

Histogrammes spatiaux couleur optimisés pour l'indexation d'images par le contenu

S. KHARBOUCHE, P. VANNOORENBERGHE, C. LECOMTE, P. MICHE

Laboratoire Perception Systèmes Information, FRE 2645 CNRS

Université de Rouen, UFR des Sciences

Place Emile Blondel, 76821 Mont Saint Aignan cedex, France

{Said.Kharbouche, Patrick.Vannoorenberghe, Christele.Lecomte}@univ-rouen.fr, Pierre.Miche@insa-rouen.fr

Résumé – Dans cet article, nous présentons une méthode qui augmente l'efficacité et la précision des systèmes d'indexation et de recherche d'images couleur par le contenu basés sur les histogrammes spatiaux. Cette technique consiste essentiellement à extraire les informations spatiales de l'image en la divisant en plusieurs sous-images rectangulaires interférées (grille des histogrammes). La méthodologie proposée optimise, via une combinaison linéaire, le découpage horizontal et vertical de l'image pour le calcul des histogrammes spatiaux. Nous augmentons ainsi la performance de la recherche en réduisant en même temps le coût de stockage et le temps de calcul par rapport aux méthodes classiques dont le choix concernant le découpage reste arbitraire.

Mots clef : Indexation d'images, Recherche d'image par le contenu, Couleur, Histogramme spatial.

Abstract – In this paper, a methodology that improves efficiency and accuracy of content-based image retrieval (CBIR) systems is presented. It is based on spatial histograms and used for colour images databases. This technique primarily consists in extracting spatial information from the image by dividing it horizontally and vertically (colour histograms grid). With a linear combination, the number of horizontal and vertical splits is optimised. It allows us to increase the performance of the CBIR system and reduce the storage cost and computing time generated by other methods based on splitting the image into equal sub-images.

Keywords: Image indexing, Content based image retrieval, Colour, Spatial histogram;

1. Introduction

L'indexation d'une base d'images a été envisagée dans un premier temps en associant à chaque image des commentaires textuels généralement sous forme de mots-clefs. Ces dernières années, une autre voie consiste à indexer et rechercher directement les images à partir de leurs contenus. Cette deuxième technique est plus intéressante parce qu'elle est indépendante des thésaurus associé aux commentaires, plus riche que le texte et permet une description moins subjective que le texte et souvent plus efficace. Cette technique semble bien adaptée à plusieurs domaines d'application (albums personnels, images biomédicales, imagerie satellitaire, télé-surveillance, média, ...).

L'indexation de la base étant basée sur le contenu des images, on étudie la similarité perceptuelle entre les images, l'image pouvant alors être caractérisée par trois attributs : la forme, la texture et la couleur [10]. L'extraction de ces attributs pour chacune des images de la base permet d'obtenir une signature de chaque image. Avant de calculer cette signature, il faut souvent résoudre les problèmes du choix de l'espace couleur, de la quantification de l'image et des traitements à effectuer au préalable sur les images :

Choix de l'espace couleur : Il n'y a pas d'espace couleur idéal pour l'indexation. Les espaces couleur les plus utilisés sont RGB (Red, Green, Blue), HSV (Hue, Saturation,

Value), Lab ou Luv (L : luminance, (a,b) et (u,v) représentent la chrominance). Quelques auteurs proposent d'utiliser des espaces couleur hybrides [3].

Quantification de l'image : Elle consiste à réduire le nombre de couleurs présentes dans l'image afin de ne conserver uniquement que les couleurs dominantes. Cette étape facilite la segmentation et le calcul des signatures. On distingue deux catégories de quantification, la quantification statique (uniforme ou non-uniforme) et la quantification dynamique (classification).

Pré-traitements : Cette étape est abordée surtout si on veut extraire des informations de forme. Elle peut permettre également d'éliminer le bruit dans l'image, de localiser et déterminer des frontières de zones ou des points d'intérêt dans l'image.

Dans cet article, nous nous intéressons uniquement à l'attribut couleur du fait de son haut pouvoir discriminant par rapport aux autres attributs mais l'approche reste concevable pour tous autres attributs extraits de l'image. Nous travaillons dans l'espace HSV qui s'avère plus dédié à l'évaluation de la similarité perceptuelle entre deux couleurs. On réalise sur les images de la base une quantification statique et uniforme sur les trois axes de l'espace HSV. La section 2 décrit les histogrammes spatiaux couleur ainsi que leur utilisation pour l'indexation d'images par le contenu. Des simulations sur une base d'images généraliste sont ensuite présentées dans la section 3 de cet article.

2. Histogrammes spatiaux couleur

Une des particularités de l'utilisation de l'histogramme en tant que signature réside dans le fait que l'image reste invariante par translation ou rotation. Ce grand avantage est obtenu par rapport aux autres approches au détriment des informations spatiales qui ne sont pas prises en considération par l'histogramme. Pour réduire les effets de cet inconvénient majeur, plusieurs travaux s'attachent à calculer les proportions de couleurs et leurs positions relatives par exemple la méthode CRT [4], spatial-chromatic histograms [5], color co-occurrence matrix [6]. Il y a également d'autres méthodes qui consistent à diviser l'image en sous-images par exemple en 9 rectangles égaux [7], en 25 rectangles égaux [3]. Après décomposition, le système associe à chaque sous-image son histogramme couleur correspondant. Avec cette technique, la dimension du vecteur des indices sera multipliée par le nombre de rectangles, ce qui augmente la taille des signatures (espace de stockage) dans la base indexée et donc allonge le temps de calcul. Un autre inconvénient généré par ces méthodes concerne la perte d'invariance de l'image. Un exemple de ce type d'inconvénient est illustré à la figure 1 où un système de découpage en 3x3 rectangles donne une image (c) plus proche de l'image (a) que de l'image (b) !

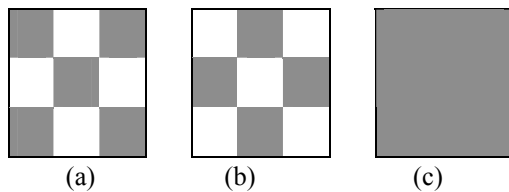


FIG 1 : Le système basé sur la division de l'image en 9 rectangles estime que l'image (c) est plus proche de l'image (a) que de l'image (b).

La méthode que nous proposons cherche à réduire la taille de la signature (par conséquent le temps de recherche dans la base) et à augmenter l'invariance de l'image. Notre méthode consiste à subdiviser l'image quantifiée (en c couleurs différentes) en m parties horizontales et en n parties verticales. A partir de ce découpage, les signatures associées à chaque sous-image sont calculées. L'effet obtenu en divisant l'image verticalement, est l'invariance de l'image par translation verticale. L'effet obtenu en divisant l'image horizontalement, est l'invariance de l'image par translation horizontale. Cet avantage est illustré sur la figure 2. Cette technique de division d'image donne une certaine souplesse au système pour mesurer la similarité perceptuelle par rapport aux méthodes classiques de découpage, parce que plus le nombre de rectangles est grand, plus l'image reste invariante. De plus, les intersections entre les n zones verticales et les m zones horizontales fournissent une idée sur la proportion des couleurs dans les nm rectangles qui forment l'image. Alors que la division classique de l'image en n^2 rectangles augmente n^2 fois la taille de la signature, la subdivision de

l'image en n parties horizontales et en m parties verticales augmente $(m+n)$ fois la taille de la signature.



FIG 2 : Illustration de l'invariance du découpage horizontal pour le calcul de la signature d'une image.

Pour la mesure de similarité, le système fait une pondération entre les parties horizontales et les parties verticales, les découpages verticaux restant les plus intéressants [8][4], surtout si on travaille avec des bases généralistes. Ainsi, la similarité entre deux images I_i et I_j est évaluée par le calcul de la distance suivante :

$$d(I_i, I_j) = \alpha \sum_{k=1}^m f(V_{i,h(k)}, V_{j,h(k)}) + \beta \sum_{q=1}^n f(V_{i,v(q)}, V_{j,v(q)}) \quad (1)$$

où $(\alpha, \beta) \in [0,1]^2$ tels que $\alpha + \beta = 1$ représentent les coefficients de pondérations entre les parties verticales et les parties horizontales des deux images. Dans cette équation, $V_{i,h(k)}$, respectivement $V_{j,v(q)}$, représentent les vecteurs des indices calculées sur la k -ème partie horizontale $I_{i,h(k)}$ de l'image I_i et respectivement la q -ème partie verticale $I_{j,v(q)}$ de l'image I_j . Une caractéristique importante d'un système de recherche concerne la fonction de similarité entre les images (les parties d'images dans cette approche) c'est-à-dire la fonction f définie par $f: (\text{espace d'indice})^2 \rightarrow [0,1]$. Le choix de cette fonction sera discutée dans la partie suivante.

3. Simulations

Dans cette section, nous proposons d'évaluer les performances de la méthode proposée en la comparant avec une méthode classique de division des images en plusieurs rectangles. La base de test à indexer contient 1115 images classées en 20 catégories sélectionnées à partir de la base COLUMBIA :

www.cs.washington.edu/research/imagetdatabase/groundtruth/

Cette base généraliste comporte essentiellement des images de paysages naturels ou des lieux habités. Pour cette raison, nous avons choisi de diviser les images horizontalement et verticalement en $m=n=3$ parties horizontales et verticales (ciel, objets, terre). En ce qui concerne la signature des images, nous avons travaillé uniquement sur l'attribut couleur. La quantification est faite dans l'espace colorimétrique HSV et est statique et uniforme sur les trois axes avec : 18 niveaux pour l'axe Hue (teinte), 3 niveaux pour l'axe Saturation et 3 niveaux pour l'axe Value (intensité). Ceci permet de quantifier l'image sous 162 couleurs différentes. Ce choix qui reste arbitraire peut être discutable mais a montré son efficacité dans plusieurs travaux [9],[4],[3]. Ainsi, chaque image I_i est représentée

par un vecteur de dimension 972 : soit le nombre des couleurs (=162) multiplié par le nombre d'histogrammes (=6). La similarité entre deux sous-images horizontales $V_{i,h(k)}$ et $V_{j,h(k)}$ extraites respectivement des images I_i et I_j est calculée par la fonction [8] :

$$f(V_{i,h(k)}, V_{j,h(k)}) = \frac{\sum_{l=1}^c \min(p_{i,h(k)}^l, p_{j,h(k)}^l)}{\min(\sum_{l=1}^c p_{i,h(k)}^l, \sum_{l=1}^c p_{j,h(k)}^l)} \quad (2)$$

Cette même équation peut être utilisée pour les parties verticales de l'image permettant ainsi de quantifier la similarité entre $V_{i,v(q)}$ et $V_{j,v(q)}$:

$$f(V_{i,v(q)}, V_{j,v(q)}) = \frac{\sum_{l=1}^c \min(p_{i,v(q)}^l, p_{j,v(q)}^l)}{\min(\sum_{l=1}^c p_{i,v(q)}^l, \sum_{l=1}^c p_{j,v(q)}^l)} \quad (3)$$

Dans les équations (2) et (3), $p_{i,h(k)}^l$ (resp $p_{i,v(q)}^l$) représente la proportion de la couleur l dans la sous-image $I_{i,h(k)}$ (resp. $I_{i,v(q)}$). Ainsi le vecteur $V_{i,h(k)}$ (resp. $V_{i,v(q)}$) représente l'histogramme de la sous-image $I_{i,h(k)}$ (resp. $I_{i,v(q)}$). Cette fonction de similarité varie entre 0 et 1. Plus les deux sous-images sont similaires, plus f est faible ($f(V_{i,h(k)}, V_{i,h(k)})=f(V_{i,v(q)}, V_{i,v(q)})=0$).

Nous avons testé cette approche sur la base généraliste COMLUBIA en utilisant, pour mesurer la similarité entre les images, le système d'équations (1), (2) et (3) et en ajustant les paramètres de la façon suivante :

$m=n=3$, $c=162$, $\alpha=0.5$ et $\beta=0.5$.

Les figures 3 et 4 montrent quelques résultats des tests, où l'on peut constater facilement que la division de l'image en sous-images horizontalement et verticalement améliore la performance de la recherche. La figure 3 illustre les résultats obtenus pour deux requêtes respectives (image en haut à gauche de chaque planche d'images). On compare respectivement de manière qualitative les images retrouvées par le système basé sur la division en 9 rectangles (images de la planche a) et la grille des histogrammes (images de la planche b). De manière visuelle, on constate que les images retrouvées par le système basé sur la grille des histogrammes est beaucoup plus robuste et permet de retrouver un nombre plus important d'images de la même catégorie. Ceci est confirmé par les courbes ^précision / rappel présentées à la figure 4 pour trois catégories respectives de la base d'images.

Ces résultats montrent clairement que la grille des histogrammes est plus discriminante que le module de sous-images non-interférées, malgré la non-pondération ($\alpha=0.5$, $\beta=0.5$) afin de ne pas passer outre le choix de la division qui est généralement évincée dans les autres méthodes.

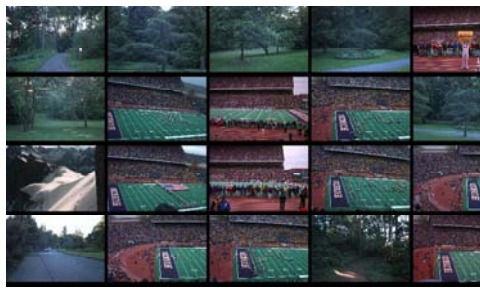
4. Conclusion

Dans cet article, nous avons présenté une méthode qui augmente l'efficacité et la précision des systèmes d'indexation et de recherche d'images couleur par le contenu basés sur les histogrammes spatiaux. Cette technique consiste à diviser l'image en plusieurs blocs rectangulaires interférés et à en calculer leurs histogrammes couleur respectifs. Ceci permet d'introduire des informations spatiales avec peu de contraintes dans les processus d'indexation et de recherche. Nous augmentons ainsi la performance de la recherche en réduisant en même temps le coût de stockage et le temps de calcul par rapport aux méthodes classiques dont le choix concernant le découpage reste arbitraire

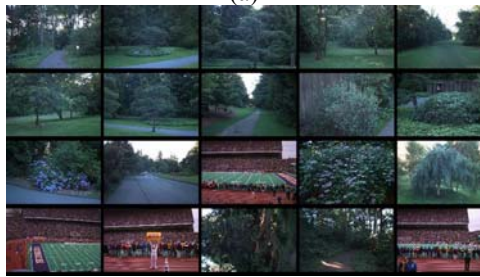
Optimiser les coefficients de pondération et le nombre de division en fonction du type d'image à indexer constituent les perspectives de ce travail.

Références

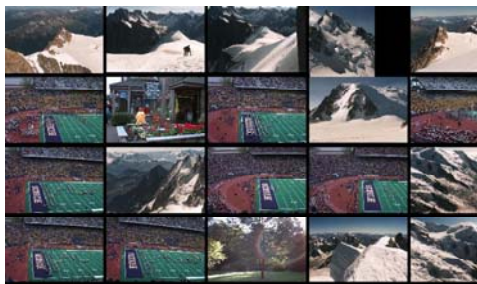
- [1] Arnold W.M and al. Content-Based Image Retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 22, N° 12. December 2000.
- [2] R.Schettini and al,"A survey of methods for colour image indexing and retrieval in image databases", (R. Luo, L. MacDonald eds.), J. Wiley, 2001. (invited paper).
- [3] Hum-Woo Yoo and al, "Visual information retrieval via content-based approach" Pattern Recognition 35 (2002) 749-769
- [4] J.R Smith and S.F.Chang Li,"Decoding image semantics using composite region templates", IEEE Workshop, CBAIVAL 1998.
- [5] L.Cinque and all. Color-based retrieval using spatial-chromatic histograms. Image and Vision Computing 19(2001) pp979-989.
- [6] Seong-Oshim and al,"Image Indexing by modified Color co-occurrence matrix", ICIP 2003
- [7] Gong Y,"Image Indexing and retrieval using color histogram", Multimedia Tools and Applications, Vol 2, pp 133-156, 1996
- [8] Jérôme FOURNIER, Indexation d'images par le contenu et recherché interactive dans les bases généralistes. Thèse : Université de Cergy -Pontoise 2002.
- [9] J.R Smith and S.F.Chang, "A fully automated content-based image query system", ACM MM 1996, pp 87-98
- [10] A.Trémeau, C.Fernandez-Maloigne, P.Bonton, "Image couleur :de l'acquisition au traitement",Editions DUNOD 2004.



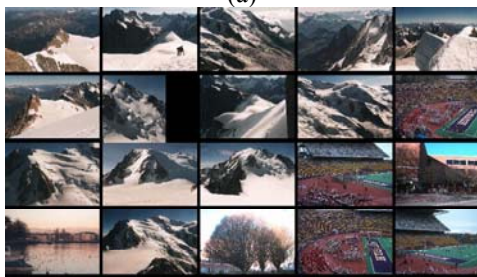
(a)



(b)



(a)



(b)

FIG 3 : Images résultats d'une requête (image en haut à gauche) donnée par les méthodes division en 9 rectangles (a) et la grille des histogrammes (b).

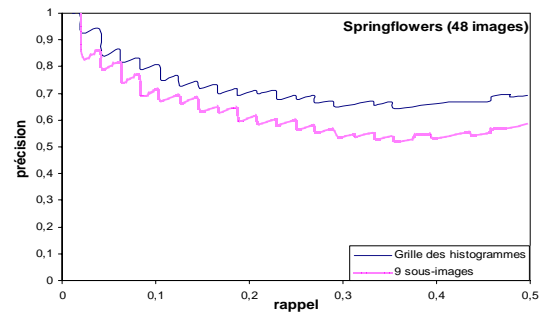
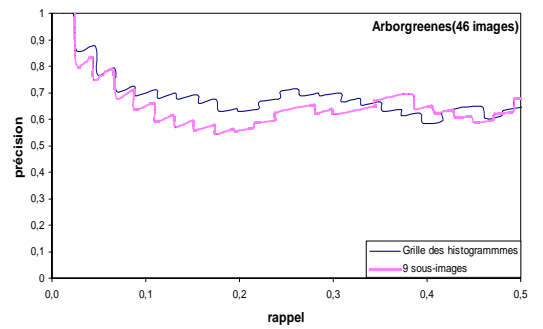
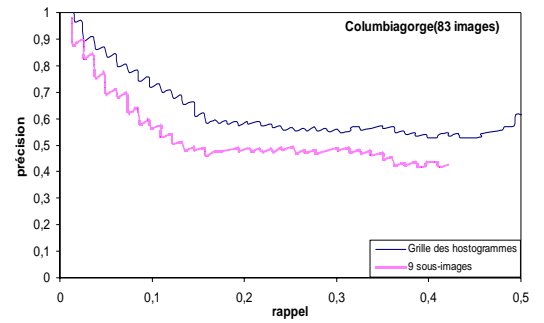


FIG 4 : Courbes précision / rappel pour trois catégories d'images dans la base COLUMBIA