

USING COLOR DISTRIBUTION TO EFFECTIVELY QUERY IMAGE DATABASES

B. Barolo, I. Gagliardi, R. Schettini

Istituto Tecnologie Informatiche Multimediali (ITIM), Consiglio Nazionale delle Ricerche (CNR)
Via Ampere 56, 20131 Milano, Italy
e-mail: centaura@itim.mi.cnr.it

ABSTRACT

We present here an effective image retrieval strategy based on the fuzzy evaluation of color image similarity. In this method both the query and the database images are displayed in device-independent space with a limited palette of perceptual significance. Image color distributions are represented by histograms, and a suitable similarity measure between histograms is also defined in order to model the perceptual similarity between their different colors. Experimental results on a database of some 200 images are reported.

1. INTRODUCTION

Recent advances in image technologies have accelerated the development of multimedia and image information retrieval systems. However, these systems still have a limited capability for accessing images by content, in addition to traditional SQL and textual queries. Research projects on indexing images by low-level perceptual features, such as color, texture, shape and relative position, have mainly addressed system efficiency, assuming that the databases to be queried are large and that the user will then find the target images by browsing a short list of some tens of candidates [5, 14]. However, in several application fields - textile printing, ceramics and the visualization of sales products, for example - color information, to be of any use, must be faithfully described and rendered on the output device. If inappropriate color description and similarity evaluation schemes are adopted, many images that are actually similar in color may be missing from the short list. Moreover, since other factors, such as background and adjacent colors, image size, and the observer's state of adaptation to lighting, all strongly influence color appearance perception [7], target images can not be easily chosen among candidates by browsing through "thumbnail" representations of candidates.

This paper presents an effective image retrieval strategy based on the fuzzy evaluation of color image

similarity. In this method, which is an extension of our earlier studies on color range and image indexing [2, 3] using perceptual correlates of the psychological dimensions of Lightness, Chroma, and Hue, both query and database images are displayed in device-independent space with a limited palette of a perceptual significance. Image color distributions are represented by histograms, and a suitable measure of the similarity between histograms is also introduced in order to model the perceptual similarity between their different colors (i.e. between histogram bins). The experimental results on a database of about 200 images are reported and compared with those of other existing methods.

2. IMAGE DATABASE ACQUISITION AND INDEXING

To obtain a perceptually meaningful, device-independent color description, both input and display devices are calibrated colorimetrically, and the CIELAB color appearance model is adopted [17]. Non-colorimetric scanner outputs are related to CIELAB standard coordinates by means of a feed-forward neural network trained by back-propagation. The algorithm uses the ANSI IT8 color chart as its training set [11]. Despite its simplicity, the calibration accuracy of the method is comparable, and, in some cases, even superior to that reported by other authors in similar experiments [13]. Using CIELAB specifications to describe colors involves some form of device modeling for the output device as well. To find a mapping between the CIE XYZ tristimulus space and the RGB color space of the display, the calibration procedure we described in [12] has also been applied here.

The image is then quantized on a limited palette of perceptual significance, reducing the burden of massive image data on storage and processing without causing notable color shift. The proposed method partitions the CIELAB color space into about 250 subspaces (categories) in each of which the color remains perceptually the same and distinct from that

of neighboring subspaces. The mapping between the CIELAB values and the color categories is obtained with a look-up table.

The proposed quantization allows color specification to remain the same for small variations that may occur during the image acquisition process (changes in lighting conditions during picture shooting, over- or under exposure in developing film, digitalization noise, etc.).

Before color indexing, the objects of interest are segmented from the background on which they are displayed. The segmented images are then scaled to the same number of pixels before quantization and histogramming are performed.

A color histogram is a vector (h_1, \dots, h_n) in a n -dimensional vector space, where each element h_j represents the pixel percentage of the image having the color j and is labeled with the coordinates of its Lightness (L^*), Chroma (C^*), Hue (h^*), Redness-Greenness (a^*), and Yellowness-Blueness (b^*), as defined in the discretized CIELAB color space.

The choice of using the color histogram as the only index in an image database has allowed us to concentrate on the perceptual and cognitive tasks of comparing the appearance of color images.

3. HISTOGRAM SIMILARITY EVALUATION

The distance between two histograms can be measured with a metric on the histogram space [16]. Stricker [15] has shown that using the RGB color space for color description and the L1 norm for evaluating histogram similarity may produce false negatives (i.e. not all the images that are similar to the query are retrieved), while the L2 norm may result in false positives (i.e. images that are not similar to the query are retrieved).

The L1 norm defines the distance between two histograms, Q and D , as:

$$d_{L1}(Q,D) = \sum_i |q_i - d_i|$$

The L2 norm defines this as:

$$d_{L2}(Q,D) = \sqrt{\sum_i |q_i - d_i|^2}$$

Hafner et Al. [6] have proposed the use of a weighted distance between histograms that takes into account the "cross-talk" between colors. However, their work has focused on the efficiency of the solution and only marginally considered the necessity of coding the perceptual similarity between colors. According to Hafner the distance between histograms Q and D is defined as:

$$d_H(Q,D) = \sqrt{\sum_i \sum_j a_{ij} (q_i - q_j)(d_i - d_j)}$$

where a_{ij} , coding the similarity between the color i and j , is expressed as:

$$a_{ij} = 1 - \frac{d_{ij}}{\max_{ij}(d_{ij})}$$

and d_{ij} represents the Euclidean distance between the colors as defined in a variant of the Munsell color space [10].

Alternatively, Hafner et Al. [6] have also proposed an alternative choice for a_{ij} :

$$a_{ij} = \exp(-\sigma(d_{ij}/d_{\max}))^2$$

for some positive constant value of σ .

In earlier studies we have addressed color range and image indexing [2,3] using perceptual correlates of the psychological dimensions of Lightness, Chroma, and Hue, defined in the CIELAB-LUV color space. We elicited, with a structured interview technique, three primary fuzzy sets corresponding to similarity in lightness (L^*), hue (h^*) and chroma (C^*) color dimensions.

The membership functions defined, $\mu_{h^*}(h_i^* - h_j^*)$, $\mu_{L^*}(L_i^* - L_j^*)$, and $\mu_{C^*}(C_i^* - C_j^*)$, associated a degree of similarity between two colors i and j as a function of their difference in a given color feature, while the product of these three membership functions was used to define the global degree of similarity between the two colors considered:

$$\mu_{h^*L^*C^*}(i,j) = \mu_{h^*}(h_i^* - h_j^*) \mu_{L^*}(L_i^* - L_j^*) \mu_{C^*}(C_i^* - C_j^*)$$

Extending these studies, the similarity between two histograms is defined here as:

$$d_F(Q,D) = \sqrt{\sum_i \sum_j \mu_{h^*L^*C^*}(i,j) (q_i - q_j)(d_i - d_j)}$$

4. EXPERIMENTAL RESULTS

The database used in our present experiment is composed of over 200 images taken from the multimedia catalogue "Ancient textile collection of Poldi Pezzoli Museum" developed at our Institute [4]. The catalogue contains mainly images of antique textiles. Although these images vary greatly in color, pattern, and texture complexity, only color distribution has been considered in our experiment.

After a preliminary phase in which the user was allowed to familiarize with the full database, ten query images were taken at random from the database, and s/he was asked to assess the similarity of each to all the other images of the databases on the basis of color similarity. For each query image the matching mechanism compared the histogram of the query with all the stored histograms assigning each a similarity score.

As the effectiveness of the method was our main concern, and the database used was small, the target images were found by sequential scanning of the stored histograms. (A filter could be applied in the case of a larger database to drastically reduce the number of images to be evaluated according to the proposed method [6]) The images most similar to the query were then displayed, ranked from the best match to the Tth best match (T being a user-settable parameter) applying, in turn, the L1 norm, the L2 norm, the two weighted L2 norms, and our method as the matching mechanism.

The evaluation of the effectiveness of system performance has not been expressed in terms of recall

$$\text{Recall} = \frac{\text{relevant retrieved}}{\text{all relevant}}$$

which measures the ability of the system to retrieve useful documents, and precision

$$\text{Precision} = \frac{\text{relevant retrieved}}{\text{all retrieved}}$$

which, conversely, measures the ability to reject useless ones. This would have required that the user set a threshold below which the histograms (i.e. the images) were considered similar [8], and we have already found that this type of threshold has a context-sensitive value [2,3].

To evaluate the effectiveness of system performance for each method, we have, instead, applied a measure proposed in a similar application by Methre et Al. [9]. We let T be the number of relevant items the user wants to retrieve when posing a query, N the total number of relevant images, and n the number of relevant images retrieved in the short list. The effectiveness measure is defined as:

$$\text{Effectiveness} = \begin{cases} n/N & \text{if } T > N \\ n/T & \text{if } T \leq N \end{cases}$$

In Table I we have summarized the experimental results for the 10 queries considered and for the different length short lists for the five methods we tested:

	T=1	T=3	T=5	T=7
Average	0.63	0.5	0.47	0.61
Best query	1	1	0.75	1
Worst query	0	0.33	0.4	0.33

a) L1 norm

	T=1	T=3	T=5	T=7
Average	0.75	0.54	0.52	0.52
Best query	1	1	0.8	0.75
Worst query	0	0.33	0.2	0.4

b) L2 norm

	T=1	T=3	T=5	T=7
Average	0.87	0.46	0.39	0.51
Best query	1	1	0.75	0.86
Worst query	0	0	0.2	0.2

c) First weighted L2 norm

	T=1	T=3	T=5	T=7
Average	0.87	0.46	0.41	0.59
Best query	1	1	0.75	1
Worst query	0	0	0.2	0.4

d) Second weighted L2 norm ($\sigma=2$)

	T=1	T=3	T=5	T=7
Average	0.87	0.58	0.46	0.54
Best query	1	1	1	1
Worst query	0	0.33	0.2	0.4

e) Our method

Tab. I: Effectiveness of different image retrieval strategies: experimental results

It can be seen, that although our method did produce the best results, the other methods did not register significantly different values.

We also noted that there was considerable disagreement among observers in evaluating color similarity, and that the set of similar images found by browsing the original images was far from concurring with that obtained by browsing the randomized version of the database (where the original image structure was changed, but not the color distribution): the percentage agreement, averaged over the queries considered, was about 65% [1].

5. CONCLUSIONS

We have presented an image retrieval strategy based on the fuzzy evaluation of color image similarity. Experimental results on a textile image database show that the strategy effectively retrieves documents of interest. However, these results also suggest that for some images color information can not be assessed by observers independently of other perceptual features, such as shape and texture: a more "global", or even semantic similarity evaluation is performed.

The experiment is currently being assessed with a larger, more statistically significant panel of at least thirty observers.

6. ACKNOWLEDGEMENTS

All the experiments described here used a prototype version of *Quicklook*, a public domain image server that we are developing at ITIM-CNR on a grant from of Hewlett-Packard Italy. We sincerely thank Dr. Anna Della Ventura and Ing. L. Coslovi for their helpful criticisms and Dr. Marco Suardi for his collaboration.

7. REFERENCES

1. B. Barolo, I. Gagliardi, R. Schettini, An effective strategy for querying image databases by color distribution, *Proc. of the EVA' 96 Florence (Electronic Imaging and the Visual Arts) Conference*, Firenze, 8-9 feb., 1996 (in print).
2. E. Binaghi, A. Della Ventura, A. Rampini, R. Schettini, A fuzzy reasoning approach to similarity evaluation in image analysis, *International Journal of Intelligent Systems*, Vol. 8(7), pp. 749-769, 1993.
3. E. Binaghi, I. Gagliardi, R. Schettini, Image retrieval using fuzzy evaluation of color similarity, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 8(4), pp. 945-968, 1994.
4. P. Carrara, I. Gagliardi, D. Gussoni "Antique Textiles of the Poldi Pezzoli Museum of Milan: the multimedia catalogue" *Proc. of the EVA' 96 Florence (Electronic Imaging and the Visual Arts) Conference* Firenze 8-9 febbraio 1996 (in print).
5. C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petrovic "Efficient and effective querying by image content" *Journal of Intelligent Systems*, Vol. 3, pp. 231-262, 1994.
6. J. Hafner, H.S. Sawhney, W. Esquitz, M. Flickner, W. Niblack, Efficient color histogram indexing for quadratic form distance functions, *IEEE Trans. Pattern Analysis and machine Intelligence*, Vol. PAMI 17, pp. 729-736, 1995.
7. R.W.G. Hunt "Revised colour-appearance model for related and unrelated colours", *Color Research and Application*, Vol. 16, pp. 146-165, 1991.
8. M.J. McGill, G. Salton "Introduction to modern Information Retrieval", **McGraw-Hill**, 1983.
9. B.M. Mehtre, M.S. Kankanhalli, A. Desai Narasimhalu, G.C. Man, Color matching for image retrieval, *Pattern Recognition Letters*, Vol. 16, pp. 325-331, 1995.
10. M. Miyahara, Y. Yoshida "Mathematical transform of (RGB) color data to Munsell (HVC) color data" *Proc. SPIE Visual Communications and Image Processing*, Vol. 1001, pp. 650-657, 1988.
11. N. Otha "Development of color targets for input scanner calibration" *Proc. 7th Congress of the International Color Association "Colour 93"*, Technical University of Budapest, Budapest (Hungary), pp. 130, 1993.
12. R. Schettini, The faithful rendition of color ranges on display, *Proc. IS&T and SID's Color Imaging Conference: Transforms & Transportability of Color*, USA, pp. 160-163, 1993.
13. R. Schettini, B. Barolo, E. Boldrin, Colorimetric calibration of color scanners by back propagation, *Pattern Recognition Letters*, Vol. 16, pp. 1051-1056, 1995.
14. J.R. Smith, S-F Chang "Single color extraction and image query" *Proc. Int. Conference on Image Processing (ICIP-95)* Washington DC, 1995.
15. M. Stricker, Bounds for the discrimination power of color indexing techniques, *Proc. SPIE* Vol. 2185, pp. 15-24, 1993.
16. M.J. Swain, Color Indexing, **Technical Report 360**, University of Rochester, Rochester, New York, 1990.
17. G. Wyszecki, W.S. Stiles, "Color science: concepts and methods, quantitative data and formulae", **Wiley**, New York, 1982.