Dans un tel $*$ -arbre, les relations spatiales dérivent des blocs englobant les objets définis avec la précision requise, i.e. niveau dans la pyramide (figure 1 c). La plus simple dépendance invariante par transformation géométrique est l'inter distance associée à l'angle relatif. Le vecteur des centres de gravité des régions est donc un attribut spatial adéquat. On applique le parcours de Peano déterminé par ailleurs et on construit le $*$ -arbre de l'image. La distance entre noeuds respectifs contenant les régions est une différence de leurs indices.

3 Caractérisation de la texture

Les parcours de Peano peuvent également caractériser la texture e.g. distribution spatiale de motifs optimaux codée par matrice de leur co-occurrence.

3.1 Les parcours de Peano localement optimaux

Différents balayages engendrent différentes relations d'ordre entre pixels le long de la séquence 1D, modulant les fréquences présentes dans l'image (voir fig 1 a). Nous tirons avantage de cette propriété pour construire un chemin minimisant les transitions de niveaux de gris. Sur $\begin{matrix} p_1 & p_2 \\ p_3 & p_4 \end{matrix}$ le scan optimal suit la permutation r^* qui correspond à :

$$g^* = \min_r \{ |p_{r1} - p_{r2}| + |p_{r2} - p_{r3}| + |p_{r3} - p_{r4}| \} \tag{3}$$

On considèrera d'autres critères selon la sémantique de segmentation ou l'application.

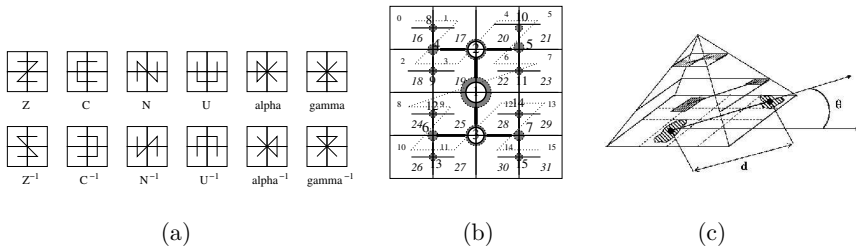


FIG. 1 – a) Motifs sur 4 pixels b) Z arbre 4 x 4 c) pyramide inter-régions

3.2 Matrice de co-occurrence des Motifs

L'image est divisée en blocs de 2×2 pixels traversés chacun selon son parcours optimal. Une image de taille $M \times N$ est donc un tableau de $M/2 \times N/2$ motifs. Leur distribution spatiale est résumée par la matrice de co-occurrence de motifs (MCM), dont l'élément (i, j) est la probabilité de trouver le "motif i " à une distance donnée du "motif j ". Le nombre de motifs étant 12 le vecteur caractéristique est de dimension 12×12 donc inférieure à celle plus classique de la co-occurrence en niveaux de gris.

4 La Segmentation

Notre segmentation en régions est une classification des pixels variante de la méthode "Fuzzy c-means" (Bezdek, 1981) avec contrainte spatiale. Ici la fonction coût est :

$$J_M(U, V) = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m \|x_j - v_i\|^2 + \alpha \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m e^{-\sum_{k \in \Omega} u_{ik}^m} \quad (4)$$

Le second terme tend à forcer un pixel dans la classe de ses voisins, d'où régularisation i.e. limitation a priori du nombre de régions (Ω : voisinage, u_{ik} : appartenance / $k \neq j$). $J_M(U, V)$ est optimisée classiquement via le Lagrangien.

5 Recherche d'image

Les images de la base sont segmentées, on calcule la MCM pour chaque région ainsi que les relations spatiales. La recherche est une mise en correspondance de graphes : on mesure la similarité visuelle par comparaison exhaustive des MCMs. On obtient un ensemble de m ($= \min\{\text{nombre régions (image requête IR)}, \text{nombre régions (image de la base IB)}\}$) couples de régions (R_i, R'_i) . Les images obtenant le meilleur score pour l'aspect visuel sont gardées dans la deuxième phase de similarité spatiale, avec distance entre sous graphes (Hafiane et al. 2004). Soit $d(\cdot)$ et $\delta(\cdot)$ deux distances entre MCMs et entre faisceaux d'arcs respectivement, E_i^k et $E_i^{k'}$ des arcs inter régions homologues

$$D(IR, IB) = \frac{1}{m} \sum_{i, k(i)} (d(R_i, R'_i) + \alpha \delta(E_i^k, E_i^{k'})) \quad (5)$$

Les tests ont concerné 300 images aériennes (<http://terraserver.microsoft.com>). La base est divisée en 35 groupes, chacun de 6 à 10 images visuellement similaires ou connues géographiquement pour telles. Une image résultat est dite pertinente si elle appartient au groupe de sa requête. Un résultat de recherche est donné figure 2 a. La requête est en haut à gauche et le score décroît de gauche à droite et de haut en bas. Les images 5 et 6 sont des intrus. La segmentation en régions peut biaiser le processus de recherche. Nous avons étudié son influence en comparant les réponses aux mêmes requêtes après segmentation automatique et segmentation humaine par un non spécialiste. On mesure les performances respectives par les courbes classiques "rappel" = $f(\text{"précision"})$. Pour comparer les performances, on calcule la différence normalisée entre les deux courbes. Une centaine d'images requête a été testée, l'histogramme

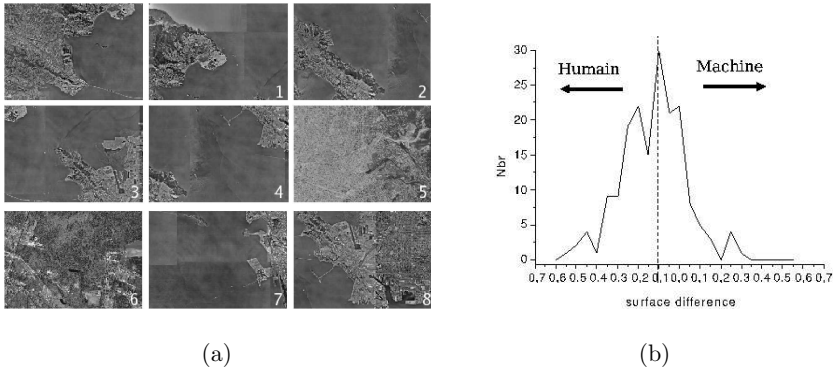


FIG. 2 – a) Résultat d'une requête b) histogramme de comparaison homme/machine

des différences est donné figure 2 b. On constate que les performances de recherche après segmentation humaine ou automatique sont comparables, si bien que la chaîne de requête peut être entièrement automatisée y compris la segmentation. Le niveau de compétence du programme est comparable à celui d'un non spécialiste.

Références

- Bezdek J.C. (1981), Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York.
- Carson C., Thomas M., Belongie M., Hellerstein J.M. et Malik J. (1999), Blobworld : a system for region-based image indexing and retrieval, 3rd ICVIS, pp. 509-516.
- Chang S.K., Shi Q. Y. et Yan C. W. (1987), Iconic indexing by 2-D strings, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, pp. 413-428.
- El-Kwae E. A. et Kabuka M. (1999), A robust framework for content-based retrieval by spatial similarity in image databases, ACM Trans. Inf. Sys, vol. 17, pp. 174-198.
- Flickner M. et al. (1995), Query by image and video content : The QBIC system, IEEE Computer, vol. 28, pp. 23-32.
- Gudivada V.N. et Raghavan V.V. (1995), Design and evaluation of algorithms for image retrieval by spatial similarity, ACM Trans. Inf. Sys, vol. 13, pp. 115-144.
- Hafiane A., Chaudhuri S., Seetharaman G. et Zavidovique B. (2004), Image retrieval using *-trees for GIS applications, 8th WCSLI, Orlando, USA, Jul 2004.
- Seetharaman G. et Zavidovique B. (1997), Image Processing in the Tree of Peano Traversed Images, IEEE CAMP, Boston, Oct 1997, pp. 19-21.

Summary

The paper deals with region based CBIR. Extracted regions from aerial imagery are characterized by the texture through Peano coding images. The latter coding supports spatial relationship too. Retrieval results are shown comparable to those of non expert humans.