The Journal of Applied Sciences Research. 2(2): 97-96, 2015

The Journal of Applied Sciences Research

Journal homepage: http://www.journals.wsrpublishing.com/index.php/tjasr

Online ISSN: 2383-2215 Print ISSN: 2423-4400

Original Article

Visual Feature Description Techniques for Content Based Image Retrieval

Latika Pinjarkar[1,*], Manisha Sharma[2] and Smita Selot[3]

[1] Associate Professor, Department of IT, Shri Shankaracharya Technical campus, Bhilai (C.G.), India
[2] Professor and Head, Department of ETC, Bhilai Institute of Technology, Bhilai (C.G.), India
[3] Professor and Head, Department of MCA, Shri Shankaracharya Technical campus, Bhilai (C.G.), India

ARTICLE INFO

Corresponding Author:
Latika Pinjarkar
latikabhorkar@gmail.com

How to Cite this Article:
Pinjarkar, L., Sharma, M., and Selot, S. (2015). Visual Feature Description Techniques for Content Based Image Retrieval. *The Journal of Applied Sciences Research*, 2(2): 97-96.

Article History:
Received: 18 August 2015
Revised: 11 September 2015
Accepted: 12 September 2015

ABSTRACT

Feature extraction technique is used in image processing to represent the image in its compact and unique form of single values for the purpose of content-based image retrieval (CBIR). The need of CBIR is because of the enormous increase in image database sizes, as well as its vast deployment in various applications. In CBIR systems, image processing techniques are used to extract visual features such as color, texture and shape from images. Images are represented as a vector of these extracted visual features. The images are retrieved on the basis of these visual feature vectors from the database. In this paper the conventional feature extraction techniques used in CBIR for visual feature description are discussed.

Keywords: Content Based Image Retrieval (CBIR), Color feature Extraction, Texture feature Extraction, Shape feature Extraction.

INTRODUCTION

CBIR or Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape. The features of the images in the database are extracted. These features are compared with the features of the query image for similarity computation.

The main steps in CBIR are:

Feature Extraction – Features are extracted from the images. The definitions of features are usually pre-defined, such as color, texture and shape. These features are usually stored in the form of real-valued multi-dimensional vectors.

Indexing – The image database may then organize the extracted features by using an indexing structure for retrieval.

Retrieval – Content-based retrieval can be performed on the indexing structure efficiently and effectively.

The rest of the paper is organized as follows:

The image processing technique for feature extraction are discussed in section II. Section III describes Color feature extraction. Section IV describes about Texture feature extraction. Section V discusses about shape feature extraction. The techniques for optimization and classification are discussed in section VI and VII respectively. Section VIII concludes the paper.

Feature Extraction

The visual content of an image is analyzed in terms of low-level features extracted from the image. These primarily constitute color, Shape and texture features (Long *et al.,* 2003). Various image processing techniques are used for extraction of these features. For Color Feature extraction, Color Histrogram, Color Moments and Color Correlogram are used. For Texture feature Extraction Gabor wavelet and Haar Wavelet process can be implemented. And for Shape feature Extraction, Fourier Descriptor, Circularity features are used. The extracted features are then optimized by optimization technique, where features are optimized and approximated to relevant features. These features are finally classified with similarity computation so as to retrieve the relevant images from the database.

Color Feature Extraction

Computing distance measures based on color similarity is achieved by computing a color histogram for each image that identifies the proportion of pixels within an image holding specific values. Examining images based on the colors they contain is one of the most widely used techniques because it can be completed without regard to image size or orientation. However, research has also attempted to segment color proportion by region and by spatial relationship among several color regions (Bimbo *et al.,* 1998).

Color feature extraction involves analyzing the absolute color value of each pixel. Color is generally represented by the color distribution of the image. Color distribution is a statistical feature and techniques such as color histogram, color moments and color correlogram can be implemented for color description (Long *et al*., 2003).

Color Histogram

In image retrieval systems color histogram is the most commonly used feature. The main reason is that it is independent of image size and orientation. Also it is one of the most straight-forward features utilized by humans for visual recognition and discrimination. Statistically, it denotes the joint probability of the intensities of the three color channels. Once the image is segmented, from each region the color histogram is extracted. The major statistical data that are extracted are histogram mean, standard deviation, and median for each color channel i.e. Red, Green, and Blue. So totally 3×3=9 features per segment are obtained.

The color histogram for color feature and wavelet representation for texture and location information of an image this reduces the processing time for retrieval of an image with more promising representatives. The extraction of color features from digital images depends on an understanding of the theory of color and the representation of color in digital images. Color spaces are an important component for relating color to its representation in digital form. The transformations between different color spaces and the quantization of color information are primary determinants of a given feature extraction method. Color is usually represented by color histogram, color correlogram, color coherence vector and color moment, under certain a color space. The color histogram feature has been used by many researchers for image retrieval. A color histogram is a vector, where each element represents the number of pixels falling in a bin, of an image. The color histogram has been used as one of the feature extraction attributes with the advantage like robustness with respect to geometric changes of the objects in the image. However the color histogram may fail when the texture feature is dominant in an image.

Color Moments

A color retrieval method based on the primitives of color moments is proposed. After dividing an image into several blocks, the color moments of all blocks are extracted and clustered into several classes based on a fast non-iterative clustering algorithm. The mean vector of each class is considered as a primitive of the image and all primitives are used as feature vectors. The first order (mean), the second order (variance) and the third order (skewness) color moments are used for effective representing color distribution of images (Long *et al*., 2003).

The first three moments are defined as:

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

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

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

Where f_{ij} is the value of the i^{th} color component of the image pixel j, and N is the number of pixels in the image.

Color Correlogram

The color correlogram is neither an image partitioning method nor a histogram refinement method. Unlike purely local properties, such as pixel position, gradient direction, or purely global properties, such as color distribution, correlograms take into account the local color spatial correlation as well as the global distribution of this spatial correlation. While any scheme that is based on purely local properties is likely to be sensitive to large appearance changes, correlograms are more stable to these changes; while any scheme that is based on purely global properties is susceptible to false positive matches, correlograms prove to be effective for content-based image retrieval from a large image database A color correlogram expresses how the spatial correlation of pairs of colors changes with distance.

A color Correlogram is a table indexed by color pairs, where the k-th entry for (i, j) specifies the probability of finding a pixel of color j at a distance k from a pixel of color i in the image. Let I represent the entire set of image pixels and Ic (i) represent the set of pixels whose colors are c (i). Then, the color Correlogram is defined as:

$$\gamma_{i,j}^{k} = \Pr_{p1 \in I_{c(i)},\, p2 \in I} \left[p2 \in I_{c(j)}, \, |p1 - p2| = k \right]$$

Where i,j$\in \{1,2,\ldots.N\}, k \in \{1,2,\ldots.d\}, and\ |p1-p2|$ is the distance between pixels p1 and p2.

Texture Features Extraction

Texture measures look for visual patterns in images and how they are spatially defined. Textures are represented by texels which are then placed into a number of sets, depending on how many textures are detected in the image. These sets not only define the texture, but also where in the image the texture is located (Ma and Manjunath, 1996).

The identification of specific textures in an image is achieved primarily by modeling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality may be estimated. The problem is in identifying patterns of co-pixel variation and associating them with particular classes of textures such as silky, or rough. For Texture feature extraction Gabor wavelet transform and Haar wavelet transform can be implemented.

Gabor Wavelet

Gabor wavelet is a classic method for multi-channel, multi resolution analysis that represents image variations at different scales (Lee, 1996). Gabor filters are a group of wavelets obtained from the appropriate dilation and rotation of Gabor function: a Gaussian modulated sinusoid. By capturing image details at specific scales and specific orientations, Gabor filters present a good similarity with the receptive fields on the cells in the primary visual cortex of the human brain.

Gabor wavelet transform provides a flexible method for designing efficient algorithms to capture more orientation and scale information. Many researches stated that Gabor wavelet transform represents one of the efficient techniques for image texture retrieval yielding good results in content - based image retrieval applications due to many reasons. Being well suited for image signal expression and representation in both space and frequency domains. Presenting high similarity with human visual system as stated above. Offering the capacity for edge and straight line detection with variable orientations and scales not being sensitive to lighting conditions of the image

Haar Wavelet

Haar wavelets, are fastest to compute and simplest to implement. In addition, user queries tend to have large constant-colored regions, which are well represented by this basis. One drawback of the Haar basis for lossy compression is that it tends to produce blocky image artifacts for high compression rates.

Haar wavelets do not have sufficiently sharp transition and hence are not able to separate different frequency bands appropriately. Daubechies' wavelets, on the other hand, have better frequency resolution properties because of their longer filter lengths.

Shape Features Extraction

Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out (Costa *et al*., 2001). Shapes will often be determined first applying segmentation or edge detection to an image. Other methods use shape filters to identify given shapes of an image. Shape descriptors may also need to be invariant to translation, rotation, and scale. Shape features of objects or regions have been used in many content-based image retrieval systems. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. For retrieval of shape feature Fourier Descriptor and circularity features can be extracted.

Fourier Descriptors

Fourier descriptors describe the shape of an object with the Fourier transform of its boundary (Bartolini *et al.*, 2005). Again, consider the contour of a 2D object as a closed sequence of successive boundary pixels (xs, ys), where 0 ≤ s ≤ N-1 and N is the total number of pixels on the boundary. Then three types of contour representations, i.e., curvature, centroid distance, and complex coordinate function, can be defined.

The curvature K(s) at a point s along the contour is defined as the rate of change in tangent direction of the contour, i.e.

$$K(s) = \frac{d}{ds}\theta(\mathrm{s})$$

Where θ (s) is the turning function of the contour. The centroid distance is defined as the distance function between boundary pixels and the centroid (xc, yc) of the object:

$$R\,(s) = \sqrt{\left(X_{S-x_C}\right)^2\left(y_{S-Y_C}\right)^2}$$

The complex coordinate is obtained by simply representing the coordinates of the boundary pixels as complex numbers:

$$\overset{s}{i} = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij-}\mu_i\right)^3}$$

The Fourier descriptor of the curvature is:

$$f_k \;\; = \left\lfloor |F_1|, |F_2|, \dots\dots, \left|F_{M/2}\right| \right\rfloor \dots..(1)$$

The Fourier descriptor of the centroid distance is:

$$f_R = \left[\frac{|F_1|}{|F_0|}\;, \frac{|F_2|}{|F_0|}, \dots..\frac{\left|F_{M/2}\right|}{|F_0|}\right] \dots\dots..(2)$$

Where Fi in (1) and (2) denotes the i^{th} component of Fourier transform coefficients. Here only the positive frequency axes are considered because the curvature and centroid distance functions are real and, therefore, their Fourier transforms exhibit symmetry, i.e., |F-i| = |Fi|.

Circularity Features

The first low-level feature derives a simple object description that includes the object size (in macro blocks), the circularity of the border and the position in the image for the first five objects. A macro block has one 64^{th} of the width and height of the image. The edge histogram has four bins for all edges in an image with length smaller than one macro block, one to two macro blocks, two to four macro blocks and more than four macro blocks. Circularity is computed as:

$$\alpha = \frac{4\pi S}{p^2}$$

Where S is the size and P is the perimeter of an object. This value ranges between 0 (corresponding to a perfect line segment) and 1 (corresponding to a perfect circle).

OPTIMIZATION

Co-occurrence Matrix

A co-occurrence matrix or co-occurrence distribution (less often co-occurrence matrix or co-occurrence distribution) is a matrix or distribution that is defined over an image to be the distribution of co-occurring values at a given offset. Really any matrix or pair of matrices can be used to generate a co-occurrence matrix, though their main applicability has been in the measuring of texture in images, so the typical definition, as above, assumes that the matrix is in fact an image. It is also possible to define the matrix across two different images. Such a matrix can then be used for color mapping.

$$C_{\Delta x,\Delta y}(i,j)=\sum_{p=1}^{n}\sum_{q=1}^{m}\begin{cases}1, & \text{if } I(p,q)=i \text{ and } I(p+\Delta x,q+\Delta y)=j\\ 0, & \text{otherwise}\end{cases}$$

Where i and j are the image intensity values of the image, p and q are the spatial positions in the image I and the offset (Δx, Δy) depends on the direction used \the ta and the distance at which the matrix is computed d. The 'value' of the image originally referred to the grayscale value of the specified pixel, but could be anything, from a binary on/off value to 32-bit color and beyond. Note that 32-bit color will yield a 2^{32} \times 2^{32} co-occurrence matrix.

Whether considering the intensity or grayscale values of the image or various dimensions of color, the co-occurrence matrix can measure the texture of the image. Because co-occurrence matrices are typically large and sparse, various metrics of the matrix are often taken to get a more useful set of features.

Classification

Many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features in recent years. Different similarity/distance measures will affect retrieval performances of an image retrieval system significantly. In this section, some commonly used similarity measures are discussed. *D* (*I*, *J*) denote the distance measure between the query image *I* and the image *J* in the database; and *fi* (*I*) as the number of pixels in bin *i* of *I* (Long *et al*., 2003).

.

Quadratic Form (QF) Distance

The Minkowski distance treats all bins of the feature histogram entirely independently and does not account for the fact that certain pairs of bins correspond to features which are perceptually more similar than other pairs. To solve this problem, *quadratic form distance is used.*

$$D(I,J)=\sqrt{(F_I-F_J)^T A(F_I-F_J)}$$

Where *A*=[*aij*] is a similarity matrix, and *aij* denotes the similarity between bin *i* and *j*. **F** and *J* **F** are vectors that list all the entries in *fi*(*I*) and *fi*(*J*).

Euclidean Distance

The Euclidean distance or Euclidean metric is the "ordinary" distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space. The associated norm is called the Euclidean norm. Older literature refers to the metric as Pythagorean metric. The Euclidean distance between point's p and q is the length of the line segment connecting them.

In Cartesian coordinates, if p = (p1, p2,..., pn) and q = (q1, q2,..., qn) are two points in Euclidean n-space, then the distance (d) from p to q, or from q to p is given by the Pythagorean formula:

$$\mathrm{d}(\mathbf{p},\mathbf{q})=\mathrm{d}(\mathbf{q},\mathbf{p})=\sqrt{(q_1-p_1)^2+(q_2-p_2)^2+\cdots+(q_n-p_n)^2}$$
$$=\sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}.$$

The position of a point in a Euclidean n-space is a Euclidean vector. So, p and q are Euclidean vectors, starting from the origin of the space, and their tips indicate two points.

Mahalanobis Distance

The *Mahalanobis distance* metric is appropriate when each dimension of image feature vector is dependent of each other and is of different importance. It is defined as:

$$D(I,J)=\sqrt{(F_I-F_J)^T C^{-1}(F_I-F_J)}$$

Where *C* is the covariance matrix of the feature vectors.

The Mahalanobis distance can be simplified if feature dimensions are independent.

In this case, only a variance of each feature component, *ci*, is needed.

$$D(I,J)=\sum_{i=1}^{N}(F_I-F_J)^2/C_i$$

CONCLUSION

Color, texture, shape, and spatial information are most widely used visual features in content-based image retrieval. Color is usually represented by the color histogram, color correlogram, color coherence vector, and color moment under a certain color space. Texture can be represented by Tamura feature, would decomposition, SAR model, Gabor and Wavelet transformation. Shape can be represented by moment invariants, turning angles, Fourier descriptors, circularity, eccentricity, and major axis orientation and radon transform. Some of the image processing techniques among these are discussed in this paper. For future research work direction these techniques can be implemented for image retrieval process in the application area of CBIR.

REFERENCES

Bartolini, I., Ciaccia, P., & Patella, M. (2005). WARP: Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27.

Bimbo, A. Del, Mugnaini, M., Pala, P., & Turco, F. (1998). Visual querying by color perceptive regions, *Pattern Recogn*., 31, 1241 -1251.

Costa, Luciano da Fontoura, and Roberto, Marcondes Cesar Jr. (2001). Shape Analysis and Classification, *Theory and Practice*, CRC Pres.

Lee, T.S. (1996). Image representation using 2D Gabor wavelets, *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp: 959-971.

Long, F., Zhang, H., and Feng, D.D. (2003). Multimedia Information Retrieval and Management: Chapter-1 Fundamentals Of Content-Based Image Retrieval, *Springer Berlin Heidelberg*, pp: 1-26.

Ma, W.Y. & Manjunath, B.S. (1995). A comparison of wavelet transform features for texture image annotation, *Proc. IEEE International Conf. On Image Processing*.

Manjunath, B.S., & Ma, W.Y. (1996). Texture features for browsing and retrieval of image data, *IEEE Trans. Pattern Analysis and Machine Intelligence,* 18(8), 837-842.