Soft Color Signatures for Image Retrieval by Content

N. Boujemaa
INRIA Rocquencourt
IMEDIA project
BP105, 78153
Le Chesnay Cedex France
Nozha.Boujemaa@inria.fr

S. Boughorbel
INRIA Rocquencourt
IMEDIA project
BP105, 78153
Le Chesnay Cedex France
Sabri.Boughorbel@inria.fr

C. Vertan
INRIA Rocquencourt
IMEDIA project
BP105, 78153
Le Chesnay Cedex France
Constantin.Vertan@inria.fr

Abstract

Content-based image retrieval primarily used color distributions as descriptors of the image content; researches have since focused on the use of various color representation spaces, color and illumination invariance, color quantization and color matching. In order to overcome the many limitations of the description by a first-order distribution, several higher-order distributions have been introduced since (like autocorrelogram or color coherence vectors). Although they can perform better, their computational complexity is prohibitive and they require parameter setting. We propose to upgrade the first order color distribution (color histogram) by embedding for each color additional information about its perceptual or statistical relevance. Such information is obtained by using local activity measures such as the Laplacian, the entropy and others. Histograms computed on windows and combined by different ways of accumulation improve the information on geometric repartition of colors. We prove that the new color distribution family is compact, robust and easy to compute and provides a superior retrieval performance, independent with respect to the color representation.

1 Introduction

Content-based image retrieval (CBIR) became a must in the last decade. Powered by the explosive development of the Internet, Web and the continuously cheaper digital imagining devices and technologies, applications such as digital libraries, image archives, video-on-demand and specific image databases emerge as a real-life fact. The basic idea of the CBIR process is to compactly describe an image by a digital signature and then match the query image to the most resemblant image within the database according to the similarity of their signatures. Traditionally, the content description is done (for either global or partial queries) according to the notions of color and texture. Thus the signatures are color distributions (histograms [1], color moments [2], color coherence vectors [3]), second-order, spatially constrained color distributions (color correlograms [4], edge correlograms [5]) or classical textural descriptors (Fourier coefficients [6], wavelet coefficients, Markov random field parameters, etc. [7]).

Starting with the works of Swain [1], the color distribution (histogram) became the main feature descriptor for image content. Given a color image f, of size M by N pixels, characterized by the color $\mathbf{c}$ at location (i,j), i.e. $\mathbf{c} = f(i,j)$, the color distribution (histogram) of the color set $\mathcal{C}$ is given by:

$$h(\mathbf{c}) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta\left(f(i,j) - \mathbf{c}\right), \ \forall\, \mathbf{c} \in \mathcal{C} \tag{1}$$

In the equation above $\delta()$ is the unitary impulse function. We notice that the $h(\mathbf{c})$ values are normalized in order to sum to one. The value of each bin is thus the number of image pixels having the color $\mathbf{c}$, or, after normalization by MN, the probability that the color $\mathbf{c}$ appears in the image. Thus, in the classical histogram defined by (1), any pixel (regardless the color representation space) contributes with a constant weight (1 before normalization, or $1/MN$ after the probabilistic normalization), invariant with respect to the local image context.

Pixels having the same color are generally not similar (since they can represent corners, edges, uniform areas, etc.); according to this observation, Pass et al proposed in [3] to classify the pixels into two classes: coherent and contour. A coherent pixel is defined by the color uniformity of its neighbourhood, whereas the contour pixel is situated close to the separation lines between the image objects and thus it is characterized by a non-uniform neighbourhood. The color coherence vector (CCV) proposed in [3] is a separate counting of contour and coherent pixels, into two color distributions (histograms). Still, there is no further distinction for the pixels of any of the two classes, the contour/ coherent decision needs the definition of some parameters (that are not necessarily constant for different images) and the color distribution is twice the size of the classical histogram. In this contribution we will revisit the use of color histograms from the perspective of embedding some local information about the statistical and visual relevance or importance of each pixel. In the following section we will describe the proposed modified histogram – the weighted color histogram and the various measures that describe the local behaviour of the colors. Section 3 will describe experimental results and section 4 will summarize the conclusions of this work.

2 The weighted histogram and weighting schemes

The approach proposed by the CCV [3] can be further refined by the classification of the image pixels in more than two classes, according to a local attribute (such as the edge strength). We can easely imagine a classification in three classes, consisting of pixels charaterized by a small, medium and high edge strength. The number of classes is thus related to the number of quantization level of the pixel attribute's. At the limit, since every pixel has acquired a supplementary, highly relevant characteristic, we can easely imagine a one pixel per class approach, which will certainly provide a very accurate description of the image, but will require a very important size. In order to keep the balance between the histogram size and the discrimination between pixels we propose to adaptively weight the contribution of each pixel of the image into the color distribution. This individual weighting allows a finer distinction between pixels having the same color and the construction of a weighted histogram that accounts both color distribution and statistical non-uniformity measures. Thus, we will use a modified histogram, defined as:

$$\tilde{h}(\mathbf{c}) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} w(i,j)\delta\left(f(i,j) - \mathbf{c}\right), \ \forall \mathbf{c} \in \mathcal{C} \tag{2}$$

In the equation above $w(i,j)$ is the weighting coefficient of the color at spatial position (i,j). We may notice that, since $w(i,j)$ must be a scalar, we cannot use any color statistics (which are necessarily vector triples). Intuitively the accounting within the color distribution of some local measures of each pixel could be considered as a way of integrating both color and texture, provided that the local measure have a textural background. In the following subsections we will describe the proposed weighting schemes: by the edge strength (Laplacian), by probabilistic measures and by fuzzy measures. Indeed, the probabilistic measures describe the local degree of chaos in the coloring; the fuzzy measures reveal the pertinence of the color of the current pixel with respect to its neighbourhood and the edge strength measures both the local non-uniformity and the presence of strong visual cues (as edges or corners). Thus, the weighting is to be computed on a neighborhood of the current pixel and must quantify the color activity (or visual importance, or non-uniformity) of that neighbourhood. That implies that $w(i,j)$ is increasing with the color non-uniformity.

2.1 Laplacian-based measures

Visual perception studies have proven that edges in general, and corners in particular, are very important to scene analysis. Such information is provided by the image Laplacian $\Delta(i,j)$ [8]. The color Laplacian operator $\Delta(i,j)$ that we will subsequently use is the L_2 aggregation of the scalar Laplacian operators computed on each color component of the image. A weighted histogram similar to (2) was presented in [9] as a tool for image thresholding; if $\Delta(i,j)$ is a Laplacian operator computed at location (i,j), the thresholding histograms were defined as:

$$\tilde{h}(\mathbf{c}) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} \delta\left(f(i,j) - \mathbf{c}\right) \frac{1}{1+\Delta(i,j)}, \ \forall \mathbf{c} \in \mathcal{C}, \quad (3)$$

or

$$\tilde{h}(\mathbf{c}) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} \delta\left(f(i,j) - \mathbf{c}\right) \Delta(i,j), \ \forall \mathbf{c} \in \mathcal{C} \quad (4)$$

The relation (3) emphasizes the weight of pixels that belong to constant (uniform) regions: their Laplacian is very small, so they sum with an unitary weight; the pixels placed on the edges are characterized by an important Laplacian and thus their contribution to the corresponding $\mathbf{c}$ bin is very small. This behavior is thought to reduce the influence of the uncertain colors, situated at the border between different objects and is derived from the gray-scale image case of choosing the segmentation thresholds as the minima of the histogram. The relation from (4) corresponds to a dual behavior, counting the colors proportionally to their edge strength. Since the edge information is highly relevant, we finally have chosen the weighting coefficient in (2) as the squared Laplacian:

$$w(i,j) = \Delta^2(i,j). \quad (5)$$

2.2 Probabilistic measures

Since the key factor in the evaluation of the local color activity is the color variability within some neighbourhood, the probability of occurrence of the current color $\mathbf{c} = f(i,j)$ within its neighbourhood could be a good estimate. If this probability is high, the neighbourhood include a significant uniform area, colored with $\mathbf{c}$ and thus the contribution of the color to the modified histogram must be reduced. If the probability of finding color $\mathbf{c}$ within its spatial neighbours is small, it follows that color $\mathbf{c}$ is rather singular (corner, isolated point), and thus its contribution to the modified histogram (2) must be increased.

Thus, we propose to use a weighting coefficient that is inverse proportional with the number of pixels $N_{ij}(\mathbf{c})$ having the same color $\mathbf{c}$ within the square window of size D, centered at the current location (i,j) (with $\mathbf{c} = f(i,j)$). We define $N_{ij}(\mathbf{c})$ as:

$$N_{ij}(\mathbf{c}) = \sum_{m=-D/2}^{D/2}\sum_{n=-D/2}^{D/2} \delta(f(i+m,j+n) - f(i,j)). \quad (6)$$

Thus, the weighting coefficient in (2) is:

$$w(i,j) = \frac{1}{N_{ij}(\mathbf{c})}. \quad (7)$$

Obviously the simple probabilistic approach defined by (7) measures the local non-uniformity form the current color's point of view. A global measure must take into account the occurrence probabilities for all the colors within the neighbourhood. The informational entropy could thus be used in order to measure the overall dissimilarity. Since the entropy is maximal when the probabilities are equal (and thus there are no identical colors within the neighbourhood) and is zero if the neighbourhood is absolutely uniform, we propose to use the informational efficiency (entropy to maximal entropy ratio) as a weighting coefficient of the colors:

$$w(i,j) = 1 - \frac{\sum_{\mathbf{c}' \in \mathcal{C}} N_{ij}(\mathbf{c}') \log N_{ij}(\mathbf{c}')}{2D^2 \log D}. \quad (8)$$

2.3 Fuzzy measures

From a numerical point of view, the usual histogram h in (1) maps the color set $\mathcal{C}$ into the interval $[0,1]$. According to Zadeh's [10] theory, such a function is a fuzzy set. In [11], Bezdeck further noticed that any such function can be a fuzzy set, but actually becomes a fuzzy set if and only if it fits a semantically plausible description for the properties of the object (the color in particular) within the universe (the color set $\mathcal{C}$). Thus, the normalized histogram from (1) cannot be a fuzzy set, as its semantical description is void from the uncertainty point of view. Digital images are mappings of natural scenes (sampled and quantized slices of 3-d reality) and thus have an important amount of uncertainty, in both value and location (spatial support) [12]. We will focus on the imprecise nature of the pixel values, since, in a complex, natural scene, most likely there is no perceivable difference between the gray levels of 99 and 100 (as suggested in [13]), or the colors - expressed as RGB triples - (201,100,199) and (200,100,199) (used for color image filtering, as suggested in [14]).

The simplest approach [15] is to normalize the usual histogram in (1) by the value of its largest bin, in such way that the most probable color will have a membership degree of 1 within the fuzzy set "image". The most predominant color can be thus considered as the most typical for the given image and the constructed fuzzy histogram (9) measures the typicality of a color within the image. For the entire image this leads to the fuzzy histogram h_1 [15], [16]:

$$h_1(\mathbf{c}) = \frac{h(\mathbf{c})}{\max_{\mathbf{c}' \in \mathcal{C}} h(\mathbf{c}')}, \ \forall \mathbf{c} \in \mathcal{C}. \qquad (9)$$

The normalization by the mode from (9) allows us to assign the maximum, unitary typicality to the color that is dominant within the image, regardless its probability of appearance (which is not the case of the usual histogram). The discussion from the previous subsection still holds if we are replacing, for the current color, the probability of occurrence within its neighbourhood by its typicality with respect to the same neigbourhood. If we denote by $N_{\max}$ the maximum number of pixels having the same color in the neighbourhood of pixel (i,j) (that is $N_{\max} = \max_{\mathbf{c}' \in \mathcal{C}} N_{ij}(\mathbf{c}')$), we have the following two expressions for the weighting coefficient $w(i,j)$:

$$w(i,j) = \frac{N_{\max}}{N_{ij}(\mathbf{c})}. \qquad (10)$$

$$w(i,j) = \frac{D^2 \log N_{\max} - \sum_{\mathbf{c} \in \mathcal{C}} N_{ij}(\mathbf{c}) \log N_{ij}(\mathbf{c})}{N_{\max} \log D^2}. \qquad (11)$$

Equation (10) is the fuzzy variant of (7). The weighting by the fuzzy entropy from (11) reduces to the entropy weighting from (8) if $N_{\max} = D^2$.

3 accumulative histograms

color histogram gives a global statistical information on color repartition in the image. Figure 1 show two images having the same color histograms. if we divide the images into windows, we notice that the information of local histograms computed on each window is richer than the global color histogram. In fact the first window of the right image contains much more of black pixels than other windows. If we combine informations on colors from the local histograms it will be possible to characterize better the geometric repartition of colors. The idea of accumulative histograms is to embed geometric information on colors by emphasizing their local agglomeration in windows.

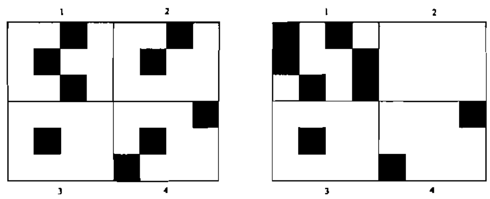

Figure 1: two images with the same color histogram but with different geometric repartitions of colors.

3.1 mutliplicative accumulation

We introduce a multiplicative accumulation of local histograms computed on windows by:

$$\tilde{h}(c) = h(c) \prod_{i=1}^{M} (1 + h_i(c)), \ \forall c \in \mathcal{C} \qquad (12)$$

where h_i is the color histogram computed on the i^{th} window. M is the total number of windows. The weighting factor $1 + h_i(j)$ is as important as the presence of the color c in the i^{th} window is high.

3.2 additive accumulation

We introduce an additive accumulation of local histograms by :

$$\tilde{h}(c) = \sum_{i=1}^{N} f(h_i(c)), \ \forall c \in \mathcal{C} \qquad (13)$$

Where f is a function that emphasizes the local presence of colors. the L_1 distance between two histograms can be written:

$$\begin{aligned} d_{L1}(\tilde{h}^1, \tilde{h}^2) &= \sum_{i=1}^{M}\sum_{j=1}^{N} |f(h_i^1(c)) - f(h_i^2(c))| \\ &= \sum_{i=1}^{M}\sum_{j=1}^{N} |f'(\alpha_i(c))||h_i^1(c) - h_i^2(c)| \end{aligned} \qquad (14)$$

The local contribution of i^{th} window in the bin

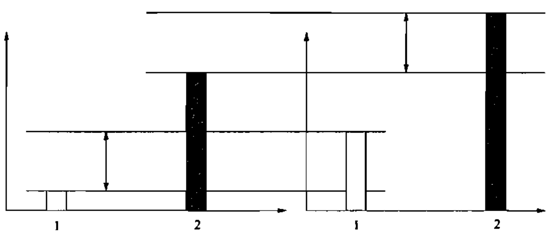

Figure 2: two histograms with the same respective difference of bins.

c is as high as $h_i(c)$ is great. So we choose f as an increasing function. Figure 2 show 2 histograms with only 2 bins. The respective difference between bins are the same but the bins 2 are higher. We can say that visual similarity between the color of the bin 2 is more important than the bin 1. So the function f should reduce the difference between higher bins. From equation (14) a sufficient condition to satisfy is that $|f'|$ should be a decreasing function. We have chosen power function $f(x) = x^m$ $0 < m < 1$, this choice satisfy the two conditions that we imposed. $f'(x) = mx^{m-1}$ is a positive function and $|f'|'(x) = m(m-1)x^{m-2}$ is a negative one. An experimental study leads us to choose $m = \frac{1}{3}$ for the parameter of the function f.

4 Experiments

The experiments performed in order to establish the retrieval capabilities of the proposed fuzzy histograms were conducted using the Ikona software platform [17] developed at IMEDIA. We investigated both the objective and the subjective retrieval quality, on two different, heterogeneous, generalist image databases. A first, small, image database consists of 210 key frames from a television broadcast (part of the AIM corpus of INA), manually grouped into similarity classes. A second image database consists of 792 color textures from both regular and irregular textures (part of the textures are from the Vistex database at MIT MediaLab). We tested the retrieval capabilities of the proposed histograms for various color representation spaces. We finally selected the RGB, HSV and Lab color space as prototypical: RGB is the primarily acquisition space, HSV is the preferred natural-language color description mode and Lab models the perceptual inter-color difference by the L_2 metric. No color invariance models were considered for the moment, although some simple and powerful models have been proposed [18], [19].

Figures 3 and 4 show retrieval results within the used image databases; the weighted histogram performs clearly better (with increased recall and precision rates).

Figures 6 and 7 show the precision-recall curves in the mentioned color spaces for the usual color distribution (1) and the weighting-updated color distributions (2).

Figures 8 shows the precision-recall curves of the usual histogram and the accumulative histograms (additive and multiplicative accumulation). All the modified histograms perform well, providing an increased effectiveness with respect to the classical color distribution. The improvement is obtained regardless the color representation.

5 Conclusions

This contribution proposes the upgrade of the usual color distribution (histogram) by an adaptive weighting of each pixel's contribution. The weighting is related to a local measure of color

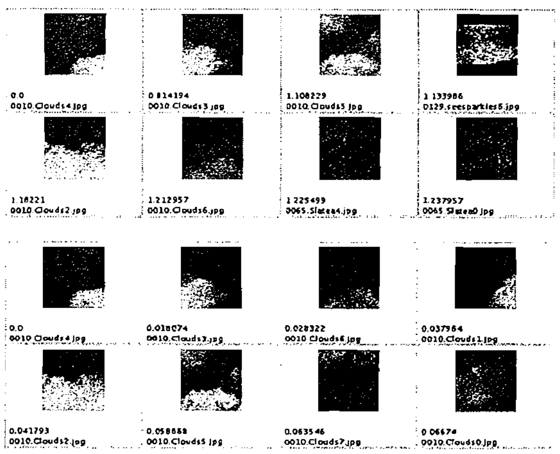

Figure 3: Image retrieval for the same query image (top left, blue contour highlighting) using the RGB uniformly quantized color space and L_1 metric. Upper picture block - retrieval by the usual histogram. Lower picture block - retrieval by the Laplacian weighted histogram, providing an increased accuracy. The images are presented in the order of decreasing similarity, from left to right and top to bottom. Thus, for the given query, the retrieval by the usual RGB histogram provides a 60 % accuracy, and the retrieval by the laplacian weighted histogram provides a 100 % accuracy. The keyframes of the television broadcast are available by courtesy of INA - Institute National de l'Audivisuel of France, who kindly provided the image database.

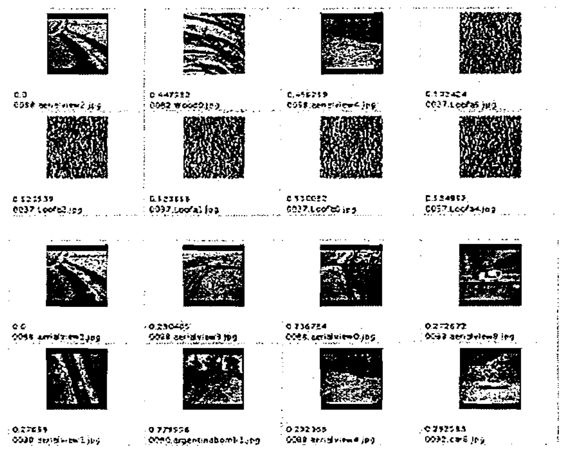

Figure 4: Image retrieval for the same query image (top left, blue contour highlighting) using the RGB uniformly quantized color space and L_1 metric. Upper picture block - retrieval by the usual histogram. Lower picture block - retrieval by the Laplacian weighted histogram, providing an increased accuracy. The images are presented in the order of decreasing similarity, from left to right and top to bottom. Thus, for the given query, the retrieval by the usual RGB histogram provides a 25 % accuracy, and the retrieval by the probability weighted histogram provides a 75 % accuracy.

non-uniformity (or color activity), computed within a neighbourhood of the pixel. The proposed non-uniformity measures are based on the evaluation of perceptual cues (corners and isolated colors, by the use of the Laplacian), statistical color area distribution (by the use of local probability of occurrence and inormational entropy) and local color relevance (by a fuzzy typicality and fuzzy entropy). The magnitude of all these measures increases with the local color variability, being minimal for uniform regions. The introduced method of accumulative histogram characterize well geometric repartition of colors within the different windows on which local histograms are computed. The performance of these modified histograms are quite better than usual histogram.

The objective quality measures (precision-recall curves) show that the proposed approaches perform better than the usual color histogram, regardless the color representation space. The new histograms have the same size as the usual color distribution and can be compared by the same metrics and its computational complexity is not excessive. Thus we claim that the weighted and the accumulative histograms can be indeed a valuable upgrade to the traditionally color distribution based image retrieval.

References

[1] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[2] B. M. Mehtre, M. S. Kankanhalli, A. D. Narasimhalu, and G. C. Man, "Color match-

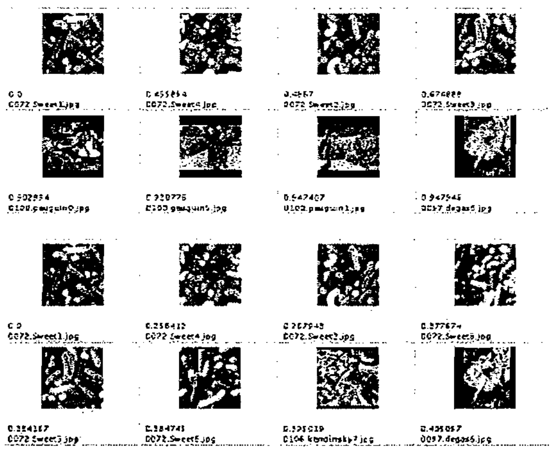

Figure 5: Image retrieval for the same query image (top left, blue contour highlighting) using the RGB uniformly quantized color space and L_1 metric. Upper picture block - retrieval by the usual histogram. Lower picture block - retrieval by the Laplacian weighted histogram, providing an increased accuracy. The images are presented in the order of decreasing similarity, from left to right and top to bottom. Thus, for the given query, the retrieval by the usual RGB histogram provides a 50 % accuracy, and the retrieval by the accumulative histogram (additive) provides a 75 % accuracy. The keyframes of the television broadcast are available by courtesy of INA - Institute National de l'Audivisuel of France, who kindly provided the image database.

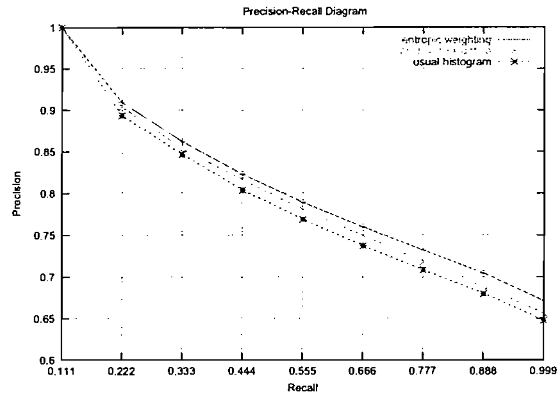

Figure 6: Precision-recall curves for the retrieval within the texture image database, using RGB color representation, uniform 6 bins/ component quantization and various color distributions: usual color histogram (1), Laplacian weighted (5) histogram and entropy weighted (8) histogram. The later two curves significantly overcome the usual color distribution.

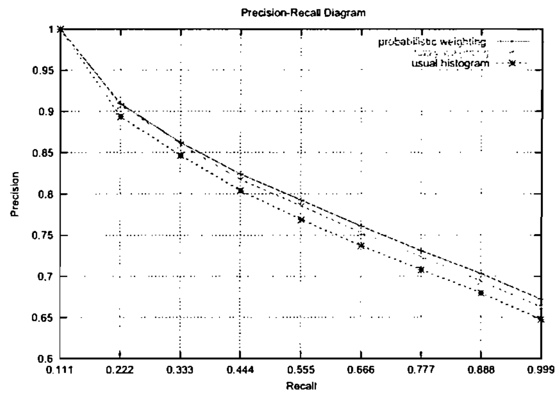

Figure 7: Precision-recall curves for the retrieval within the key-frames image database, using RGB color representation, uniform 6 bins/ component quantization and various color distributions: usual color histogram (1), probability weighted (7) histogram and typicality weighted (10) histogram. The later two curves significantly overcome the usual color distribution.

ing for image retrieval," *Pattern Recognition Letters*, vol. 16, pp. 325–331, Mar. 1995.

[3] G. Pass and R. Zabih, "Histogram refinement for content based image retrieval," in *IEEE Workshop on Applications of Computer Vision*, 1996, pp. 96–102.

[4] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using correlograms," in *Computer Vision and Pattern Recognition CVPR '97*, San Juan, Puerto Rico, 17-19 Jun. 1997.

[5] J. Huang, S. R. Kumar, M. Mitra, and W.-J. Zhu, "Spatial color indexing and applications," in *IEEE International Conference on Computer Vision ICCV '98*, Bombay, India, 4-7 Jan. 1998.

[6] C. Vertan and N. Boujemaa, "Color texture classification by normalized color space representation," in *Proc. of ICPR'2000*, Barcelona, Spain, 3-8 Sept. 2000, vol. 3, pp. 584–587.

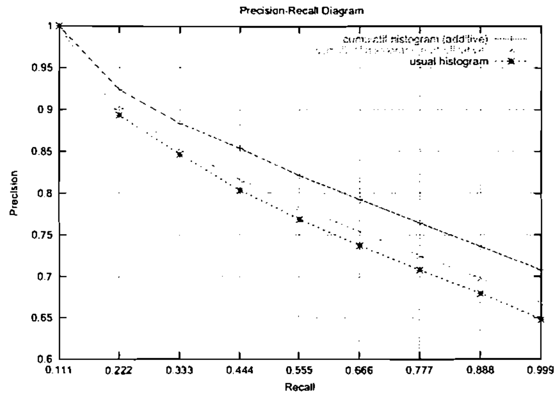

Figure 8: Precision-recall curves for the retrieval within the key-frames image database, using *RGB* color representation, uniform 6 bins/ component quantization and various color distributions: usual color histogram (1), accumulative histograms (multiplicative (13) and additive (12)). The later curve significantly overcomes the usual color distribution.

[7] A. del Bimbo, *Visual Information Retrieval*, Morgan Kaufmann, San Francisco, CA, 1999.

[8] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall Intl., Englewood Cliffs NJ, 1989.

[9] F. M. Wahl, *Digital Image Signal Processing*, Artech House, Boston, 1987.

[10] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[11] J. C. Bezdek, "Fuzzy models - what are they and why ?," *IEEE Trans. on Fuzzy Systems*, vol. 1, no. 1, pp. 1–5, Feb. 1993.

[12] I. Bloch and H. Maitre, "Fuzzy mathematical morphologies: a comparative study," *Pattern Recognition*, vol. 28, no. 9, pp. 1341-1387, Sept. 1995.

[13] C. V. Jawahar and A. K. Ray, "Fuzzy statistics of digital images," *IEEE Signal Processing Letters*, vol. 3, no. 8, pp. 225–227, Aug. 1996.

[14] C. Vertan and V. Buzuloiu, "Fuzzy nonlinear filtering of color images: A survey," in *Fuzzy Techniques in Image Processing*, E. Kerre and M. Nachtegael, Eds., Heidelberg, Germany, 2000, pp. 248–265, Physica Verlag.

[15] C. Vertan and N. Boujemaa, "Using fuzzy histograms and distances for color image retrieval," in *Proc. of CIR'2000*, Brighton, United Kingdom, 4-5 May 2000.

[16] C. Vertan and N. Boujemaa, "Embedding fuzzy logic in content based image retrieval," in *Proc. of NAFIPS'2000*, Atlanta, Georgia, 13-15 Jul. 2000, pp. 85–90.

[17] M. Ferecatu N. Boujemaa F. Fleuret, "Ikona-a modular architecture for content based image retrieval systems," *Systemics Cybernetics and Informatics SCI2001*, July 22-25 2001.

[18] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 522–529, May 1995.

[19] T. Gevers and A. W. M. Smeulders, "A comparative study of several color models for color image invariant retrieval," in *First Inetrnational Workshop on Image Databases and Multimedia Search*, Amsterdam, Holland, 1996, pp. 17–26.