Available online at www.sciencedirect.com
SCIENCE DIRECT
Pattern Recognition Letters 27 (2006) 259–267
Pattern Recognition Letters
www.elsevier.com/locate/patrec

# Region-based CBIR in GIS with local space filling curves to spatial representation

Adel Hafiane [a,*], Subhasis Chaudhuri [b], Guna Seetharaman [c], Bertrand Zavidovique [a]

[a] Institut d'Electronique Fondamentale, Université de Paris-Sud 11, Bât 220, 91405 Orsay Cedex, France
[b] Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400 076, India
[c] Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433-7765, USA

Available online 30 September 2005

**Abstract**

In this paper we present a region-based retrieval method for satellite images using motif co-occurrence matrix (MCM) in conjunction with spatial relationships. Each image is decomposed into coherent segments, MCM is computed for each region and the spatial relationship among them are evaluated by using a $*$-tree representation. The image is represented by an attributed relational graph (ARG) where nodes contain the visual feature (MCM) and edges represent spatial relationship. Principal component analysis show the usefulness of MCM as a feature.


*Keywords:* Motif co-occurrence matrix; Peano scan; Region-based image retrieval; Spatial relationship; $*$-tree; Fuzzy $c$-means clustering

## 1. Introduction

The rapid growth of computational power and storage has resulted in a large number of applications which use images. Among these applications, satellite imaging and geographical information system (GIS) applications can be particularly noted for using and storing huge amounts of digital image data. Retrieving such images by content in data bases has lately triggered an important research. Visual features constitute the main information for content-based image retrieval (CBIR). Most common approaches are based on low level image features to extract visual information such as color histogram, texture, and shape (Rui et al., 1999). Color histogram is the most popular since it is simple and effective (Swain and Ballard, 1991). Current advanced commercial systems such as QBIC, Virage and Photobook (Flickner et al., 1995; Bach et al., 1996; Pentland et al., 1996) tend to combine several features to obtain a more efficient CBIR. However these visual features are computed over an entire image and do not take into account any image semantics. On the other hand, there are some systems such as Netra (Ma and Manjunath, 1999) and Blobworld (Carson et al., 1999) which segment the images into meaningful regions to improve the results in a semantic sense. They require users to point out a set of regions of interest. The feature vectors of these regions are compared over the database. However, region-based features alone do not prove to be efficient enough to handle precise queries in a large image database. Spatial relationships provide a limited amount of semantic information for relative object positions in the image. Although it is more of a structure than semantics, it supports a more accurate CBIR when used in conjunction with textural features. Methods employing the criterion include 2D-string matching (Chang et al., 1987), $\theta R$-strings based on directional relations (Gudivada and Raghavan, 1995), and the topological extension method proposed in (El-Kwae and Kabuka, 1999) aimed at increased robustness. Petrakis et al. (1997) prefer the Attributed Relational Graph (ARG) that demonstrates a better precision and recall performance in medical images. Texture is an important feature in satellite images. Ma and Manjunath (1998) show the effectiveness

* Corresponding author. Tel.: +33 01 69 15 78 07; fax: +33 01 69 15 40 00.
*E-mail address:* adel.hafiane@ief.u-psud.fr (A. Hafiane).

0167-8655/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.08.007

of texture in aerial image retrieval and browsing. An interesting texture feature description technique is motif co-occurrence matrix (MCM) based on Peano space filling curves. Its efficiency for image retrieval has been demonstrated in (Jhanwar et al., 2004). Further, the MCM has been shown to be invariant to illumination changes a useful property for analyzing remote sensing images. However the method is used for global feature extraction. In the case of real world images there is frequently a variety of textures present in a single image. In such a case the global MCM technique would fail since it assumes a common global texture in an image.

In this paper we present a region-based CBIR technique using local space filling curves to extract relevant local information in images leading to computation of the motif co-occurrence matrix (MCM) for texture description. Spatial relationship is combined with the visual features for extended semantics. A new method of image segmentation based on MCM is proposed to handle region-based CBIR. During the retrieval process each region is assigned its own MCM which is used as a feature vector. Relations among regions are also quite simple: they stem from the regular tree structure binding them into square blocks at different resolutions. The retrieval accuracy is improved by combining the local and global properties through the use of Peano motifs.

The outline of the paper is as follows: feature extraction is discussed in Section 2, and Section 3 describes properties of MCM. In Section 4 we give a description of the segmentation method. Section 5 describes the retrieval process. In Section 6 we discuss results. The paper concludes in Section 7.

## 2. Feature extraction

### 2.1. Motif co-occurrence matrix

As in any other pattern recognition problem, CBIR is greatly affected by the choice of features. In the case of aerial images, lighting conditions for instance vary depending on the weather. That causes the pixel color to be unreliable and the popular color histogram method fails. Texture conveys an important information about structural arrangements of surfaces. It turns out to be a major feature in satellite images. We focus on local textural properties, enhanced by family of motifs shown in Fig. 1. The inner similarity between sub-regions was found to be well captured by co-occurrence matrices of such motifs (MCMs). The Peano space filling curves are used to traverse pixels along a specific local path made out of the six primitive scans represented in Fig. 1.

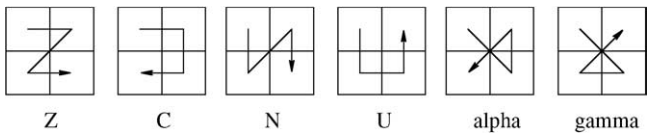


Fig. 1. Illustration of scan motifs to traverse a $2 \times 2$ grid.

One of the six possible motifs will represent the $2 \times 2$ pixel grid optimally with respect to a suitable criterion. For instance in the present application, the relevant motifs minimize the local intensity variation along the scan line. More precisely, given $\begin{matrix} p_1 & p_2 \\ p_3 & p_4 \end{matrix}$, the optimal scan follows the permutation $\alpha^*$ corresponding locally to

$$\delta = \min_{\alpha}\{|p_{\alpha 1} - p_{\alpha 2}| + |p_{\alpha 2} - p_{\alpha 3}| + |p_{\alpha 3} - p_{\alpha 4}|\} \quad (1)$$

Therefore the method primarily encodes the local texture properties in terms of an appropriate motif. And then, it encodes similarity relations among intensity variations along the specified scan directions in the image. MCMs are typically computed over the whole image for retrieval purposes. In order to increase the CBIR performance we introduce an MCM-based region representation. The region-based retrieval is an interesting concept for increased semantics at a low cost. Local properties of given regions could help understanding the image structure better, thus contributing to a more meaningful image retrieval. The image is first segmented, and for each region we do the following.

1. The image is divided into $2 \times 2$ pixel grids and each pixel quadruplet is traversed by the optimal Peano scan, and is replaced by the corresponding motif.
2. The MCM is constructed for each region from the motif transformed image: Let $(i, j)$ be an entry. We define the MCM-based representation of a region as the probability of finding "*motif i*" at a given distance from "*motif j*".

Feature vectors can now be derived from the MCM representation to encode textural information about each region. The way it was described above, the method would be very sensitive to translation by an odd number of pixels. A shiftedimage is likely to have a very different MCM. A translation by one pixel causes the $2 \times 2$ neighborhood to vary significantly. To compensate for translation effects we construct four MCM feature vectors, for shifted versions of the original image by one pixel horizontally, vertically and diagonally. Among four feature vectors, one of them would correspond to the query vector independent of the amount of translation, because all translations are modulo-2 invariant.

### 2.2. Region-based spatial relationship

#### 2.2.1. Definitions

In this section we describe how we define and extract the spatial-relation among the regions. The approach should be helpful for more precise queries like "Find images displaying similar regions crossed by the sea". Two kinds of representations apply here. The first is based on topological

relations: for instance, Egenhofer and Franzosa (1991) proposed eight topological positions, namely "equal, cover, inside, overlap, touch, disjoint, covered-by and contains". They are invariant under translation, scaling, or rotation but are quite fuzzy in definition. The second one is based on the spatial orientation describing where regions are located relative to one another. Frequently used directional relations chosen are north, northeast, etc., as per classical Freemans codes. The location of a region is specified by its centroid. The angle "$\theta$" between line joining the centroids of two regions with the $x$-axis marks the direction. One can complete the polar representation with a distance "$d$" (e.g. Euclidean) between two regions.

Within the frame of Peano scans of images, a tree representation by regular blocks called *-tree emerges naturally (Seetharaman and Zavidovique, 1997). $(\theta, d)$ can be trivially recovered from the *-*tree* shown in the next section and illustrated in Fig. 2. This representation is easier to handle than the so-called topological relations as introduced in (Zavidovique and Seetharaman, 2001). Should an application require, more topological representations can be introduced in the binary or quad tree. For instance "cover" or "inside" or "overlap" are nothing but relations of higher order between indexes along the scan over the tree (see Section 2.2.2 for indexes) to describe square or rectangle blocks the regions would lay in or intersect with. Yet such relations are not as straightforward as orientations or distances.

### 2.2.2. Distance and angle in a tree

Let us assume a $Z$-bin tree notation like in Fig. 3. Quad nodes (square blocks) are labelled with bold rectangular numbers and bin nodes (rectangular blocks) are in italics. The root (image) is 1 and the bottom right pixel is $2_{2l+1} - 1$ for a $2_l$ by $2_l$ image ($l = 4$ in the figure).

Other indexing schemes are also available for bin or quad *-tree. But this one is of particular interest with respect to Peano codes and their use (Zavidovique and Seetharaman, 2001). In all cases, the addressing scheme and subsequent distance or angle computation will be similar. The label of a block is its rank along the scan over tree layers. We consider two blocks of labels $p$ and $n$, respectively, and $p > n$. Computations of their distance and relative orientation consist of

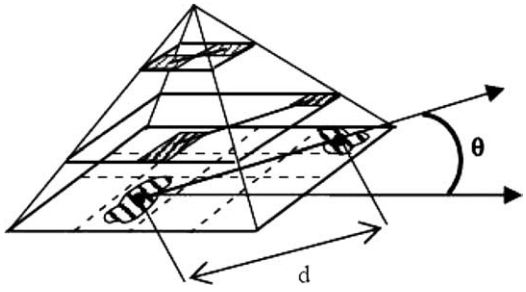


Fig. 2. Orientation and distance between two regions.

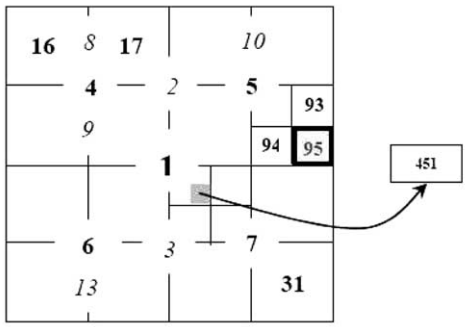


Fig. 3. Illustration of $Z$-bin tree indexing.

1. Exhibiting the blocks of descendent pixels and their centre.
2. Deriving the respective $x$ and $y$ coordinates, and computing their distance and relative angle. More precisely:

$$\alpha_p = \log_2 p \tag{2}$$

Then the block of kins at the bottom layer ranges from $p2^{2l-\alpha p}$ to $(p+1)2^{2l-\alpha p} - 1$. In this block, the first pixel next to centre pixel gets the label $c_p = p2^{2l-\alpha p} + 2^{2l-\alpha p-2} - 1$. Its index along the Peano scan in the picture is $i_p = c_p - 2^{2l}$. It is shown in (Seetharaman and Zavidovique, 1997) how the so called bit spreading (or collation) operator makes an invertible mapping from the $x$ and $y$ coordinates of a pixel to its index $i_p$ in the $Z$-scan (or $i_p$ to $(x, y)$). The binary expression of $x$ is bit spread to $s_x$ in inserting $0_s$ between the bits of the string: for instance 5 would give $\underline{0}1\underline{0}0\underline{0}1$. Then $i_p(x, y)$ is the integer for which the bit string is $s_x + s_y0$, where $s_y0$ is the one shifted version of $s_y$ by bit to the left. Conversely given a bit-string, there is a unique decomposition of it into the pair $(s_x, s_y)$. For example **195** is 11000011, without any ambiguity since $s_y$ starts with a 0 and component strings have one bit 0 out of two: (1001, 1001). It is enough to ignore either even or odd bits. For an example, the relative distance and angle of **95** and **451** require the sequence of computations:

$\alpha_{95} = \log_2 95 = 6,$
$c_{95} = 95.2^{2_x4-6} + 2^{2_x4-6-2} - 1 = 380,$
$380 - 256 = 124 \rightarrow 1111100 \rightarrow 1110, 110 \rightarrow 14, 6.$

The distance and angle between (9,9) and (14,6) are:
$d = (34)^{1/2} = 6$ and $\theta = 0.54$ rd $= \pi/6$.

## 3. MCM properties

### 3.1. Illumination invariance

Variation in illumination has a major influence on the appearance of an object and amount of information presented in an amplitude quantized image. The same region captured under different lighting conditions will exhibit various levels of information. The aim is to bring out some

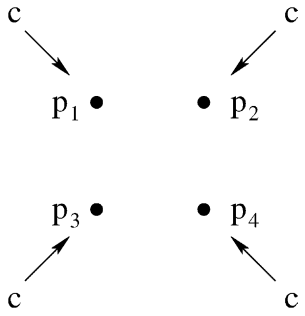


Fig. 4. Four adjacent pixels receiving the same quantity $c$ of light.

consistency in feature representation. Suppose that the light illuminates the surface uniformly and the change in lighting conditions is considered additive or multiplicative in intensity locally. It is obvious that the histogram will be affected by the brightness change. Methods based on grey level co-occurrence matrix (GLCM) will not be efficient because they depend on pixel values. Let $p_1$, $p_2$, $p_3$, $p_4$ intensity values of four adjacent pixels. Let $c$ be an additive value of intensity to the four pixels shown in Fig. 4. According to the Eq. (1) we have

$$\delta = \min_{\alpha}\{|(p_{\alpha 1}+c)-(p_{\alpha 2}+c)|+|(p_{\alpha 2}+c)-(p_{\alpha 3}+c)| + |(p_{\alpha 3}+c)-(p_{\alpha 4}+c)|\} \quad (3)$$

It is obvious that Eq. (3) is equivalent to Eq. (1). Therefore Eq. (3) yields the same motif as Eq. (1). The same is true if one works with a multiplicative change. The motifs do not change under such lighting conditions. Other advantages of MCM over conventional co-occurrence matrices and other intensity-based texture descriptions can be found in (Jhanwar et al., 2004).

### 3.2. Invariance to resolution changes

The normalized MCM is computed for a homogenous region at multiple resolutions in Fig. 5. The resolutions of images from left to right respectively are: 4 m, 2 m, 1 m, 0.5 m, 0.25 m as obtained from Terraserverwebsite. Note that for the first two images MCM is computed in the highlight rectangle. We notice that the MCM values are quite close even at different levels of resolution, thereby denoting that the characteristic features are fairly invariant under MCM representation. This allows a better accuracy for similarity measure and retrieval of regions. However, the invariance still depends on the image content and the scale variation. Qualitatively over all experiments, on both textures and aerial images (cf. Section 5), it appears that for non-homogenous regions the MCM stays stable when the scale factor remains small respective to the frequencies involved. A large scale variation yielding different representation at different resolution ends affecting MCM.

### 3.3. Separability of classes

Satellite images contain different types of terrain (vegetation, mountains, buildings...). Our retrieval process is based on similarity between regions containing texturally homogenous areas represented by an MCM. The goal is to check the separability between classes using the MCM feature. As the MCM vector has a dimensionality (36 dimensions) difficult to be shown graphically, we use the principal component analysis (PCA) to reduce the dimensionality, thereby allowing us to show the scatter diagram of classes. Five types of images have been picked randomly from the database to form five different classes shown in Fig. 6a. Each class contains five similar images from the database. So the input data contain 25 points. Fig. 6b shows the scatter diagram of classes, and it illustrates that the points belonging to a particular class tend to be clustered separately. This is another MCM property which is useful while defining the similarity measure for image retrieval.

## 4. Segmentation

The goal of segmentation is to split images into segments that correspond to the real world objects or regions. We propose the addition of a new feature for data to be classified or clustered. The image is scanned by overlapped

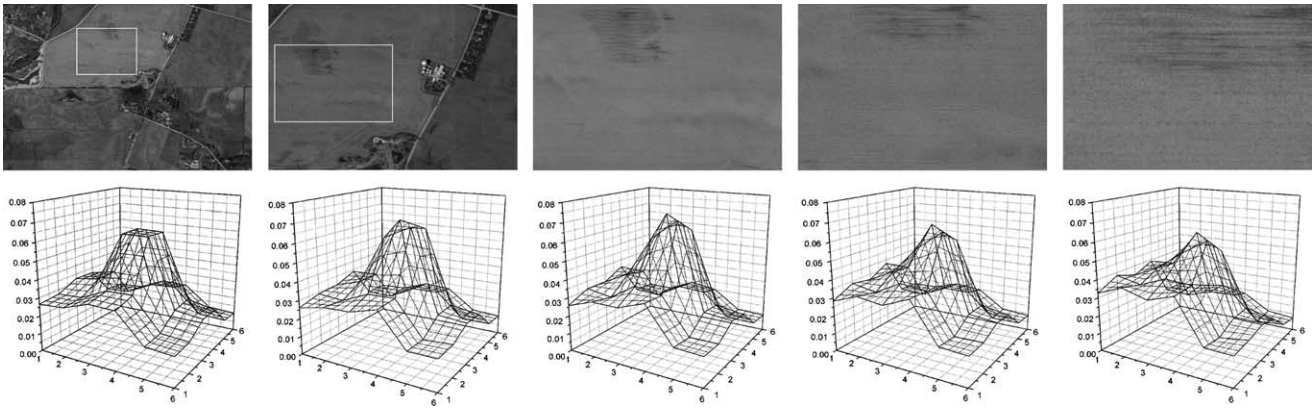

Fig. 5. Illustration of scale space invariance of the MCM feature. The curves represent MCM values computed for each of images shown exactly above them.

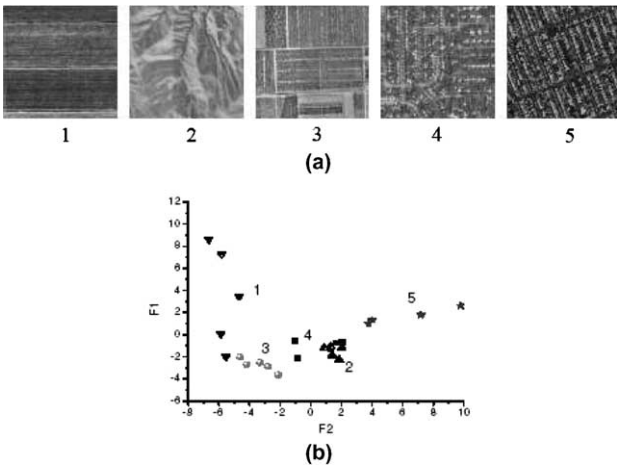


Fig. 6. (a) Sample images for different types of area and (b) scatter diagram for five classes using two first principal components. Legends are shown under each image.

windows of size $w \times w$ as shown in Fig. 7. At each pixel, the MCM is computed locally over the window. Further, to preserve color information, the mean value and the variance are calculated in the same window and added to the feature space with a weak weight. The data set which is represented by (MCM, mean, variance) will be classified into clusters, each of them is likely to produce several regions following the spatial repartitioning (see Fig. 8). To partition the data space we use Fuzzy $C$-Means (FCM) clustering (Bezdek, 1981). It is an unsupervised method allowing each point to belong to a particular class with a degree of membership.

Let us underline that the number of clusters $C$ is set apriori depending on some domain knowledge about the terrain, such as crop variety, various physical, agricultural and urban attributes. Typically such a domain knowledge is available by the local GIS user group. In case a manual initialization turns out to be inaccurate due to unavailability of such a knowledge, and a higher level of retrieval automation is looked for, techniques do exist to predetermine C without any human intervention. They may rely on a histogram-based preprocessing or be adaptive in incrementing the parameter up to meeting some satisfaction criterion (Gath and Geva, 1989; Xie and Beni, 1991).

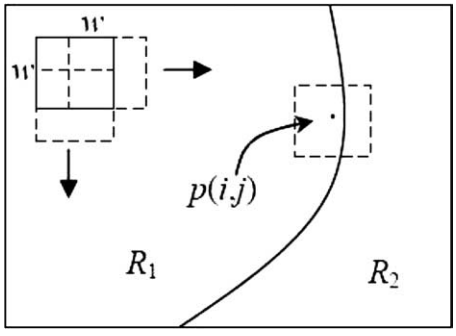


Fig. 7. A window with size $w \times w$ used to scan the local textural property. The dashed widow around the pixel $p(i, j)$ represents the zone of possible misclassification.

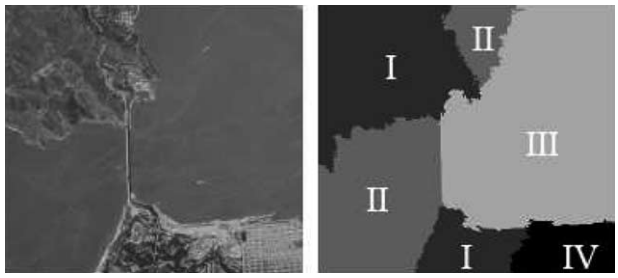


Fig. 8. shows segmentation results. The window size $w = 64$. The number of classes is set to $C = 4$, the fuzziness factor is set to $m = 2$ and the difference between successive instances of the fuzzy partition matrix is set to $\epsilon = 0.001$ for the termination criterion. Observe that regions I and II involve spatially disjoint regions.

Let $X = \{x_1, x_2, \ldots, x_N\}$ a set of $N$ feature data vectors $x_k \in R^n$, $n$ is the dimension of feature space. The FCM algorithm minimizes the flowing objective function:

$$J(U, V) = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^m \|x_j - v_i\|^2 \tag{4}$$

where $v_i$ represent the prototype, known as the cluster centre. The norm $\|\cdot\|$ denote the Euclidian distance ($d_{ij}$) between data point and prototype, $m$ controls the fuzziness of membership must be >1. $u_{ij}$ represent the partition matrix and satisfies the following conditions:

$$U\left\{u_{ij} \in [0,1]\ \forall i,j\ \ 0 < \sum_{j}^{N} u_{ij} < N\ \forall i\ \ \sum_{i}^{C} u_{ij} = 1\ \forall j\right\} \tag{5}$$

the cluster prototypes are given by

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m} \tag{6}$$

and the membership degrees are given by

$$u_{ij} = \frac{1}{\left(d_{ij}^2 \Big/ \sum_{k=1}^{C} d_{kj}^2\right)^{\frac{1}{m-1}}} \tag{7}$$

The window defined around the pixel needs to have enough size to capture a texture information, it depends on texture properties such as scale, coarseness, etc. However windows cause problem in borders of regions because it includes other regions characteristics as shown in Fig. 7. If the window involves more than one region the feature vector will contain mixed information, this may lead to a misclassification of the point. The proposed approach proceeds in two stages. Firstly FCM is used to compute the membership function for each pixel. Then the pixel is assigned to a particular region through defuzzification assigning it to class $c$ if its membership degree is more than 70%, if it is

less than 70% the pixel is tagged as belonging to an ambiguous zone. The second stage consists in classifying ambiguous points with a method similar to region growing. The candidate pixel $p$ is merged to the neighboring region if it satisfies the following condition:

$$|p - m_i| < \beta_k \tag{8}$$

where $m_i$ is the mean value of region $i$. $p$ is the pixel intensity and $\beta_k$ represent the threshold at iteration $k$. If the number of non-classified pixels stays the same in iteration $k$ and $k+1$, this means no point can be merged with this threshold, $\beta_k$ is then incremented allowing more pixels to be merged.

**Algorithm**

**Stage 1**: FCM with MCM features
- *Step* 1: Chose the number of clusters, initialize the membership matrix with random values with respect to (5). Chose appropriative values of $m$ and $\epsilon$
- *Step* 2: Compute clusters centers by using (6)
- *Step* 3: Compute $d_{ij}$
- *Step* 4: Update $u_{ij}$ using (7)
- *Step* 5: If $|u_{ij}^k - u_{ij}^{k+1}| > \epsilon$ go to Step 2
- *Step* 6: Assign each point to its class according to the maximum membership value. If this value is less than threshold (70%) do not assign the point to any class.

**Stage 2**: Classification of pixels not assigned to any class:
- *Step* 1: Initialize $\beta_k$ to 0
- *Step* 2: If one region or more in the pixel neighborhood satisfy (8) merge the pixel to the region $R_i$ having the minimal difference.
- *Step* 3: If the number of candidate pixels is stable in iteration $k$ and $k+1$ increment $\beta_k$ by one and go to Step 2
- *Step* 4: If there are no candidate point stop

## 5. Image retrieval

During the retrieval process each image in the database is split into different regions by the method described in Section 4. Two kinds of characteristics, visual and spatial characteristics are used for similarity measure. The first approach exploits visual features represented by MCM in each region of image and the second combines visual features and spatial relationship. Let $Q$ denote the query image and $DB$ an image in the data base: $Q$ and $DB$ contain, respectively, $M$ and $N$ segments. A weighted Bhattacharyya distance is used to compute the similarity measure (see Eq. (9)). Let $\omega$ and $\omega'$ denote area of respective segments in $Q$ and $DB$. The distance between two regions $R^i$, $R'_j$ is

$$D_r(R_i, R'_j) = -\alpha_i \log \sum_p V_i(p) V'_j(p) \tag{9}$$

with $V_i, V'_j$ as the feature vectors (MCM). The weight is set to

$$\alpha_i = \frac{\min(\omega_i, \omega'_j)}{\sum_{i=1}^{M} \omega_i} \tag{10}$$

The MCM is normalized when the distance is computed, so that the region size does not matter. Consider the approach in which the query is done without spatial constraint. In this approach each region in $Q$ is compared with all segments in $DB$. That requires $m$ regions of $Q$ to match with $m$ regions of $DB$, where $m \leqslant \min\{M, N\}$ (see Fig. 9). The distance between two images is computed from distance of the most resembling pairs, restricted to those for which the distance is less than a given threshold

$$D(Q, DB) = \frac{1}{m} \sum_{i=1}^{M} \min_{j \in [1,N]} \left\{ D_r(R_i, R'_j) \cdot 1_{\{(i,j)/D_r(R_i,R'_j) < \text{th}\}} \right\} \tag{11}$$

$1_x(i,j)$ the set function of $X$ : $\begin{cases} 1_X(i,j) = 1 & \text{iff } \in X \\ = 0 & \text{else} \end{cases}$

Query by regions with spatial relationship requires that similar images must have a segment-wise textural similarity as well as similar spatial arrangements of segments. The visual characteristics and the spatial relationships are jointly represented by the attributed relational graph. The ARG of image $P$ is denoted by $GP(V, E)$ where $V(V1, V2, \ldots)$ is a set of attributed nodes (features) and $E(E1, E2, \ldots)$ is the set of attributed edges. In the present case, the ARG nodes contain regional textural features (MCM) and edges encode the spatial relationship $(\theta, d)$ among regions. Thus the graph contains all the information. Measuring the similarity between two images is tantamount to graph matching. To measure the similarity between the two graphs, firstly nodes are compared. A more complete similarity measure, including spatial constraints, is then computed on the sub-graphs of matching nodes. The spatial similarity is measured again by a weighted distance among edges:

$$DG(Q, DB) = D(Q, DB) + \frac{1}{m\beta_i} \sum |E_i - E_j| \tag{12}$$

for $E_i$ and $E_j$, respective edges between nodes of corresponding pairs in $Q$ and $DB$

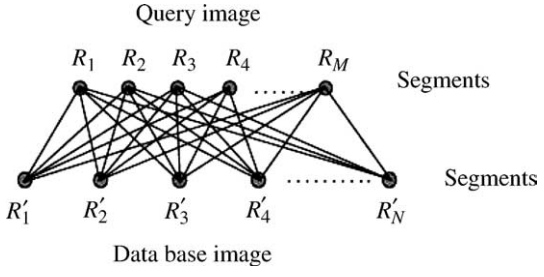


Fig. 9. Comparison of features over regions.

$$\beta_i = \frac{\omega_i}{\sum_{i=1}^{M} \omega_i} \tag{13}$$

gives more weight to the larger region in $Q$.

## 6. Experiments and results

In our experiments we use real data, 400 aerial images from http://terraserver.microsoft.com and 30 from the Brodatz data base http://sipi.usc.edu. Similar images are sorted in groups where each group contains between 6 and 10 images. A relevant retrieved image means that it belongs to the same group as the query image. Generally a group of images is selected from the close geographic area where the landscapes show similar visual aspects, the selection being based on human judgment. Fig. 10 shows an example of a group (the images from rank 1 to rank 6 are relevant). Let us underline that a proper framework for evaluation of the relevance of answers to queries is still to be found and is an open research. The measures extensively used in the field are called recall and precision.

$$\text{Recall} = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of relevant documents}}$$

$$\text{Precision} = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of documents retrieved}}$$

The experimentation sequence described below aims at checking four properties of our method: improved process with MCM by regions, better retrieval than with classical methods, more precise retrieval with spatial relationship added, impact of resolution.

(1) To show the retrieval performance of this region-based approach in which MCM is computed for each region, we compare the technique to a previous global approach in which MCM is computed over the entire image. Region-based retrieval shows a better efficiency than the global approach. An example of performance comparison is given in Fig. 11. Note that the query image is the same as shown in Fig. 10.

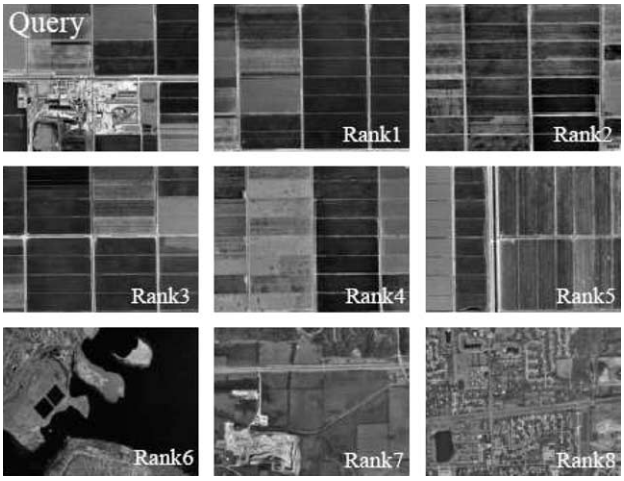


Fig. 10. Retrieval results of query by regions.

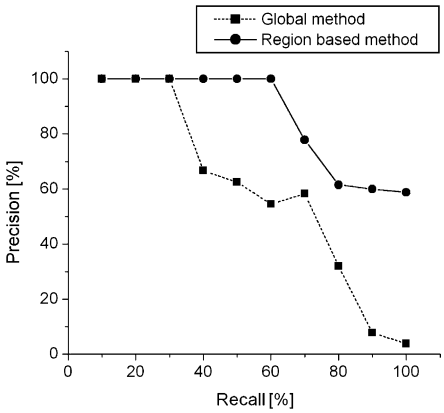


Fig. 11. Comparison of retrieval results of query using global method and region-based method.

(2) To asses the "MCM by region" performance we compare it with three more classical methods. The retrieval process is kept the same for all four cases, only region features and the representation vary—MCM, histogram, GLCM or Gabor filters. Meastex (http://www.cssip.uq.edu.au/meastex/meastex.html) source code was used to compute GLCM and Gabor filter results. The methods depend on several parameters to be set. For GLCM the distance between two pixels $d$ is set to 1 and orientations are 0°, 45°, 90° and 135°. GLCM is computed by averaging matrices over all angles with symmetric option. Then features are derived from this matrix as proposed in (Conners et al., 1984). For the Gabor filters energy method parameters are the mask size $w$, wave length $\lambda$, orientation $\theta$ and phase shift $\phi$. We chose $w = 17 \times 17$, $\lambda = 2, 4, 8$, $\theta = 0, 45, 90, 135$ and $\phi = \pi/2$. The feature vector of a given region is obtained by computing the energy for each $\lambda$ and $\theta$ values. Curves shown in Fig. 12 are the average result on 20 different query images: we can notice that the performance of GLCM and Gabor filters methods decrease faster than histogram and MCM. MCM shows better results and more stability in any retrieval case.

(3) In addition to having a better retrieval performance, spatial relationship has an advantage to get a more precise query. Retrieval performance has been tested for another

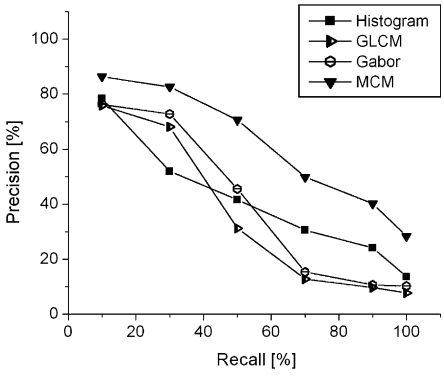


Fig. 12. Performance comparison between MCM, GLCM and Gabor filters-based method.

series of images where regions layout is required. In this case we have compared the performance of query by regions with and without spatial relationship. Adding region and spatial constraints, the retrieval shows an improved performance. Fig. 13 shows an example of the performance measure.

(4) To test the multi-resolution effect we process 100 more images of areas at two different scales 1 m and 0.5 m. Note that such resolutions make images difficult compared to the previous examples at 4–8 m (cf. Section 3.2). Groups of relevant images are made again accounting for both resolutions. Conventional curves "precision = *f*(recall)" are shown Fig. 14. The retrieval for 0.5 m is globally better in these cases than for 1 m, and both are better than when resolutions 0.5 and 1 m are mixed. Indeed, two scales bring out additional ambiguities. A large range of the resolution results into severe modifications of the perceived textures, asking for a different semantics to be implied in the process (Figs. 15 and 16).

The behavior of the curves is found to be consistent on the 4–5 queries taken from every groups of 6–10 image. The present method shows in this case an improvement by at least 2 more relevant images (in the same group), compared to other methods (see Figs. 11–13) tested under similar conditions.

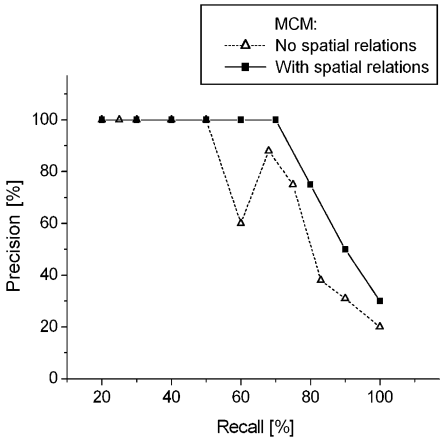


Fig. 13. Retrieval results of query by only regions, and using both spatial and regional properties.

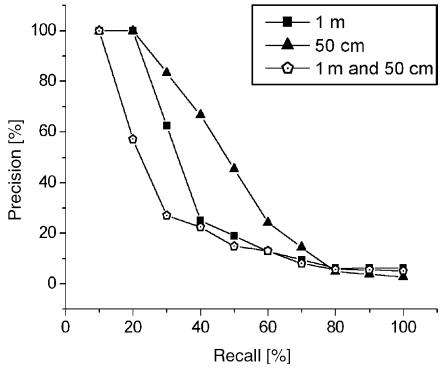


Fig. 14. Retrieval results for resolution 1 m and 0.5 m.

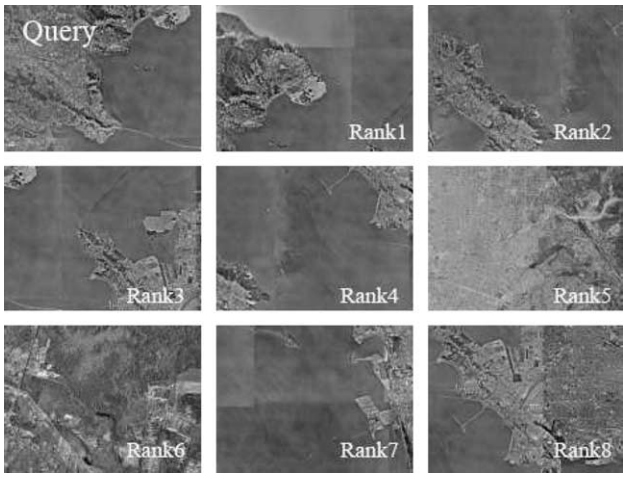


Fig. 15. Another retrieval results of query by regions.

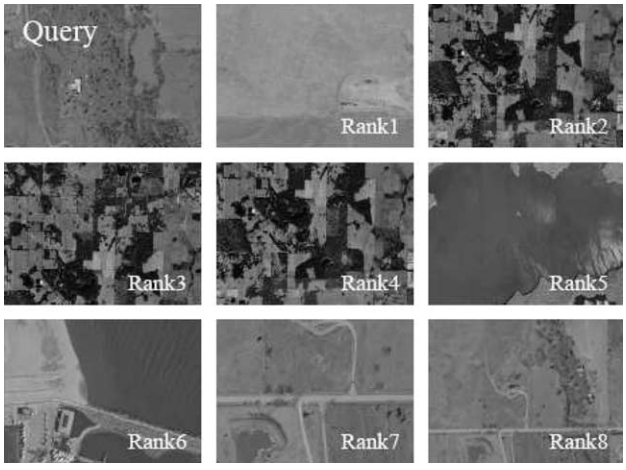


Fig. 16. Retrieval results of query by regions at 1 m, images 1, 8 are at 1 m, 6, 7 at 0.5 m and 2 to 5 at 8 m. 1, 7 and 8 are correct answers, 2, 3 and 4 resembling forests.

Note that we do not address the computational complexity related issues in the present paper. Our team has been quite successful in building massively parallel machines, such as the SPHINX Pyramid Machine (Clermont et al., 1990; Merigot and Zavidovique, 1992). Its connection network fits exactly the tree structure stemming from Peano scans, and hence the computation poses no difficulty.

## 7. Conclusion

The MCM-based description of regional textures proves to be of particular interest in the case of aerial images, and more generally, natural images such as landscapes. The choice of Peano coding the picture results in an easy computation of the spatial relationship after segmentation. Thanks to the data structure involved, the whole process is computationally efficient, and the complexity remains low despite the handling of ARGs. Future work deals with

evaluating the outcome of extending the process to multi-spectral images. More fundamentally, we shall endeavor to making the segmentation and coding even more closely related to Peano scanning process. Indeed, the resolution pyramid naturally associated with Peano scanning an image is used solely to accelerate the process of finding the spatial relations. Apart from the invariance of MCMs to resolution (see Fig. 5), benefits from extending to such a texture coding in multi-resolution should be worth investigation. We are currently investigating the possibility of including the scan semantics over a multi-resolution while performing the segmentation using FCM.

## References


Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., 1996. The virage image search engine: An open framework for image management. In: Proc. Storage and Retrieval for Still Image and Video Databases, SPIE, vol. 2670, pp. 76–87.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.

Carson, C., Thomas, M., Belongie, S., Hellerstein, J.M., Malik, J., 1999. Blobworld: A system for region-based image indexing and retrieval. In: Proc. Third Internat. Conf. on Visual Information Systems 1999, pp. 509–516.

Chang, S.K., Shi, Q.Y., Yan, C.W., 1987. Iconic indexing by 2-D strings. IEEE Trans. Pattern Anal. Machine Intell. 9 (3), 413–428.

Clermont, Ph., Mrigot, A., Roussel, J.C., Zavidovique, B., 1990. Parallel implementation of a region merging algorithm on a pyramid machine. In: Progress in Image Analysis. Scientific World Publishing, pp. 721–732.

Conners, R.W., Trivedi, M.M., Harlow, C.A., 1984. Segmentation of a high-resolution urban scene using texture operators. Comput. Vision Graphics Image Process. 25, 273–310.

Egenhofer, M., Franzosa, R., 1991. Point-set topological spatial relations. Internat. J. Geograph. Inform. Systems 5 (2), 161–174.

El-Kwae, E.A., Kabuka, M., 1999. A robust framework for content-based retrieval by spatial similarity in image databases. ACM Trans. Inform. Systems 17 (2), 174–198.

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P., 1995. Query by image and video content: The QBIC system. IEEE Comput. 28, 23–32.

Gath, I., Geva, A.B., 1989. Unsupervised optimal fuzzy clustering. IEEE Trans. Pattern Anal. Machine Intell. 11 (7), 773–781.

Gudivada, V.N., Raghavan, V.V., 1995. Design and evaluation of algorithms for image retrieval by spatial similarity. ACM Trans. Inform. Systems 13 (2), 115–144.

Jhanwar, N., Chaudhuri, S., Seetharaman, G., Zavidovique, B., 2004. Content based image retrieval using motif cooccurence matrix. Image Vision Comput. 22 (14), 1211–1220.

Ma, W.Y., Manjunath, B.S., 1998. A texture thesaurus for browsing large aerial photographs. J. Amer. Soc. Inform. Sci. 49 (7), 633–648.

Ma, W.Y., Manjunath, B.S., 1999. Netra: A toolbox for navigating large image databases. Multimedia Systems 7 (3), 184–198.

Merigot, A., Zavidovique, B., 1992. Image analysis on massively parallel computers: An architectural point of view. J. Pattern Recognition Artificial Intell. 6 (2), 387–393, World Scientific.

Pentland, A.P., Picard, R., Sclaroff, S., 1996. Photobook: Content-based manipulation of image databases. Internat. J. Comput. Vision 18 (3), 233–254.

Petrakis, Euripides, G.M., Faloutsos, C., 1997. Similarity searching in medical image databases. IEEE Trans. Knowledge Data Eng. 9 (3), 435–447.

Rui, Y., Huang, T.S., Chang, S.-F., 1999. Image retrieval: Current techniques, promising directions and open issues. J. Visual Comm. Image Represent. 10, 39–62.

Seetharaman, G., Zavidovique, B., 1997. Image processing in a tree of Peano coded images. In: IEEE-CAMP 97, Boston, pp. 229–234.

Swain, M., Ballard, D., 1991. Color indexing. Internat. J. Comput. Vision 7 (1), 11–32.

Xie, X.L., Beni, G., 1991. A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Machine Intell. 13 (8), 841–847.

Zavidovique, B., Seetharaman, G., 2001. $*$-Trees for Peano based image processing. CACS Technical Report and Working Paper, University of Louisiana.