Research Article *ISSN: 2319-507X*
Prajakta Thakare, IJPRET, 2013; Volume 1(8): 623-631 *IJPRET*

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

IDEAS, INFLUENCES AND PROMISING DIRECTIONS FOR IMAGE RETRIEVAL USING MULTIPLE FEATURES REPRESENTATIONS.

PROF. PRAJAKTA R. THAKARE, PROF. TANAJI D. KHADTARE

Department of Information Technology, Sinhgad Institute of Technology and Science, University of Pune, India

Accepted Date:

27/02/2013

Publish Date:

01/04/2013

Keywords

Content Based Image Retrieval,

Feature Extraction,

Color,

Texture,

Shape,

Retrieval Efficiency.

Corresponding Author

Prof. Prajakta R. Thakare

Abstract

Content-based image retrieval (CBIR), is any technology that in principle helps to organize digital picture archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR. It aims at retrieving the similar set of images from the database corresponding to the users query. To do so, a set of features need to be extracted from the images and stored in the database prior to accepting users query. Color, texture and shape information have been the primitive image descriptors in content based image retrieval systems. In this paper we are focusing on improving image retrieval performance by using low level features like color, texture and shape. The current state-of-the-art in CBIR holds enough promise and maturity to be useful for real-world applications if aggressive attempts are made

1. INTRODUCTION

Recent years have seen a rapid increase in the size of digital image collections. However, we cannot access or make use of the information unless it is organized so as to allow efficient browsing, searching, and retrieval. Image retrieval has been a very active research area since the 1970s, with the thrust from two major research communities, database management and computer vision. However, there exist two major difficulties, especially when the size of image collections is large. One is the vast amount of labor required in manual image annotation. The other difficulty, results from the rich content in the images and the subjectivity of human perception. That is, for the same image content different people may perceive it differently. The perception subjectivity and annotation impreciseness may cause unrecoverable mismatches in later retrieval processes. Due to emergence of large-scale image collections, manual annotation approach became more and more acute. To overcome these difficulties, content-based image retrieval was proposed. That is, instead of being manually annotated by text-based key words, images would be indexed by their own visual content, such as color, shape and texture. Since then, many techniques in this research direction have been developed and many image retrieval systems, both research and commercial, have been built. Regarding content-based image retrieval, we feel there is a need to survey what has been achieved in the past few years and what are the potential research directions which can lead to compelling applications. Since excellent surveys for text-based image retrieval paradigms already exist, in this paper we will devote my effort primarily to the content-based image retrieval paradigm. CBIR is considered as the process of retrieving desired images from huge databases based on extracted features from the image themselves without resorting to a keyword [4]. CBIR aims at searching image libraries for specific image features like colors and textures and querying is performed by comparing feature vectors of a search image with the feature vectors of all images in the database. The visual features are classified into low and high level features according to their complexity and the use of semantics. This paper considers low level features like Color, Texture and Shape. The remainder of this paper is organized as follows. In Section II, we review various visual features like Color, Shape, Texture, their corresponding representation and Image Retrieval techniques. Section III gives concluding remarks.

II. FEATURE EXTRACTION

This section describes various feature extraction techniques using Color, Texture and Shape.

2. Robust Color Feature

Color is the most widely used visual content for retrieving images because it does not depend on image size or orientation. Global Color Histogram (GCH) is the most traditional way of describing the color attribute of an image [4]. It is constructed by computing the normalized percentage of the color pixels in an arrange corresponding to each color element. An example of a true colored (RGB) image and the corresponding histograms of each component are displayed in Fig. 1.

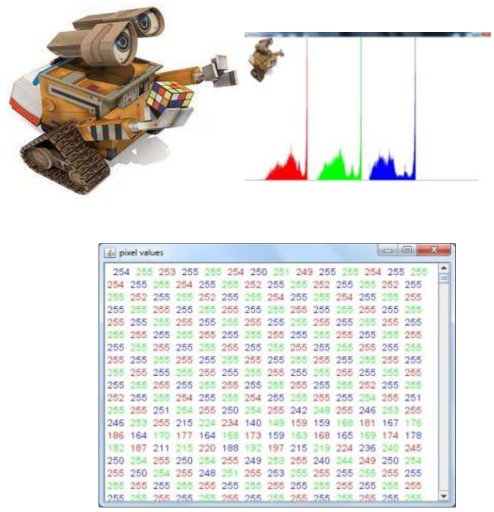

Fig.1 A colored image , corresponding histogram for red green and blue components , and the corresponding pixel values.

To construct the color feature vector for both the query image and all images in the database, the three-color components (R, G, and B) are identified and corresponding histograms of these components is computed. A GCH considers neither the color similarity across different bins nor the color dissimilarity in the same bins. Hence it is found to be sensitive to noisy interference such as illumination changes and quantization errors. To address these concerns, a new color histogram representation, called Fuzzy Color Histogram (FCH) is proposed[5], by considering the color similarity of each pixel's color associated to all the histogram bins through fuzzy-set membership function in comparison with the(GCH), which assigns each pixel into one of the bins only. FCH spreads each pixel's total membership value to all the histogram bins. Also, to reduce the computational complexity, Fuzzy C–Means (FCM) clustering algorithm can be used. Taking a color space containing n different color bins, the color histogram of image I containing N pixels is represented as H (I) = [h1, h2, …hn], where hi = Ni / N is the probability of a pixel in the image belonging to the i^{th} color bin, and Ni is the total number of pixels in the i^{th} color bin. According to the probability theory, h_i can be defined as

$$h_i = \Sigma P_{i/j} P_j = (1/N) \Sigma P_{i/j}$$

where P_j is the probability of a pixel chosen from image I being the j^{th} pixel, which is 1/N and $P_{i/j}$ is the conditional probability of the chosen j^{th} pixel belonging to the i^{th} color bin. FCH, on the other hand considers each of the N pixels in image I, related to all the n color bins via Fuzzy set membership function such that the degree of "belongingness" of the j^{th} pixel to the i^{th} color bin is found by distributing the membership value of the j^{th} pixel, μ_{ij}, to the i^{th} color bin. Thus the Fuzzy color Histogram (FCH) of image I can be expressed as F(I) = [f1, f2… fn], where

$$f_i = \Sigma \mu_{ij} P_j = (1/N) \Sigma \mu_{ij}$$

Thus when compared to GCH, FCH considers not only the similarity of different colors from different bins but also the dissimilarity of those colors from the same bin. Lu et al[1] proposed perceptually weighted histogram or PWH. CIEL*u*v space is chosen because it is an uniform color space in terms of color distance. In order to use L*u*v* space, color values are first converted from RGB space into CIEXYZ space with a linear transform and then from CIEXYZ space into L*u*v* color space using the following transform:

$$
\begin{array}{lll}
L^* = 116 \times \sqrt[3]{Y/Y_n} - 16 & Y/Y_n > 0.008856 & u' = 4X/(X+15Y+3Z) \\
L^* - 903.3 \times (Y/Y_n) & Y/Y_n \le 0.008856 \quad \textit{where} & v' = 9Y/(X+15Y+3Z) \\
u^* = 13 \times L^*(u' - u'_n) & & u'_n = 4X_n/(X_n+15Y_n+3Z_n) \\
u^* - 13 \times L^*(v' - v'_n) & & v'_n = 9Y_n/(X_n+15Y_n+3Z_n)
\end{array}
$$

and (X_n, Y_n, Z_n) is the reference white in XYZ space.

In the L*u*v* space, representative colors are used instead of quantizing each color channel by a constant step. The number of representative colors is given by the combinations (512) of the three components L,u,v components in L*u*v* space[1]. These representative colors are uniformly distributed in L*u*v* space. In contrast to the conventional histogram building which assigns the color of each pixel to a single color bin, the PWH assigns the color of each pixel to 10 neighboring color bins based on the following weight:

$$w_i = \frac{1/d_i}{1/d_1 + 1/d_2 + \ldots + 1/d_{10}}$$

$$\text{where } d_i = \sqrt{(L_0 - L_i)^2 + (u_0 - u_i)^2 + (v_0 - v_i)^2}$$

and (L_0, u_0, v_0) is the color of the pixel to be assigned, (L_i, u_i, v_i) is the color of bin i. The use of PWH overcomes the drawback of conventional histogram methods which would in many situations assign a pixel color to a bin of a quite different color and having to assign two quite different colors to a same color bin. As the result, PWH is much more

accurate in representing the image than conventional histograms.

3. Texture

Two images with different content can usually be distinguished by their texture features even when the images share similar colors. The term 'texture' is used to specify the roughness or coarseness of object surface. Frequency domain techniques using Gabor filters or Gabor wavelets[1] have been particularly popular for CBIR. Basically, Gabor filters are a group of wavelets, with each wavelet capturing energy at a specific frequency and a specific direction. Expanding a signal using this basis provides a localized frequency description, therefore capturing local features/energy of the signal. Texture features can then be extracted from this group of energy distributions. For a given image

I(x, y) with size P×Q, its discrete Gabor wavelet transform is given by a convolution:

$$G_{mn}(x,y) = \sum_{s=0}^{K}\sum_{t=0}^{K} I(x-s, y-t) g^*_{mn}(s,t)$$

where, K is the filter mask size, and g^*_{mn} is the complex conjugate of g_{mn} which is a class of self-similar functions generated from dilation and rotation of the following mother wavelet

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})]\cdot\exp(j2\pi Wx)$$

where W is called the modulation frequency.

After applying Gabor filters on the image with different orientation at different scale, an array of magnitudes is obtained. These magnitudes represent the energy content at different scale and orientation of the image.

The mean μ_{mn} and standard deviation $\acute{o}_{mn}$ of the magnitude of the transformed coefficients are used to represent the homogenous texture feature of the region. A feature vector f (texture representation) is created using μ_{mn} and $\acute{o}_{mn}$ as the feature components. Five scales and six orientations are used in common implementation and the Gabor texture feature vector is thus given by:

f = (μ_{00} , $\acute{o}_{00}$,μ_{01} , $\acute{o}_{01}$, …, μ_{45}, $\acute{o}_{45}$). The similarity between two texture patterns is measured by the Euclidean distance between their Gabor feature vectors. To make the above extracted texture features robust to image rotation, a simple circular shift on the feature map is used [6]. Texture co-occurrence matrix can also be used for describing the texture of an image[3]. Usually a small patch of finite area of an image is required to feel or measure local texture value. So in order to compute the texture co-occurrence matrix, the intensity image is divided into blocks of size 2×2 . Then grey level of the block is converted to binary by thresholding at the average intensity. By arranging this pattern in raster order a binary string is formed. It is considered as the gray code and corresponding decimal equivalent is its texture value. Thus, by the virtue of gray code, blocks with similar texture are expected to have closer values. The problem with this approach is that smooth intensity block and coarse textured block may produce same binary pattern and hence same texture value.

Hence Wavelet' transform can be used to characterize textures using statistical properties of the gray levels of the pointed pixels comprising a surface image [4]. The wavelet transform is a tool that cuts up data or functions or operators into different frequency components and then studies each component with a resolution matched to its scale. There are different types of wavelet families whose qualities vary according to several criteria. Daubechies is one of the brightest stars in the world of wavelet research. Daubechies family includes the Haar wavelet, written as 'DBI, Formulas (1) and (2) illustrate the mother wavelets for the Haar wavelet:

$$\phi := x \rightarrow \left\{ \begin{array}{l} 0 \le x \textbf{ and } x < 1 \\ otherwise \end{array} \right. \quad (1)$$

$$\psi := x \rightarrow \begin{cases} 0 \le x \textbf{ and } x < \frac{1}{2} \\ \frac{1}{2} \le x \textbf{ and } x < 1 \\ otherwise \end{cases} \quad (2)$$

Fig. 2 The Haar (DB1) wavelet ψ and its scale function Φ

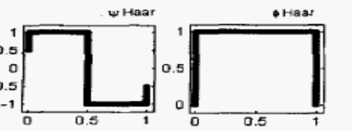

Fig.2 The Haar wavelet and its scale function Φ

Where $\emptyset$ is called the scale of the **Haar** wavelet and **Ψ,** is the actual wavelet in (Fig. 2) [4]. Fig. 3 shows an example of different Daubechies wavelets (DB2, DB4, and DBS**).** A Daubechies wavelet representation of a function is a linear combination of the wavelet function elements

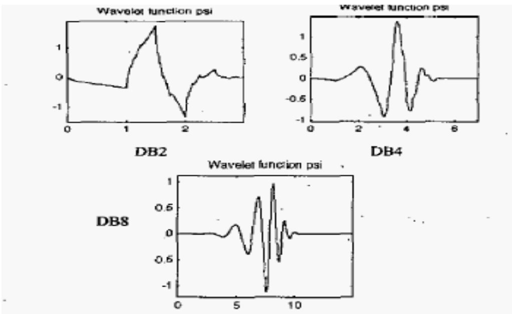

Fig. 3 The mother wavelets of Daubechies wavelets DB2, DB4, and DB8.

4. *Shape*

Compared with color and texture features shape feature are usually described after images have been segmented in to regions or objects. Shape does not refer to the shape of

an image but to the shape of a particular region that is being sought out. Wang Xiaoling and Xie Kanglin proposed minimized sum statistical direction code (MSSDC) to represent the shape of object [2]. It has the advantage of being invariant to position and rotation of object and is proportional to image scaling. To compare the query and the database image shape representation, an entropy based similarity measure is used in which first the entropy of the MSSDC is computed and then the similarity S(Q, I) between the query and the database image is defined as

$$S(Q, I) = \Sigma [(E_{Qi} - E_{Ii})^2]^{1/2}$$

where E_Q and E_I are the Entropy of the statistical directional code (SDC) of the query and database image respectively. We can also derive shape information from phase congruency. Phase congruency is a dimensionless quantity that is invariant to changes in image brightness or contrast; hence, it provides an absolute measure of the significance of feature points, thus allowing the use of universal threshold values that can be applied over wide classes of images. Congruency of phase at any angle produces a clearly perceived feature[5].

$$s(x) = \sum_{0}^{n} \frac{1}{(2n+1)} sin[(2n+1)x + \phi]$$

where Φ, the offset at which congruence of phase occurs, is varied from 0 to ∏/2. Phase congruency function in terms of the Fourier series expansion of a image at some location *x* is defined as

$$PC(x) = \{\max_{\Phi(x) \in [0, 2\pi]} \Sigma A_n \cos(\varphi_n(x) - \varphi'(x))\}/\Sigma A_n$$

where A_n represents the amplitude of the *n*th Fourier component, and $\Phi n(x)$ represents the local phase of the Fourier component at position *x*. The value of $\Phi(x)$ that maximizes this equation is the amplitude weighted mean local phase angle of all the Fourier terms at the point being considered. Values of phase congruency vary from a maximum of 1 (indicating a very significant feature) down to 0 (indicating no significance). This allows one to specify a threshold to pick out features before an image is seen.

CONCLUSION

In this paper, we have reviewed a CBIR system for images containing multiple objects using multiple features like Color, Shape and Texture along with past and current achievements in visual feature extraction. Successful Image Retrieval System requires seamless integration of multiple research community's efforts. We have presented a comprehensive survey highlighting current progress, emerging directions, the spawning of new fields, and methods for evaluation relevant to the young and exciting field of image retrieval and open research issues are identified. For better Retrieval efficiency features can be combined and novel method can be designed. We will like to deal with this issue in near future.

REFERENCES:

1. Dengsheng Zhang, "Improving Image Retrieval Performance by Using both Color and Texture Features ",Proceedings of Third international conference on Image and Graphics, 2004

2. Aradhana Katare, Suman K Mitra, Asim Banerjee, "Content based Image Retrieval Systems for Multi Object Images using Combined Features", Proceedings of international conference on Computing: Theory and applications,2007.

3. Sanjoy Kumar Saha, Amit Kumar Das, Bhabatosh Chanda, "CBIR using Perception based Texture and Colour Measures", Proceedings of international conference on Pattern Recognition, 2004.

4. Mohamed A. Tahoun, Khaled A. Nagag, Taha, El-Arie, Mohammed A-Megeed,"A Robust Content-Based Image Retrieval System Using Multiple Features Representations", Proceedings of IEEE international conference, 2005.

5. K. Satya Sai Prakash, RMD. Sundaram "Combining Novel features for Content Based Image Retrieval", 2005.

6. D.S. Zhang and G.Lu,"CBIR using Gabor Texture Features", in Proceedings of First IEEE Pacific-Rim Conference on Multimedia,2000.