Image Indexing and Retrieval Using Spatial Chromatic Histograms and Signatures

*G. Ciocca°, R. Schettini°, and L. Cinque**
°Istituto Tecnologie Informatiche Multimediali, Consiglio Nazionale delle Ricerche Milano, Italy
**Universita' di Roma "La Sapienza", Dip. Scienze dell'Informazione Rome, Italy*

Abstract

We introduce an innovative image retrieval algorithm, the Dynamic Spatial Chromatic Histogram, that makes it possible to take into account spatial information in a flexible way without drastically increasing the computation cost. The results of experimental trials on a database of about 3000 images are here reported and compared with Spatial Chromatic Histogram matching and more standard techniques based on human perceptual judgments to demonstrate the effectiveness of the algorithm proposed.

1. Introduction

The research presented here focuses on the pictorial feature color, which, together with shape, texture, size, distance and relative position, characterizes scenes and pictures. We introduce an innovative image retrieval algorithm, the Dynamic Spatial Chromatic Histogram or DSCH, based on our previous work with the Spatial Chromatic Histogram[1] (SCH). The DSCH makes it possible to take into account spatial information in a flexible way without drastically increasing the computation cost. The main features that it inherits from the SCH are:

(1) it takes into account, in an efficient and effective way, both chromatic and spatial features, and
(2) the similarity function designed makes it possible to perform partial image queries such as: search for images having a blue region on the top, and a green region on the bottom.

The extension we introduce allows the algorithm to compare two images on the basis of their most salient colors as identified by color clustering and/or spatial image segmentation, for a more precise and compact characterization of the content of the images.

The results of experimental trials on a database of about 3000 images are here reported and compared with Spatial Chromatic Histogram matching and more standard techniques based on human perceptual judgments to demonstrate the effectiveness of the algorithm proposed.

2. Image Indexing and Retrieval

The image retrieval process comprises two steps (Fig. 1):
(1) indexing, in which for each image in the database, a set or a vector of features summarizing the properties of its contents is computed and stored;
(2) retrieval, in which the features of the query image are extracted and compared with the others in the database. The database images are then ranked in order of their similarity.

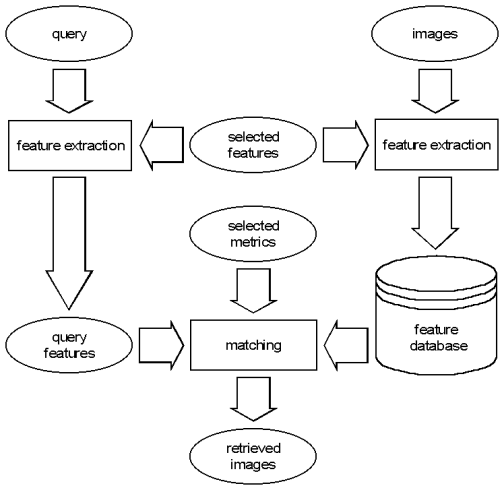

Figure 1. Schematic description of the image retrieval process.

2.1 Image Indexing

Image indexing implies reducing the number of colors used to represent the image's color content. This reduction can be done by color space quantization, using a predefined color palette (static quantization), or by clustering and/or spatial segmentation (dynamic quantization).[2]

Static quantization is achieved by partitioning the color space into equivalence classes. As the SCH algorithm required a predefined color palette, we chose to partition the sRGB color space in 11 basic colors: red, orange, yellow, green, blue, purple, pink, brown, black, gray and white.

Dynamic quantization is achieved by clustering the colors of the image to the k most representative colors (color signature), where k can be set by the user (the default value is 11). To do so the R, G and B, values are first converted to H, S and V color coordinates, and then a competitive learning clustering algorithm is then applied[3].

Both quantization strategies produce a coarse, but completely unsupervised image segmentation. Obviously image-dependent color quantization characterizes the color content of the image much more precisely.

To reduce the noise, while preserving edges and chromaticity contents (without adding any new colors), a vector median filter can be applied to the segmented image, using a working window *W* of 3 x 3 pixels.[4]

2.2 Image Retrieval by Matching Spatial Chromatic Histogram

Spatial Chromatic Histograms are extended histograms that preserve not only information about the color content of the image, but also the spatial distribution of each color within the image. Each entry in a SCH is composed of three values: $h_I(k)$, the ratio of pixels in *I* of color *k*; $\mathbf{b}_I(k)=(\bar{x}_k,\bar{y}_k)$, the baricenter (in relative coordinates) of the spatial distribution of color *k*; and $\sigma_I(k)$, the standard deviation of the distribution of color *k*. The entries in the SCH for an image are then:

$$S_I(k)=\left(h_I(k),\mathbf{b}_I(k),\sigma_I(k)\right) \text{ where } k=1\ldots n. \quad (1)$$

The similarity function proposed takes into account both color and spatial information as follows:

$$f(I_1,I_2)=\sum_{k=1}^{n}\left[min\left(h_{I_1}(k),h_{I_2}(k)\right)\times\left(\frac{\sqrt{2}-d\left(b_{I_1}(k),b_{I_2}(k)\right)}{\sqrt{2}}+\right.\right.$$
$$\left.\left.+\frac{min\left(\sigma_{I_1}(k),\sigma_{I_2}(k)\right)}{max\left(\sigma_{I_1}(k),\sigma_{I_2}(k)\right)}\right)\right] \quad (2)$$

where the second term is a weighting function. This function of enhances the spatial component that presents large populations of equally colored pixels by increasing the component when the baricenters are closely located.

2.3 Image Retrieval by Dynamic Spatial Chromatic Signatures Matching

Each entry is composed of four values: the three values of a SCH entry and the coordinates of the color considered. Given an image *I*, the new descriptor is defined as $DSCH_I=\{S_I(k)\}$, with $S_I(k)$ indicating the single entry *k=1…n,* and *n* the number of colors of the segmented image (*n* may be less than the maximum number of colors allowed). Each entry $S_I(k)$ is then:

$$S_I(k)=\left\langle h_I(k),\mathbf{b}_I(k),\sigma_I(k),c_I(k)\right\rangle \quad (3)$$

where *k=1…n,* $h_I(k)$ is the percentage of pixels of color k, $\mathbf{b}_I(k)$ are the coordinates of the baricenter of color k, $\sigma_I(k)$ is the standard deviation of the distribution of color k, and $c_I(k)$ are the HSV coordinates of the color *k*. DSCH similarity takes into account both color and spatial proximity. These factors are modeled in two terms, The first considers the spatial relationship between the two entries compared and the percentage of corresponding colors. The second term weights the formula with the actual similarity between the respective colors, *k* and *u,* of the two entries.

In formula, given two images I and J, the similarity between two entries $S_I(k)$ and $S_J(u)$ is then computed as:

$$f_s\left(S_I(k),S_J(u)\right)=\frac{1}{3}\left[\frac{\min\left(h_I(k),h_J(u)\right)}{\max\left(h_I(k),h_J(u)\right)}+\right.$$
$$\left.+\frac{\sqrt{2}-d\left(\mathbf{b}_I(k),\mathbf{b}_J(u)\right)}{\sqrt{2}}+\frac{\min\left(\sigma_I(k),\sigma_J(u)\right)}{\max\left(\sigma_I(k),\sigma_J(u)\right)}\right]\times \quad (4)$$
$$\times Sim_{HSV}\left(c_I(k),c_J(u)\right)$$

where the chromatic similarity between two colors (described in terms of HSV color coordinates) is computed as follows:

$$Sim_{HSV}(k,u)=1-\frac{1}{2}\left[\left(Val_k-Val_u\right)^2+\right.$$
$$+\left(Sat_k\cos Hue_k-Sat_u\cos Hue_u\right)^2+ \quad (5)$$
$$\left.+\left(Sat_k sinHue_k-Sat_u sinHue_u\right)^2\right]$$

and the spatial similarity term, which resembles that defined above in the SCH formulation, has been normalized in order to be equal to 1 when the two DSCH entries have exactly the same value.

The similarity between two images *I* and *J* represented in terms of DSCH is computed as:

$$f_s\left(DSCH_I,DSCH_J\right)=\frac{1}{\left\|DSCH_I\right\|}\times$$
$$\times\sum\nolimits_{S_I(k)}\max\left\{f_s\left(S_I(k),S_J(u)\right)\middle|\,S_J(u)\in DSCH_J\right\} \quad (6)$$

Each entry of the first DSCH is compared with each entry of the second DSCH. The pair that has the highest similarity value is taken to compute the global similarity score.

Since Eq. (6) is not symmetric, we finally compute the similarity between two images as:

$$f_s(I,J)=$$
$$\max\left\{f_s\left(DSCH_I,DSCH_J\right),f_s\left(DSCH_J,DSCH_I\right)\right\} \quad (7)$$

3. Experimental Results

For our tests we used a database containing 3000 still images representing various subjects and scenes, such as flowers, animals, landscapes and people. A subset of 25 images, chosen at random from the database, were used as the query images, and the system performed the similarity evaluation with respect to each query image. The ground truth similarity has been assessed jointly by authors. When a query was submitted, the system we implemented reordered the database images by similarity with respect to the query and then proposed to the user the most similar images. We considered an image assessed as truly

similar to the query correctly retrieved when it appeared in the first set of 21 displayed images.

To evaluate the effectiveness of our retrieval method we applied a measure proposed by Methre[5] et al. We let S be the number of relevant items the user wanted to retrieve when posing a query; $\mathbf{R}_\mathbf{q}^\mathbf{I}$, the set of relevant images; and $\mathbf{R}_\mathbf{q}^\mathbf{E}$, the set of relevant images retrieved in the short list. The effectiveness measure was defined as:

$$\eta_S = \begin{cases} \dfrac{\left|\mathbf{R}_\mathbf{q}^\mathbf{I} \cap \mathbf{R}_\mathbf{q}^\mathbf{E}\right|}{\left|\mathbf{R}_\mathbf{q}^\mathbf{I}\right|} & \text{if } \left|\mathbf{R}_\mathbf{q}^\mathbf{I}\right| \le S \\ \dfrac{\left|\mathbf{R}_\mathbf{q}^\mathbf{I} \cap \mathbf{R}_\mathbf{q}^\mathbf{E}\right|}{\left|\mathbf{R}_\mathbf{q}^\mathbf{E}\right|} & \text{if } \left|\mathbf{R}_\mathbf{q}^\mathbf{I}\right| > S \end{cases} \quad (8)$$

In the first case, the effectiveness was reduced to the traditional recall measure, while, in the second case, the effectiveness corresponded to precision. In our implementation S was set at 21.

For sake of comparison the following retrieval methods were applied:

Color Coherence Vectors[6] (CCV), a histogram-based method for comparing images that incorporates spatial information. Quantization is performed by uniformly sampling sRGB color cube (using 16x16x16 color bins) and then clustering in 64 colors using a K-means clustering algorithm. The authors classify each pixel of the quantized image as either coherent or incoherent, depending upon whether or not it is a part of a "large" similarly-colored region (a region is classified as large if its size exceeds 1% of the image size). By counting coherent and incoherent pixels separately the method offers a finer distinction between images than color histograms alone can provide. For each color c_i the number of coherent pixels, α_{c_i}, and the number of non-coherent pixels, β_{c_i} are computed; each entry in the CCV is thus a pair $(\alpha_{c_i}, \beta_{c_i})$, called a *coherence pair*. The whole coherence vector is defined as:

$$CCV(\tilde{I}) = \left\langle (\alpha_{c_1}, \beta_{c_1}), \ldots, (\alpha_{c_i}, \beta_{c_i}), \ldots, (\alpha_{c_n}, \beta_{c_n}) \right\rangle \quad (9)$$

The sum $\alpha_{c_i} + \beta_{c_i}$ is clearly the number of pixels of color c_i present in the image; the set of the sums for i=1..n represents the color histogram. The L1 distance can be used to compare two CCV:

$$\Delta CCV(\tilde{I}, \tilde{I}') = \sum_{i=1}^{n} \left(\left|\alpha_{c_i} - \alpha'_{c_i}\right| + \left|\beta_{c_i} - \beta'_{c_i}\right| \right) \quad (10)$$

The *Spatial-Chromatic Histogram SCH11***,** where the colors of the images are quantized in 11 colors according to our static quantization method, and the similarity is computed as described in Section 2.2.

The *Spatial-Chromatic Histogram SCH64,*: where quantization is performed as described above for DSCH, and the similarity is computed as described in Section 2.2.

The *Dynamic Spatial-Chromatic Signature DSCH,* where the colors of the images are quantized into 11 colors according to our dynamic quantization method, and the similarity is computed as described in Section 2.3.

Table I summarizes the experimental results, reporting the Min, Mean and Max values for Retrieval Efficiency, for each of the retrieval techniques we used. We observe that both the SCH and DSCH outperform the CCV, the DSCH achieving the best scores. An example of query result with the four algorithms is shown in Figures 2a-d.

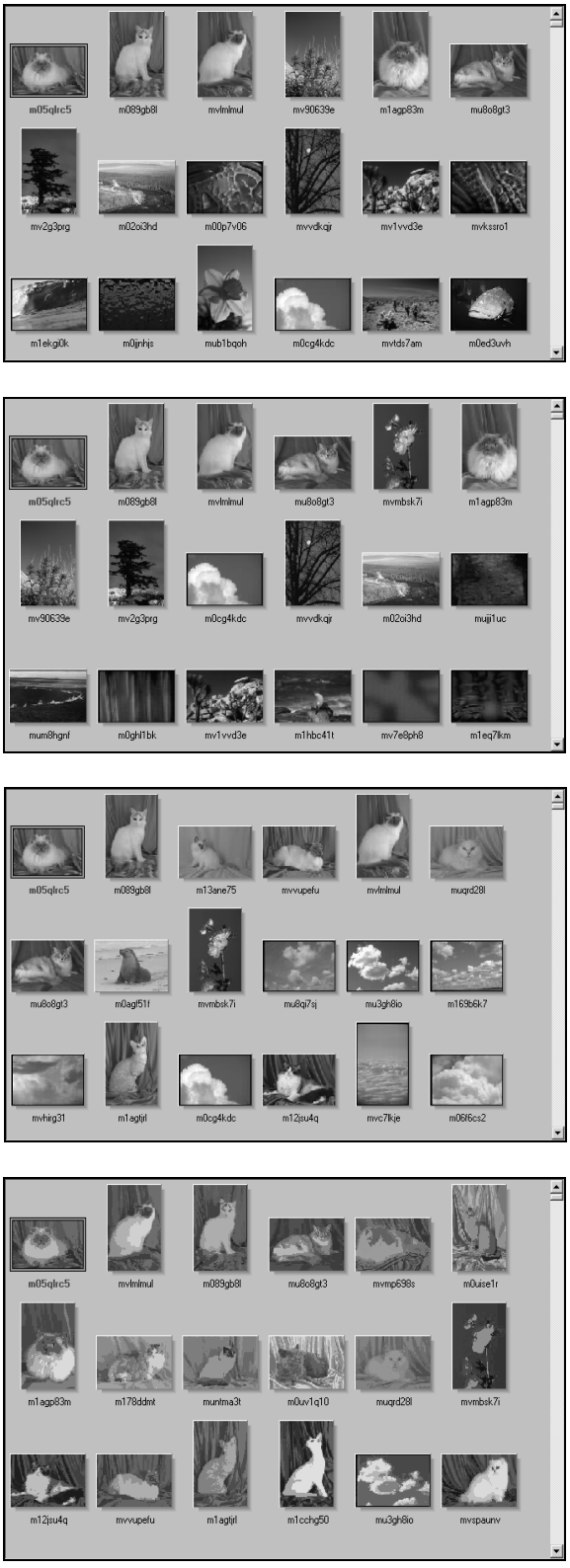

Figure 2. a) Results of the top left image with the CCV64 method; b) the same query with the SCH64 method; c) results with the SCH11 method; d) results with the DSCH11 method.

Table I. Experimental Results

	CCV64	SCH64	SCH11	DSCH11
Min	0.090909	0.181818	0.142857	0.190476
Mean	0.319073	0.414057	0.472625	0.502646
Max	0.523810	0.714286	0.761905	0.904762

4. Conclusions

Although color distribution does not capture the whole meaning of images, it is well correlated with their semantic contents in many practical situations. Our preliminary results show that the proposed indexing and retrieval strategy, which employs dynamic color quantization and takes into account spatial information, is a powerful approach. We feel that the system should also allow the user to refine the response to his query by indicating the relevance, or irrelevance, of the retrieved items and consequently updating the similarity measure. We are currently studying a solution to this problem.

References

1. R. Schettini, G. Ciocca, S. Zuffi, Color-based image retrieval using spatial-chromatic histograms, Image and Vision Computing, 19, pp. 976-986.(2001).
2. R. Schettini, G. Ciocca, S. Zuffi, A survey on methods for colour image indexing and retrieval in image databases, Color Imaging Science: Exploiting Digital Media, (R. Luo, L. MacDonald eds.), J. Wiley, 2001 (in print).
3. T. Uchiyama, M.A. Arbib, Color Image Segmentation Using Competitive Learning, IEEE Transaction on PAMI, 16, pp. 1197-1206 (1994).
4. J. Astola, P. Haavisto, Y. Neuvo, Vector median filters, Proceedings IEEE, 78, pp. 678-689. (1990).
5. B. M. Mehtre, M.S. Kankanhalli, Content based image retrieval using composite color shape approach, Inform. Process. Managment, 34, pp. 109-120. (1998).
6. G. Pass, R, Zabih, J Miller, Comparing Images Using Color Coherence Vectors. Proc. Fourth ACM Multimedia Conference. (1996).

Biography

Raimondo Schettini has been associated with the Italian National Research Council (CNR) since 1987. With 1994 he moved to the Institute of Multimedia Information Technologies, where he is currently the director of the Image and Color Analysis Lab. He has been team leader in several research projects and published more than 120 refereed papers on image processing, analysis and reproduction, and on image content-based indexing and retrieval. Since 1997 he also teaches a course on multimedia design at the Faculty of Industrial Design of the Polytechnic of Milan. He is member of The Technical Committee 8-03 On Gamut Mapping, Division 8 (Image Technology) of the Commission Internationale de l'Eclairage (CIE). He has been General Co-Chairman of the 1st Workshop on Image and Video Content-based Retrieval (1998), and general co-chair of the first and second EI Internet Imaging Conferences (2000, 2001). He is general co-chair of the First European Conference on Color in Graphics, Imaging and Vision (CGIV'2002) and of the Third EI Internet Imaging Conference (2002).