ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Similarity of Image Multiple Feature Extraction and Retrieval Perspectives

S.Kousalya[1], Dr.Antony Selvadoss Thananmani[2]

Research scholar, Computer science (Aided), NGM College, Coimbatore, India [1]

Associate professor & Head, Computer science (Aided), NGM College, Coimbatore, India [2]

Abstract: Retrieval of images based on ocular qualities such as color, texture and shape have proven to essential its own set of limitations under different conditions. The other area in the Image mining system is the Content-Based Image Retrieval (CBIR).The color feature is one of the most widely used visual features in image retrieval using coherence vectors or color correlograms.A set of shape features from the contour image and retrieve the images by using Region-based shape descriptors. Gabor functions for texture feature extraction from the given image queries and with image to remove the effects of sensor noise and gray level deformation. The Retrieval system is used to find the similarity between a query image and database images.

Keywords: Content Based Image Retrieval; Coherence vectors; Gabor filters; Region based shape descriptors

I. INTRODUCTION

In image mining applications roughly used is the process of retrieving preferred images from a large collection on the origin of features that can be automatically extracted from the images themselves[1][2]. CBIR (Content Based Image Retrieval) have received intensive courtesy in the literature of image information retrieval, and consequently an expansive range of techniques has been implied. The image algorithms used in these systems are commonly divided into three tasks: extraction- selection, and classification. Each image is described by its visual features (color,shape,texture)[4].A similarity measure is used to find the similarity between a query image and database images .The feature extraction is most critical because the particular features made available for discrimination directly influence the efficacy of the classification task. The end result of the extraction task is a set of features, commonly called a feature vector, which constitutes a representation of the image. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps :

Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.

Similarity measure: The second step Measure the difference between images for determining the significance between images.

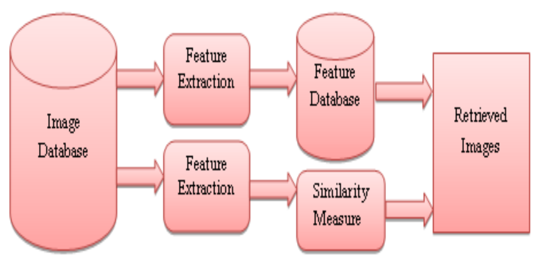

Figure 1: Extraction and Retrieval System Process

Content based indexing Low-level visual features like color, shape, texture are being used for indexing and retrieving images. Content Based Image Retrieval (CBIR) refers to a technique which uses visual contents to search an image from large scale image database according to users' interests &based on automatically-derived image features[4].

II. RELATED WORK

Ryszard S. Chora´s[1] main contributions of this work are the identification of the problems existing in CBIR and Biometrics systems describing image content and image feature extraction. A possible approach to mapping image content on to low-level features. Swati V. Sakhare &Vrushali G. Nasre[2] aims this application performs a simple color-based search in an image database for an input query image, using color, texture and shape to give the images which are similar to the input image as the output. The number of search results may vary depending on the

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

number of similar images in the database. M. Narayana[3]The retrieval performance of the proposed method is system architecture for the Content Based Image Retrieval by gray level Co-occurrence Matrix (GLCM) derived four image features.

III. IMAGE CONTENT DESCRIPTORS

A. Color

The important features that construct feasible the establishing of images by folks is color. Color is a property that depends on the contemplation of light to the eye and the processing of that information by the brain. Usually colors are defined in three dimensional color spaces. These could either be ***RGB*** (Red, Green, and Blue), ***HSV*** (Hue, Saturation, and Value) or ***HSB*** (Hue, Saturation, and Brightness). The last two are dependent on the human perception of hue, saturation, and brightness. A color image has three values per pixel and they measure the intensity and chrominance of light. The actual information stored in the digital image data is the brightness information in each spectral band. Quantization in terms of color histograms refers to the process of reducing the number of bins by taking colors that are very similar to each other and putting them in the same bin. Default, the maximum number of bins one can obtain using the histogram function. For the purpose of saving time when trying to compare color histograms, one can quantize the number of bins. Obviously quantization reduces the information regarding the content of images but it is a trade-off between processing time and quality.

In content-based image retrieval, images are automatically indexed by generating a feature vector (stored as an index in feature databases) describing the content of the image. The similarity of the feature vectors of the query and database images is measured to retrieve the image.

Let {F(x, y); x = 1, 2, . . . ,X, y = 1, 2, . . . , Y}be atwo-dimensional image pixel array. For color images F(x, y) denotes the color value at pixel (x, y)

$$F(x,y) = \{FR(x,y), FG(x,y), FB(x,y)\}$$

For black and white images ,F(x, y)denotes the grayscale intensity value of pixel (x, y).The problem of retrieval is following: For a query image Q,We find image T from the image database, such that distance between corresponding feature vectors is less than specified threshold,

$$D(Feature(Q), Feature(T) \leq t$$

A color model is specified in terms of3-D coordinate system and a subspace within that system where each color is represented by a single point. The more commonly used color models are *RGB* (red, green, blue),*HSV* (hue, saturation, value) and Y,Cb,Cr (luminance and chrominance). Thus the color content is characterized by3-channels from some color model. One representation ofcolor content of the image is by using color histogram. Statistically, it denotes the joint probability of the intensities of the three color channels.Color is perceived by humans as a combination of three color stimuli: Red, Green, Blue, which forms a color space

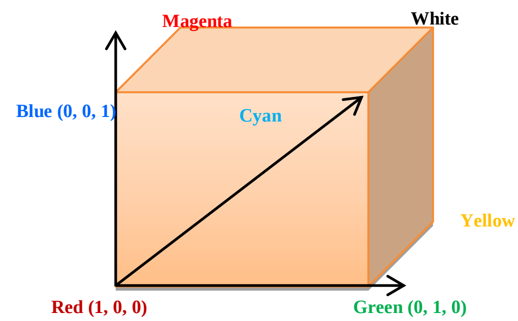

Figure 2: Color Space

Color Histogram

The color histogram is easy to evaluateand vigorousin illustratingboth the global and the local distribution of colors in an image. The color histogram extraction algorithm involves three steps: partition of the color space into cells, association of each cell to a histogram bin, andcounting of the number of image pixels of each cell and storing this count in the corresponding histogram bin. This descriptor is invariant to renovationand rotation. The similarity between two color histograms can be performed by computing the L1, L2, or weighted Euclidean distances.

Color Moments

Color moments have been successfully used in many retrieval systems (like QBIC), especially when the image contains just the object. The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images. Mathematically, the first three moments are defined as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij}$$

$$\sigma_i = (\frac{1}{N}\sum_{j=1}^{N}(f_{ij} - \mu_i)^2)^{\frac{1}{2}}$$

$$s_i = (\tfrac{1}{N}\sum_{j=1}^{N}(f_{ij} - \mu_i)^3)^{\frac{1}{3}}$$

Where f_{ij} is the value of the *i*-thcolor component of the image pixel j, and N is the number of pixels in the image. Usually the color moment performs better if it is defined by both the L*u*v* and L*a*b* color spaces as opposed to solely by the HSV space Using the additional third-order moment improves the overall retrieval performance compared to using only the first and second order moments.

B. Shape

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Many content-based image retrieval Systems use shape features of object or region. Shape features are usually described after images have been segmented into regions or objects as compared with color and texture features. The most frequently used methods for shape description can be boundary-based (rectilinear shapes, polygonal approximation, finite element models and Fourier-based shape descriptors) or region-based (statistical moments). A good shape representation feature for an object should be invariant to translation, rotation and scaling.

Shape descriptor

Features calculated from objects contour:circularity, aspect ratio, discontinuity angle irregularity,length irregularity, complexity, right-angleness,sharpness, directedness. Those are translation, rotation(except angle), and scale invariant shape descriptors. It is possible to extract image contours from the detected edges. From the object contour the shape information is derived. We extract and store a set of shape features from the contour image and for each individual contour. These features are

Circularity $cir = \frac{4pA}{p^2}$

Aspect ratio $ar = \frac{p_1 + p_2}{c}$

Discontinuity Angle irresularity

dar = $\sqrt{\frac{\sum|\theta_i - \theta_{i+1}|}{2\pi(n-2)}}$

Length Irregularity $lir = \frac{\sum|L_i - L_{i+1}|}{K}$

Right-Angleness$ra = \frac{r}{n}$

Sharpness $sh = \sum \frac{\max(0, 1 - \frac{2|\theta - \pi|}{\pi}}{n}$

Directedness $dir = \frac{M}{\sum p_i}$

C. Texture

Texture is the property of an image, characterized by the existence of basic primitives whose spatial distribution creates some visual patterns defined in terms of granularity, directionality, and repetitiveness. Generally, texture representation methods are classified into two main categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primaries and their rules. Statistical methods, including Tamura feature, shift-invariant principal component analysis (SPCA), Fourier power spectra, Wold decomposition, co-occurrence matrices, Markov random field, fractal modeland multi-resolution filtering techniques such as Gabor and wavelet transform, define texture by the statistical distribution of the image intensity. The co-occurrence matrix C(i, j) counts the co-occurrence of pixels with gray values i and j at a given distance d. The distance d is defined in polar coordinates (d,θ), with discrete length and orientation. In practice, θ takes the values0°; 45°; 90°; 135°; 180°; 225°; 270°; and 315°. The co-occurrence matrix C(i, j) can now be defined as follows:

$$c(i,j) = card \left\{ \begin{matrix} ((x_1, y_1), (x_2, y_2)) \in (XY) \times (XY) \\ for\ f(x_1, y_1) = i, f(x_2, y_2 = \\ j\ (x_2, y_2) \\ = (x_1, y_1) + (d\ cos\,\theta, d \sin\theta); \end{matrix} \right\}$$

Gabor Filter:

Gabor filters there has been increased interesting deploying Gabor filters in various computer vision applications and to texture analysis and image retrieval. The general functionality of the 2D Gabor filter family can be represented as a Gaussian function modulated by a complex sinusoidal signal. It can be normalize an image to remove the effects of sensor noise and gray level deformation. The two-dimensional Gabor filter is defined as:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right]$$

Where, σ_x and σ_y are the standard deviations of the Gaussian envelopes along the x and y direction.

IV. CONCLUSION

In the retrieval system and multiple feature extraction has been tremendous growth in the quality (resolution and color depth), nature (dimensionality) and throughput (rate of enervation) of the images acquired and this tendency of rising growth is likely to endure. The future of CBIR will depend on the progress made in each aspect of image retrieval, and the extent to which the individual user benefits by it.

REFERENCES

1. Ryszard S. Chora´s, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems", International Journal of Biology and Biomedical Engineering, Issue 1, Vol. 1, 2007
2. Swati V. Sakhare&Vrushali G. Nasre,"Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture",International Journal of Computer Science & Informatics, Volume-I, Issue-II, 2011
3. Nishchol Mishra, Dr. Sanjay Silakari, "Image Mining in the Context of Content Based Image Retrieval:A Perspective", International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.
4. T. N. Manjunath, R. S. Hegadi, and G. K. Ravikumar,"A Survey on Multimedia Data Mining and Its Relevance

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Today", Journal of Computer Science, vol.10, no. 11, 2010, pp. 165-170.
5. E. Saber, A.M. Tekalp,"Integration of color, edge and texture features for automatic region-based image annotation and retrieval," Electronic Imaging, 7, pp. 684–700, 1998.
6. Dr.Fuhui Long, Dr.Hongjiang Zhang and Prof. David Dagan Feng, "Fundamentals Of Content-based image retrieval"
7. ArtiKhaparde, B.L.Deekshatulu, M.Madhavilath,ZakiraFarheen, SandhyaKumari V, "Content Based Image Retrieval Using Independent Component Analysis", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.4, April 2008.
8. S.A.Pine, R C Patil, "Content Based Image Retrieval Using Wavelet Transform", Thadomal Shahani Engineering College, Bandra (East), Mumbai-51.
9. Dr.H.B.Kekre, Ms.SwapnaBorde," Content Based Image Retrieval," National Conference on Applications of Digital Signal Processing, NCDSP-2007, January 19-20, 2007
10. S.kousalya, Dr.Antony Selvadoss Thanamani, "Image Color Extraction and retrieval using classification Techniques" International Journal of Computer Science and Mobile Computing, Vol. 2, Issue. 8, August 2013, pg.169 – 173