ROBUST COLOR IMAGE RETRIEVAL FOR THE WORLD WIDE WEB

B. Smolka, M. Szczepanski

Silesian University of Technology,
Department of Automatic Control,
Akademicka 16 Str, 44-100 Gliwice, Poland

R. Lukac, A.N. Venetsanopoulos

The Edward S. Rogers Sr. Department
of Electrical and Computer Engineering,
University of Toronto,
10 King's College Road,
Toronto ON, M5S 3G4, Canada

ABSTRACT

The rapid growth of image archives increases the need for efficient and fast tools that can retrieve and search through large amount of visual data. In this paper we propose an efficient method of extracting the image color content, which serves as an image digital signature, allowing to efficiently index and retrieve the content of large, heterogeneous multimedia Internet based databases. We apply the proposed method for the retrieval of images from the *WEBMUSEUM* Internet database, containing the collection of fine art images and show that the new method of image color representation is robust to image distorsions caused by resizing and compression and can be incorporated into existing retrieval systems which exploit the information on color content in digital images.

1. INTRODUCTION

Successful queries on large, distributed databases cannot rely on textual information and therefore color image indexing is one of the methods used for automatic content based image retrieval. In this paper we focus on the image indexing based on the global color distribution, which is applied for cases when the user provides a sample image for the query.

The majority of the systems exploiting the image color information work using various kinds of color histograms. However the histogram based approach has many drawbacks, as the histogram representation is sensitive to illumination changes, image resizing through interpolation and compression induced artifacts. Therefore, in this paper we propose a nonparametric approach to the problem of the estimation of the distribution of image colors.

2. COLOR HISTOGRAMS

Color indexing is a process through which the images in a database are retrieved on the basis of their color content. The indexing process must enable the automatic extraction of features, efficient assigning of digital signatures to images and effective retrieval of images within a database.

In order for an image retrieval system to retrieve images that are visually similar to the given query, a proper representation of the visual features is needed and a measure that can determine similarity between a given query and the images from a database set has to be chosen. Assuming that no textual information about the image content are given, image features such as color [1, 2, 3], texture [4, 5] and shape [6, 7] are commonly used.

These features are dependent on illumination, shading, resizing manipulations and compression induced artifacts. Thus, the visual appearance of an image is better described by the distribution of features, rather than by individual feature vectors.

Color feature has proven to be efficient in discriminating between relevant and non-relevant images. One of the widely used tools for image retrieval is the color histogram, which describes the distribution of colors in an image using a specific color space. The colors of an image are mapped into a discrete color space containing m colors. In this way, a color histogram is a m-dimensional vector, whose elements represent the number of pixels of a given color in an image.

In this paper we use the RGB color space, which although not perceptually uniform, is the most commonly used, primarily to retain compatibility with computer display systems. Let us assume a color image $\mathbf{Q}$ of size $n_1 \times n_2 = N$, composed of three RGB channels $\mathbf{Q} = \{Q^R_{i,j}, Q^G_{i,j}, Q^B_{i,j}\}$, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$.

An image histogram in the RGB color space $H(r, g, b)$ is the approximation of the density function of the image RGB channels intensities

$$H(\rho, \gamma, \beta) = N^{-1} \# \left\{ Q^R_{i,j} = \rho, Q^G_{i,j} = \gamma, Q^B_{i,j} = \beta \right\},$$

where N is the total number of image pixels, and $\#$ denotes the number of pixels with color channel values $\{\rho, \gamma, \beta\}$.

For the analysis of colors, independent of image brightness, it is convenient to transform the RGB values into normalized components r, g, b defined as: $r = R/I,\ g = G/I,\ b = B/I,\ I = R{+}G{+}B$, where $R, G, B \in [0, 255]$. The normalized color values can be expressed using only r and g values as $g = 1 - r - b$ and the normalization makes the r, g variables non-dependent on the brightness value I.

Using the normalized rg reduced color space, we can map the color pixel on a two-dimensional plane and obtain a two-dimensional histogram $\Phi(x, y) =$

$$= \frac{\#\left\{\mathrm{int}\left(MQ^R_{i,j}/I_{i,j}\right) = x, \mathrm{int}\left(MQ^G_{i,j}/I_{i,j}\right) = y\right\}}{N} =$$

$$= \frac{1}{N}\#\left\{\mathrm{int}\left(Mr_{i,j}\right) = x, \mathrm{int}\left(Mg_{i,j}\right) = y\right\},$$

$x, y = 0 \ldots M$, where $M + 1$ is the dimension of the 2-dimensional histogram, (for true-color images $M = 255$). The similarity between two images is very often expressed through the similarity of their color histograms. One of the most popular ways to measure the similarity between two histograms H_P and H_Q is the Minkowski distance

$$\Delta(P, Q) = \sum_{x=0}^{M}\sum_{y=0}^{M}\left\{\left[\Phi_P(x, y) - \Phi_Q(x, y)\right]^p\right\}^{\frac{1}{p}}.$$

For $p = 2$ we obtain the Euclidean distance $\Delta(P, Q) = \sum\limits_{x=0}^{M}\sum\limits_{y=0}^{M}\left\{\left[\Phi_P(x, y) - \Phi_Q(x, y)\right]^2\right\}^{\frac{1}{2}}$. Another measure of the similarity of two histograms is the histogram intersection, defined as

$$\Delta(P, Q) = 1 - \sum_{x=1}^{N}\sum_{y=1}^{N}\min\left\{\Phi_P(x, y), \Phi_Q(x, y)\right\}.$$

3. KERNEL DENSITY ESTIMATION OF THE COLOR DISTRIBUTION

The drawback of the histogram representation is that the shape of the histogram strongly depends on the method used for lossy image representation and on the image size or more precisely on the number of image pixels, as for small image sizes there is too few points to build the histogram, which makes that the comparison of histograms is inapplicable.

To alleviate the problems, we propose in this paper to estimate the color distribution not through the discrete histogram, but to use a smooth nonparametric estimate, based on the concept of nonparametric density estimation [8, 9, 10]. In this formulation, the similarity measure between two estimates of the color distribution will be the distance between two surfaces of the two-dimensional kernel density estimation in the normalized rg color space.

Density Estimation describes the process of modelling the probability density function $f(x)$ of a given sequence of sample values drawn from an unknown density distribution. The simplest form of density estimation is the histogram: sample space is first divided into a grid, then the density at the center of the grid cells is approximated by the number of sample values that fall into one bin divided by the width of one grid cell. The main disadvantage of the histogram is the strong dependence of the histogram's shape on the chosen bin-width and the origin of the grid.

Kernel Density Estimation, avoids this disadvantage by placing a kernel function on every sample value in the sample space and then summing the values of all functions for every point in the sample space. This results in a smooth density estimates that are not affected by an arbitrarily chosen partition of the sample space, (Figs. 2, 3). The multivariate kernel density estimator in the q-dimensional case is defined as

$$\hat{f}_{\mathbf{h}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1 \cdots h_q}\mathcal{K}\left(\frac{x_{i_1} - x_1}{h_1}, \ldots, \frac{x_{i_q} - x_q}{h_q}\right),$$

with $\mathcal{K}$ denoting a multidimensional kernel function $\mathcal{K}: \mathbb{R}^q \to \mathbb{R}$ and $h_1, \ldots, h_q$ denoting bandwidths for each dimension and n is the number of samples in the sliding window. A common approach to build multidimensional kernel functions is to use a *product kernel* $\mathcal{K}(u_1, ..., u_q) = \prod_{i=1}^{q} K(u_i)$, where K is a one-dimensional kernel function. Intuitively, the kernel function determines the shape of the 'bumps' placed around the sample values and the bandwidths $h_1, ..., h_q$ their width in each dimension. In case bandwidth is equal for all dimensions, multivariate radial-symmetric kernel functions can be used,

$$\hat{f}_h(\mathbf{x}) = \frac{1}{nh^q}\sum_{i=1}^{n}K\left(\frac{||\mathbf{x}_i - \mathbf{x}||}{h}\right).$$

The shape of the approximated density function depends heavily on the bandwidth chosen for the density estimation. Small values of h lead to spiky density estimates showing spurious features. On the other hand too big values of h produce over-smoothed estimates that hide structural features.

The unknown density function is assumed to be the standard normal distribution re-scaled to have the same variance as the sample values. Choosing the Gaussian kernel function for K, the optimal bandwidth is in the one-dimensional case: $h_{opt} = 1.06\hat{\sigma}n^{-\frac{1}{5}}$, where $\hat{\sigma}$ denotes the standard deviation, and for the q-dimensional case [8, 9]

$$h_{opt} = \left(4/(q + 2)\right)^{\frac{1}{q+4}}\hat{\sigma}n^{-\frac{1}{q+4}},$$

in this paper we use the rg color space, so $q = 2$. In this work we used the Gaussian kernel, although we have obtained similar results using other kernel shapes commonly used in nonparametric density estimation.

Using the kernel based estimation, a smooth estimate of the color distribution is obtained as shown in Figs. 2, 3. As can be seen in these Figures, the density distribution is insensitive to resizing and lossy image coding, which are the basic operations performed when preparing large Internet multimedia databases. This distribution can also be used for the image retrieval purposes, as it can serve as an image signature, as depicted in Figs. 1, which shows the rg distributions of some well known color test images.

4. RESULTS

To evaluate the efficiency of the proposed color density estimation, we used as the testbed the images comprising the well known *WEBMUSEUM* Internet database, (available at http://www.ibiblio.org/wm). This art database contains a collection of about 3000 images of fine arts of various famous artists. Each image is coded in JPEG of moderate compression ratio (the blocking artifacts are hardly visible) with width or height of about 1000 pixels. Each image is accompanied by a thumbnail of width or size of 100 pixels, also compressed with JPEG.

From the database, the image of the painting *Starry Night* of V. van Gogh was chosen as the query image, (see Fig. 4). Using the kernel density estimation, we applied the histogram Euclidean distance as the similarity measure and ordered the retrieved images according to the distance values. The results are very promising, as the second image, most similar to the query, was another painting of van Gogh, (*Road with Cypress and Star*, Fig 4 3).

In the second experiment we used the thumbnail of the *Starry Night* image and used it as query. Although this picture is small (122×100) and heavily jpeged, the proposed scheme was able to find the image of full resolution, (first image in the ordered sequence) and the majority of images retrieved using the full resolution image, (Fig 5). Very similar results, were obtained using the histogram intersection method, so as expected the two methods of similarity evaluation yield comparable results.

5. CONCLUSIONS

In this paper we proposed a robust way of color density estimation. To enable fast retrieval of large databases we used the normalized rg color space. The experiments show that the method of nonparametric density estimation is insensitive to image compression and resizing. This makes the proposed framework interesting for image retrieval applications. Especially, the ability to retrieve images using a heavily distorted thumbnail is interesting, as it enables extremely fast retrieval of large databases.

The new approach can be used to store image signatures in such a way as to facilitate the fastest possible retrieval time in order to facilitate online browsing of Internet databases. In this way a database of image signatures can be built, to enable that only image signatures are being processed, which ensures a low computational load of the image retrieval process.

Fig. 1. Kernel density estimation of the color distribution in the rg color space for some test images.

6. REFERENCES

[1] X. Wan, C.C.J. Kuo, Color distribution analysis and quantization for image retrieval, Proceedings of SPIE, Vol. 2670, February 1996.

[2] Swain M., D. Ballard, Color indexing, International Journal of Computer Vision, 7, 1, 11-32, 1991.

[3] M. Stricker, M. Orengo, Similarity of color images, in SPIE Conference on Storage and Retrieval for Image and Video Databases III, 2420, 381-392, February 1995.

[4] B.S. Manjunath, W.Y. Ma, Texture features for browsing and retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, 8, 837-842, 1996.

[5] B.S. Manjunath, J.R Ohm, V.V. Vasudevan, A. Yamada, Color and texture descriptors, IEEE Trans. CSVT, 11, 6, 703-715, June 2001.

[6] J.E. Gary, R. Mehrotra, Similar shape retrieval using a structural feature index, Information Systems, 18, 7, 525-537, October 1990.

[7] A.K. Jain and A. Vailaya. Image retrieval using color and shape. Pattern Recognition, 29, 8, 1233-1244, 1996.

[8] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Stat., Vol. 33, 1065-1076, 1962.

[9] D.W. Scott, "Multivariate Density Estimation", New York, John Wiley, 1992.

[10] B.W. Silverman, "Density Estimation for Statistics and Data Analysis", London, Chapman and Hall, 1986.

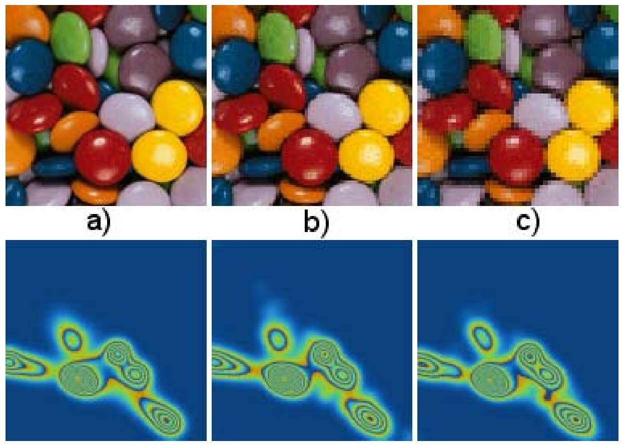

Fig. 2. Density estimation in the rg color space for the test image of size 1536×1536 (**a**), 96×96 (**b**) and 48×48 (**c**), (bilinear interpolation was used).

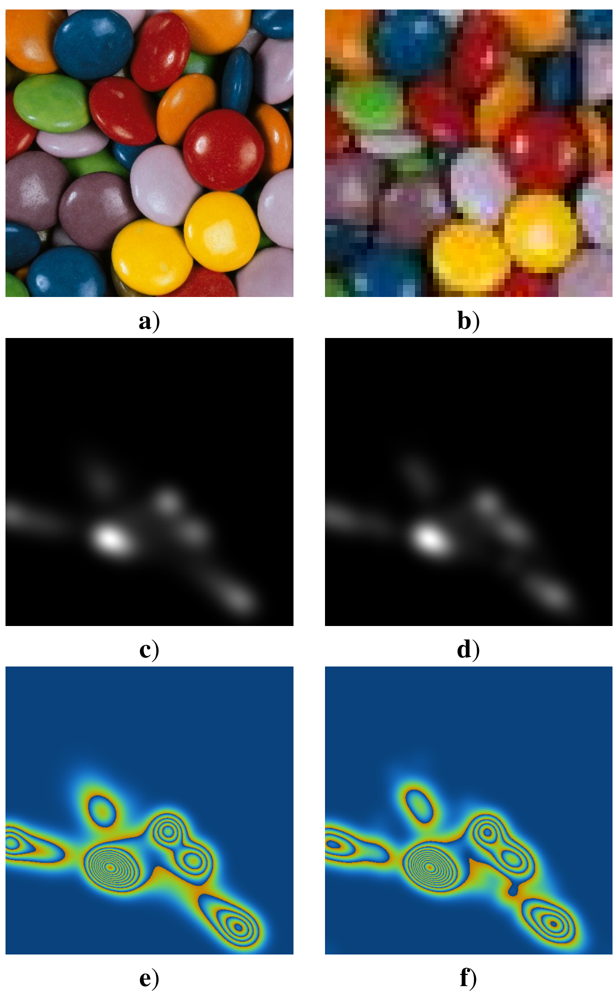

Fig. 3. Robustness of density estimation to image scaling and compression: **a**) test image of size 512x512, **b**) resized and compressed (JPEG) image of size 48x48, **c**), **d**) gray-scale and pseudo-color representation, **e**) and **f**).

Fig. 4. Results for the query for images similar to the Vincent van Gogh *Starry Night* painting of size (700 × 576).

Fig. 5. Representative results for the query for images similar to the thumbnail of size (122 x 100) of *Starry Night*.