Proceedings of the IEEE
International Conference on Automation and Logistics
Qingdao, China September 2008

New Similarity Measure for Illumination Invariant Content-Based Image Retrieval

Leila Sabeti, Q.M. Jonathan Wu
Department of Electrical and Computer Engineering
University Of Windsor, Ontario, Canada N9B 3P4
{sabeti, jwu}@uwindsor.ca,

Abstract

Similarity measure is used to study the similarity between patterns and forms the basis of content-based image retrieval systems. We have investigated existing similarity measures, and proposed a new similarity measure for illumination invariant content-based image retrieval that does not consider any prior knowledge about the camera or the illuminant. Normalized cumulative colour histogram is adopted in this paper for image feature modeling, while the new similarity measure compares the query and target images to search among large databases. Our algorithm is tested on the SFU database, and the experimental results prove the efficiency of the proposed technique during successful image retrieval.

1. Introduction

Content-Based Image Retrieval (CBIR), also known as Query By Image Content (QBIC) is the process of searching and retrieving desired images from large collections or databases like photographs and multimedia archives, retail catalogs, art collections, medical records, etc.

In order to study the similarity measures during image retrieval, image features or actual contents of images such as colour, texture or shape, are analyzed. Colour is the most popular feature as it can be processed quickly and efficiently, and is stable with respect to geometric variations of the object pattern. The degree of similarity between two images is usually measured by the distance between their contents. To find the highest similarity match, colour histograms of the query and target images are compared instead of the statistical comparison of actual image including the mean square error (MSE) or peak signal-to-noise ratio (PSNR) techniques, because as these measures do not always correspond to the human perception of similar objects [1]. An image's colour histogram is composed of bins—each counting the number of pixels that represent a particular colour for the whole image or an object in the image. Bins, forming a histogram are referred to as the lookup table (LUT).

Different histogram-based similarity measures are explained in the later sections of the paper. All these methods try to improve the CBIR's efficiency by selecting the best match; however, there are only a few that consider the challenge of object recognition under gradual or severe illumination variations. In this paper we address the problem of object recognition in CBIR systems under various lighting conditions by utilizing the cumulative colour histogram or spatial pixel information to model the object feature. We also introduce a new similarity measure that compares the histograms of the query and target images containing the same object under various brightness changes, and then searches for the best object match.

The organization of this paper is as follows: existing histogram-based similarity measures are described in section 2; our adopted feature modeling methods including the normalized cumulative colour histogram and spatial ranks are explained in section 3; section 4 clarifies the proposed similarity measure, and its success rate test results are provided in section 5; section 6 concludes the paper.

2. Histogram-based similarity measures

Various similarity measures have previously been proposed in the literature. In this section we briefly explain the specifications of the existing methods.

2.1 Similarity measures

Histogram-based similarity measures include histogram intersection, Manhattan, Euclidian and quadratic distances, center moment method, χ^2 statistical distance measure and Bhattacharyya distance measure [2].

Histogram intersection distance for two images $\mathbf{I}$ and $\mathbf{I}'$ is defined as:

$$d(H[\mathbf{I}],H[\mathbf{I}']) = \sum_{l=0}^{L-1} \min\left(\frac{H[\mathbf{I}](l)}{\sum_{l=0}^{L-1} H[\mathbf{I}](l)}, \frac{H[\mathbf{I}'](l)}{\sum_{l=0}^{L-1} H[\mathbf{I}'](l)}\right) \quad (1)$$

where L is the total number of bins. Two images are considered similar if the result of the intersection is close to 1, and considerably different if it is close to 0.

City-block or Manhattan, Euclidian and quadratic distances are respectively defined as:

$$d(H[\mathbf{I}],H[\mathbf{I}']) = \sum_{l=0}^{L-1} | H[\mathbf{I}](l) - H[\mathbf{I}'](l) | \quad (2)$$

$$d(H[\mathbf{I}],H[\mathbf{I'}])=\sqrt{\sum_{l=0}^{L-1}(H[\mathbf{I}](l)-H[\mathbf{I'}](l))^2} \quad (3)$$

and,

$$d(H[\mathbf{I}],H[\mathbf{I'}])= \sum_{l=0}^{L}\sum_{m=0}^{L}a_{lm}(H[\mathbf{I}](l)-H[\mathbf{I'}](l))(H[\mathbf{I}](m)-H[\mathbf{I'}](m)) \quad (4)$$

In the center moment method, the first three rank moments of an image histogram are as follows [2]:

$$M_1=\frac{1}{L}\sum_{l=0}^{L-1}H[\mathbf{I}](l)$$
$$M_2=\sqrt{\frac{1}{L}\sum_{l=0}^{L-1}(H[\mathbf{I}](l)-M_1)^2}$$
$$M_3=\sqrt[3]{\frac{1}{L}\sum_{l=0}^{L-1}(H[\mathbf{I}](l)-M_1)^3} \quad (5)$$

and the distance value between two images is:

$$d(H[\mathbf{I}],H[\mathbf{I'}])=\sqrt{\sum_{i=0}^{3}(M_i[\mathbf{I}]-M_i[\mathbf{I'}])^2} \quad (6)$$

The χ^2 statistical distance measure is also defined as

$$d(H[\mathbf{I}],H[\mathbf{I'}])=\sum_{i=0}^{L-1}\frac{(H[\mathbf{I}](l)-m(l))^2}{m(l)} \quad (7)$$

where,

$$m(l)=\frac{H[\mathbf{I}](l)+H[\mathbf{I'}](l)}{2},\quad l=0,\ldots,L\text{-}1 \quad (8)$$

The Bhattacharyya coefficient also considers discrete densities such as two colour histograms $H[\mathbf{I}]$ and $H[\mathbf{I'}]$, and is described as:

$$\rho(H[\mathbf{I}],H[\mathbf{I'}])=\sum_{l=0}^{L-1}\sqrt{H[\mathbf{I}](l).H[\mathbf{I'}](l)} \quad (9)$$

The larger the ρ is, the more similar the distributions are. For two identical histograms $\rho = 1$ is obtained, indicating a perfect match. The distance between two distributions is measured using:

$$d(H[\mathbf{I}],H[\mathbf{I'}])=\sqrt{1-\rho(H[\mathbf{I}],H[\mathbf{I'}])} \quad (10)$$

which is called the Bhattacharyya distance.

Based on the evaluation results of Bao and Guo [2], from the average retrieval efficiency, χ^2 statistical distance measure gives the best result and center moment and histogram intersection show the poorest performance. The Bhattacharyya and the center moment measure distances also show efficient results using cumulative histograms. The above equations, however, provide inaccurate, poor or incorrect results when colour levels shift as a result of brightness changes, especially when the variation is drastic due to a significant change in the colour vectors. Colour constancy approaches were therefore introduced to overcome the challenge of illumination variation in CBIR systems.

2.2. Colour constancy

To estimate the illuminant, colour constancy methods transform image features or colour vectors so that a true surface reflectance properties of the image is obtained. A number of approaches have been proposed to estimate the image illuminant including: Gray World [3], retinex [4], gamut mapping [5], Bayesian colour constancy [6], and neural network-based algorithms [7].

All these algorithms are based on the assumptions of the existence of camera characteristics, the illuminant properties, or the distribution of the colour values. These algorithms show poor results when no prior information about the acquisition conditions is available, and their assumptions about the camera and illumination are restrictive [8, 9].

Recently, Muselet et al. [8, 10, 11] introduced three methodologies based on the intersection between the histograms of ranks, histograms of fuzzy ranks and histograms of fuzzy spatial ranks. To find the similarity measure, colour histograms are adapted to the query and target images. Fuzzy spatial ranks method shows the best result and considers both the colour of image pixels and the spatial interaction between them. It also estimates the real value of colour responses to CCD sensors using fuzzy functions.

In this paper we focus on similarity measures, and compare our proposed method with the adapted histogram intersection of the recent work of Muselet et al. [8]. As [10,11] proposed, success rate results can be increased by using fuzzy techniques and improving the feature modeling using spatial ranks, however in this paper, we focus mostly on the comparison of similarity measure techniques.

3. Normalized cumulative colour histogram and spatial ranks

For a colour frame or image $\mathbf{I}$, three separate colour component matrices $\mathbf{I}_k=\{\mathbf{I}_R,\mathbf{I}_G,\mathbf{I}_B\}$ are obtained, where each pixel P is described with one colour value $c(P_k)$. The normalized cumulative colour histogram or rank measure [12] of the pixel P is expressed as:

$$r^k[\mathbf{I}](p)=\frac{\sum_{l=0}^{c^k(p)}H^k[\mathbf{I}](l)}{\sum_{l=0}^{L-1}H^k[\mathbf{I}](l)} \quad k=R,G,B \quad (11)$$

where L is the number of levels used to quantize the colour components (L is generally set to 256), and $H^k[\mathbf{I}](l)$ is the histogram count or number of pixels in the image that contain the R,G or B level or value of l. Finlayson et al. [12] showed that in two images of the same object, rank measures are equal—an interpretation based on the fact that the colour of a pixel in the first image can be deduced from the colour of the corresponding pixel in the second image by a nonlinear monotonic increasing function [12]. Therefore, if the colour level is higher in one region or pixel of the object than another region or pixel in the first frame, the colour level will be higher in the corresponding region or pixel in the second image. So if,

$$c^k(p1) > c^k(p2) \tag{12}$$

then,

$$f(c^k(p1)) > f(c^k(p2)) \tag{13}$$

and,

$$\therefore\ c^k(p1') > c^k(p2') \tag{14}$$

This means that normalized cumulative histogram values are coarsely preserved for an object even under brightness changes (Figure1). Later Muselet et al. [8] showed that the rank measures of two corresponding pixels in two different images are not, in fact, equal—only closest among all other rank measures. Even if the same image is captured under the same illumination, there will be shifts in rank measures resulting from the different responses of the CCD sensors or noise,

$$c^k(p_1') = f(c^k(p_1)) + \rho(p_1') \tag{15}$$

where $\rho(p_1')$ is the modification.

Same fact is true for the case of spatial colour ranks. The spatial rank $SR^i\,[\mathbf{I}_K]$ of the level i within the colour component image $\mathbf{I}_K$ is defined as the sum of the cells in the co-occurrence matrix $Co[\mathbf{I}_K]$ that represent the spatial interactions between pixels characterized by levels ranging from 0 to i:

$$SR^i\,[\mathbf{I}_K] = \sum_{u=0}^{i}\sum_{v=0}^{i} Co[\mathbf{I}_K](u,v) \qquad K = R,G,B \tag{16}$$

Where $i \in \{1,2,...,L\}$. The co-occurrence matrix $Co[\mathbf{I}_K]$ characterizes the local spatial interaction between pixel levels within each colour component of matrix $\mathbf{I}_k$, and $Co[\mathbf{I}_K](u,v)$ indicates the number of times in the image **I,** a pixel P'—whose level $c(P_K')$ is equal to v—is located in the 8-neighborhood of a pixel P whose level $c(P_K)$ is equal to u. The total number of co-occurrences normalizes this number so that the matrix does not depend on the number of pixels that represent the object. The relationships between levels of neighboring pixels within the three colour component images are represented by the three matrices, $Co[\mathbf{I}_R]$, $Co[\mathbf{I}_G]$ and $Co[\mathbf{I}_B]$. The spatial rank increases with respect to i and ranges from 0 to 1, since the matrices $Co[\mathbf{I}_k]$ are normalized by the total number of co-occurrences, and L is the total number of *RGB* levels in the image—usually 256.

4. The proposed similarity measure

Our experiments revealed that a drastic change in illumination will produce a shift in the colour values, while the rank measures and spatial rank measures will be preserved. We also found that while searching a database using CBIR, the rank measures of the query and target object have a higher similarity compared to other candidates, despite a severe brightness change. The proposed similarity measure considers this fact to compare colour ranks or spatial colour ranks.

In this method, the rank measures or spatial rank measures of the query and candidate image are quantized into N sections. We name these areas, boxes I_1 and I_2:

$$\zeta = 0, \frac{1}{N}, \frac{2}{N}, ..., 1 \qquad \text{where} \qquad N << L$$

Total number of *RGB* levels in the image—usually 256, is denoted by L. Note that the normalized cumulative colour histogram value or spatial rank measure of a pixel p for colour k in image $\mathbf{I}$ are between 0 and 1:

$$r^k[\mathbf{I}](p) \in [0,\,1] \qquad k = R,G,B$$

For each image, the rank measures closest to the quantized values are computed. So, for $n \in [0, N-1]$,

$$l_1' = \underset{l=0,...,L-1}{\arg\min} \left\| r^k[\mathbf{I}_1](l) - \zeta(n) \right\| \tag{17}$$

$$q^k[\mathbf{I}_1](n) = r^k[\mathbf{I}_1](l') \tag{18}$$

and,

$$l_2' = \underset{l=0,...,L-1}{\arg\min} \left\| r^k[\mathbf{I}_2](l) - \zeta(n) \right\| \tag{19}$$

$$q^k[\mathbf{I}_2](n) = r^k[\mathbf{I}_2](l') \tag{20}$$

The summation of all the rank measures at each segmented section is obtained using the following equations:

$$a^k[\mathbf{I}_1](n) = \sum_{\left\{l \mid r^k[\mathbf{I}_1](l) \in \left[q^k[\mathbf{I}_1](n),\, q^k[\mathbf{I}_1](n+1)\right[\right\}} r^k[\mathbf{I}_1](l) \tag{21}$$

$$a^k[\mathbf{I}_2](n) = \sum_{\left\{l \mid r^k[\mathbf{I}_2](l) \in \left[q^k[\mathbf{I}_2](n),\, q^k[\mathbf{I}_2](n+1)\right[\right\}} r^k[\mathbf{I}_2](l) \tag{22}$$

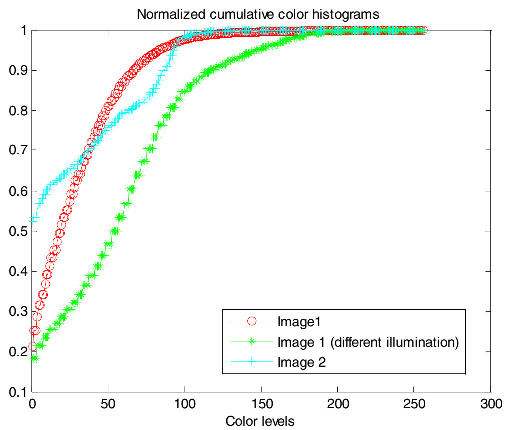
Finally, the computed sums are compared using Euclidian distance:

$$d(a[\mathbf{I}_1] - a[\mathbf{I}_2]) = \sum_k \sqrt{\sum_{n=0}^{N-1} (a^k[\mathbf{I}_1](n) - a^k[\mathbf{I}_2](n))^2} \tag{23}$$

for $k=R,\ G,\ B$

Image1 Image1 Image2

(a)



(b)

Figure 1. (a) Candidate images, Image 1, Image 1 with illumination change, and image 2; (b) Normalized cumulative colour histograms for the candidate images. It can be seen that the values are coarsely preserved for an object after illumination changes.

5. Experimental results

We have tested our proposed algorithm on the Simon Fraser University (SFU) database. This database utilized by [8, 10, 11, 13] contains three more objects, and can be found at http://www.cs.sfu.ca/~colour/data. It includes 20 different objects, shown in Figure 2, each under 11 different illuminants as shown in Figure 3. All of the 220 images were captured with a Sony DXC-930 3 CDD digital video camera.

For each iteration, the database of 20 objects was considered as target, and one of the 20 objects illuminated with one of the 10 remaining illuminants was selected as the query image. Therefore, to test the efficiency of the proposed technique using image retrieval, 10×11×20 or 2200 tests were performed to compute the success rate.

We first compared the proposed similarity measure with the technique called adapted histograms intersection proposed by Muselet et al. [8]. Both works utilize normalized cumulative colour histograms or histograms of ranks for image feature modeling. Spatial ranks method was then employed as an object feature integrated with our proposed similarity method and its result was also compared with the adapted histograms intersection technique used in [8]. Table 1 shows the quantitative comparison results of all four methods tested on SFU database for the application of image retrieval. As [8, 10, 11] proposed, success rates can be increased by using fuzzy techniques however, in this paper we have mostly focused on comparing similarity measure techniques.

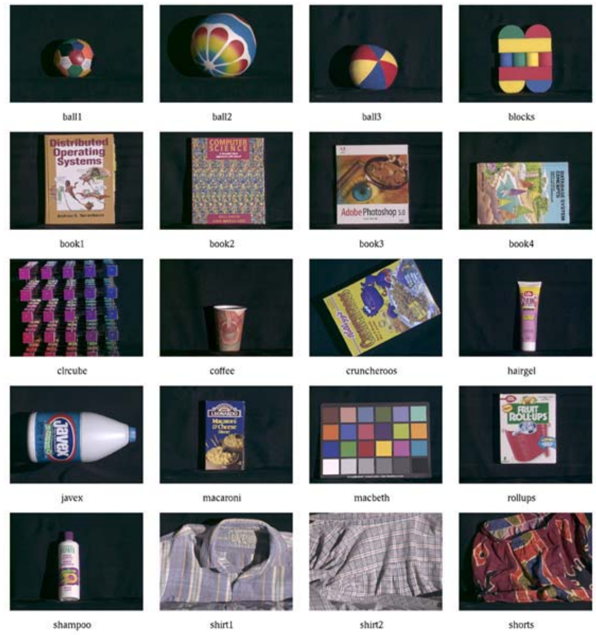


Figure 2. Samples of 20 objects in the SFU database.

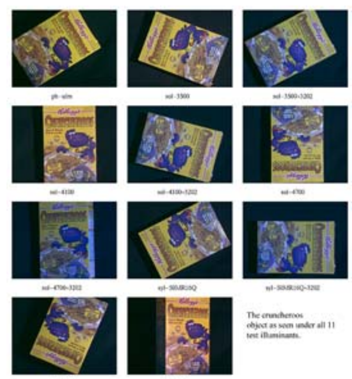


Figure 3. Objects under 11 different illuminants.

Table 1. Success rate comparisons for various methods, as applied to image retrieval.

Methods	**Success rates(%)**
Adapted histograms intersection using histograms of ranks	45.3636
Adapted histograms intersection using histograms of spatial ranks	47.0909
The proposed similarity measure using histograms of ranks	85.7727
The proposed similarity measure using histograms of spatial ranks	87.4545

6. Conclusions

The objective of this paper is to propose a new similarity measure to be applied to content-based image retrieval under random illumination variations. Most of the existing techniques either do not fully consider the brightness changes or need extra information about camera characteristics or the illuminant properties that result in very low success rates, when no prior information about the acquisition conditions is available. We have overcome this challenge by introducing a new and original similarity measure, and have tested its efficiency on a publicly available image database. The proposed scheme utilizes the normalized cumulative colour histogram or spatial pixel information to model image features. It adopts the fact that in a database containing different objects, rank measures of similar objects are the closest even after object translation or illumination variation.

7. References

[1] D. Van der Weken, M. Nachtegael, and E.E. Kerre, "Using Similarity Measures for Histogram Comparison", *Lecture Notes in Artificial Intelligence*, Springer, 2004, pp. 396-403.

[2] Q. Bao, and P. Guo, "Comparative Studies on Similarity Measures for Remote Sensing Image Retrieval", *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1, 2004, pp. 1112-1116.

[3] G. Buchsbaum, "A spatial processor model for object colour perception", *J. Franklin Inst*. 310, 1980, pp. 1–26.

[4] D.H. Brainard, and B.A. Wandell, "Analysis of the retinex theory of color vision", *J. Opt. Soc. America 3*, 1986 pp. 1651–1661.

[5] G.D. Finlayson, and S. Hordley, "Improving gamut mapping colour constancy", *IEEE Trans. Image Process,* 2000, pp. 1774–1783.

[6] D.H. Brainard, and W.T. Freeman, "Bayesian color constancy", *J. Opt. Soc. Am. A 14*, 1997, pp. 1393–1411.

[7] V. Cardei, "A neural network approach to colour constancy", *Ph.D. Thesis*, Simon Fraser University, January 2000.

[8] D. Muselet, L. Macaire, and J.G. Postaire, "Color Histograms Adapted to Query-Target Images for Object Recognition across Illumination Changes", *EURASIP Journal on Applied Signal Processing*, 2005, pp. 2164–2172.

[9] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods", *Pattern Recognition*, Elsevier Science, Volume 40, Issue 3, 2007, pp. 1106-1122.

[10] D. Muselet, and L. Macaire, "Fuzzy spatial ranks for object recognition across illumination changes", *IEEE International Conference on Multimedia and Expo ICME,* July 2006.

[11] D. Muselet and L. Macaire, "Histogram of Fuzzy Ranks for Object Recognition Across Illumination Changes", *Third European Conference on Color in Graphics*, Imaging and Vision, CGIV 2006, pp. 220-225.

[12] G. Finlayson, S. Hordley, G. Schaefer and G. Y. Tian, "Illuminant and device invariant colour using histogram equalization", Pattern Recognition, 38, 2005, pg. 179.

[13] G. Finlayson and G. Schaefer, "Colour indexing across devices and viewing conditions", *in Proc. 2nd International Workshop on Content-Based Multimedia Indexing, CBMI '01*, Brescia, Italy, September 2001, pp. 215–221.