Content Based Image Retrieval System Using Fusion of Texture Features with Various Distance Metrics

Harshdeep¹, Ankur Gupta²

¹P.G. Student, Department of Computer Science & Engineering, Doon Valley College of Engineering & Technology, Karnal, Haryana, India
²Asst. Professor, Department of Computer Science & Engineering, Doon Valley College of Engineering & Technology, Karnal, Haryana, India

ABSTRACT

Abstract- From previous few years, the necessity of a large database used to reserve vast images has been developed very fast and will also increase in the coming generation. From the enormous database, extracting and querying of the images perfectly is supreme to produce the visual content. Content Based Image Retrieval System (CBIRS) provides us utmost outcome to fetch the images from massive dataset. In these systems, various features like texture, shape, color, spatial information, edge etc. can be used to represent an image. This paper provides the review of basic features of an image denoted by color, texture and shape. For color feature extraction, techniques like color histogram (CH), dominant color descriptor (DCD), Color correlogram etc. have been described. Tamura, steerable pyramid, Gabor wavelet transform, grey level co-occurrence matrix (GLCM) are the techniques described for texture features extraction and lastly Zernike moments and MPEG-7 for shape feature extraction. If we are using single feature to retrieve the image, it takes more retrieval time. In this paper, a review of Content based retrieval system (CBIRS) is explained.

Keyword:- Content Based Image Retrieval System (CBIRS), Grey Level Co-occurrence Matrix (GLCM), Dominant Color Descriptor (DCD), Zernike moments (ZMs).

1. INTRODUCTION

Since last few decades, systems working with retrieving large amount of multimedia data have been growing rapidly. Systems such as search engines, e-business systems, online tutoring system, GIS, and image archive are among few to them. These systems involve retrieving multimedia data based on pictorial content. In the image archive for example, a simple query such as searching for bird with yellow feathers requires the system to be able to find all images in the database which contains a bird with yellow feathers. This is a challenging task since it requires system to browse every single image in database and compare it to query image. Manual browsing the database to search for identical images would be impractical since it takes a lot of time and requires human efforts. A more practical way is to use Content based image retrieval (CBIR) technology. CBIR has provided an automated way to retrieve images based on the content or features of the images itself. The CBIR system simply extracts the content of the query image matches them to contents of the search image [2].

[long 2017] Content-based image retrieval uses the visual contents of an image such as color, shape, texture to fetch and index the image. In typical content-based image retrieval systems, the visual contents of the images in the database are fetched and narrated by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database [2].

To fetch images, users assign the retrieval system with example images or sketched figures. The system then represents these examples or sketch figures by internal feature vectors. The similarities distances between the feature
vectors of the query example or sketch and those of the images in the database are then deliberated and retrieval is performed with the aid of an indexing strategy. The indexing strategy assigns an efficient way to fetch for the image database [1].

![Fig. 1 Block diagram of basic CBIRS [1]](image)

The organization of this paper is as follows: Section I describes introduction. Section II describes the related work. In section III, feature extraction techniques related to various image features are discussed. Distance metrics for Similarity calculation are discussed in section IV. Applications are discussed in section V.

2. RELATED WORK

Naushad Varish et al, proposed a CBIR system based on discrete cosine transformation (DCT) and Grey Level Co-occurrence Matrix (GLCM). By using a hybrid CBIR system based on DC and GLCM, the accuracy of a system is improved. Corel (1k) dataset has been used in which there are 1000 images. The highest precision (100) is obtained [24]. Annrose J. and S. C, presented two different methods to reduce the time and space constraints. In the first method, structured query language (SQL) is used for feature selection for creating normalized feature set through which an efficient CBIR system is developed. In the second method, SQL range query and Euclidean distance are used to filter out initial level relevant image and to smooth the filtered subspace to obtain the relevant image. The experiment is performed on Corel dataset and lowest retrieval time (1.11s) is obtained [4]. P. Chandana et al. introduced GLCM technique for texture features extraction in which conversion of the test and query image into grey level image takes place and then extracting texture features. After that similarity distance is calculated by Euclidean distance measure and KL divergence approach. Highest precision (98.67) is obtained [5]. G.S. Priyadarshini et al. proposed a system based on Dominant Color Descriptor. By using Dominant Color Descriptor, color details were extracted and therefore, performance of CBIR system is improved up to 62% [6]. Y. Mistry et al. introduced Hybrid scheme for CBIR using several distance metrics. For spatial domain features like color auto-correlogram, color moments, HSV histogram features and for frequency domain features SWT, Gabor wavelet transform, CEDD and BSIF features are used to improve precision. Wang dataset is used in which 1000 images are containing with 10 different subjects. By using hybrid scheme highest precision are obtained in Euclidean distance (0.99), city block (0.98), minkowski (0.94), mahalanobis (0.95) [7]. Ali Ahmed Alfaki et al. introduced two techniques: HSV color moment for color feature extraction and Gabor technique for the texture features extraction. This paper provides combination of texture and color features. The mean similarity for hybrid method (0.656 and 0.576 for top 10 and top 20) are retrieved [8]. Jigisha M. Patel et al. introduced color and texture feature and discussed their comparison. Extraction techniques like color histogram, color correlogram, color co-occurrence matrix are used for color features extraction and Tamura texture feature, steerable pyramid, wavelet transform, Gabor wavelet transform are used for texture feature extraction [10]. Nitish Kumar Saini et al. introduced an integration of color moment (CM) and local binary pattern (LBP) which is used for color feature and texture feature extraction and tested on Wang and UCID databases which contained 1000 images [11]. T. Karthikeyan et al. introduced both systems: TBIRS and CBIRS. Just because of some limitations in TBIRS such as task of determining image content, CBIR system came to the lime light to solve this problem. CBIR deliver higher efficiency in covering the semantic gap between high level human intelligence and basic low level features. [14]. S. Mangijao Singh et al.
introduced color moments (CM) and Gabor texture features (GTF). For encoding, a minimal amount of spatial information, we divide the space of an image as three equal regions which are non-overlapping horizontal manner to improve the color indexing techniques. Average retrieval of: GTF (43.6); CMW (55.4); CMW+GTF (58.2) [15].

3. FEATURES EXTRACTION

Feature fetching is a backend operation applied for image database. In this section color, texture and shape techniques used in CBIR are explained.

3.1 Color Features

The significance of this feature is that its extraction as well as matching is simple. Indexing and searching is also effective. The motive of color feature extraction is to extract all images of database whose color composition is almost identical to the query image [15].

3.1.1 Color Histogram: Color histogram is a technique which is used to express the color information. This is widely used method because of its easy implementation, non-sensitive and rapid nature. Rotation, translation and scaling of an image can be changed but histogram of an image cannot be changed [10] [23]. In this method, three color channels namely R, G, B are considered and intensity calculation of their independent probabilities is taken out. Histogram is obtained by computing every color pixels exist in the image and all color pixels are carried in separate bins. This method can be given by:

\[
h_{X,Y,Z}(x,y,z) = N \cdot \text{Prob}(X=x, Y=y, Z=z) \tag{1}
\]

Where \( N \) is represents the no. of pixels in the image and \( X, Y \) and \( Z \) represent three color channels [10].

3.1.2 Dominant Color: DCD accommodates only one component, named as representative color and is based on the clustered mean value. In this technique, first RGB (Red, Green and Blue) color image is partitioned in 10 parts which are to be represented into quantized colors and quantized colors are calculated by taking centroid point for each partition (color bin). RGB are the color components and represent the color of a pixel \( I_{\text{pixel}} \in [R, G, B] \).

By using the formula center value of each partition can be calculated [4];

\[
\text{Cluster Center} = \frac{\sum_{i=1}^{N} I_{x, y}}{\sum_{i=1}^{N}} \tag{2}
\]

\( I(x, y) \) represent the pixel value present in original. \( N \) denotes the totality of picture elements available in specific cluster.

3.1.3 Color Correlogram: A color correlogram provides the information regarding how the colors pairs are changed with distance. The characteristics of color correlogram can be explained as: (i) defines correlation of colors in spatial plane, (ii) implemented to evaluate the overall distribution by local spatial correlation of colors, (iii) computation complexity is less and (iv) its dimensions are small evenly [10].

3.1.4. Color Co-occurrence Matrix (CCM): CCM is a general method that is used to obtain variations in color of the image, where between pixels of same color and their respective neighbors, calculation of probability of occurrence is made [10] [16].

3.2 Texture Features

Texture is a main component of human visual perception. Like color, this also makes it an essential feature to consider when querying image databases. Everyone can recognize texture but, it is not easy to define. Unlike color, texture occurs over a region rather than on a point. It is normally defined purely by grey levels and as such is orthogonal to color. Texture has qualities like periodicity and scale; it can be described in terms of coarseness, direction, contrast. However texture can be considered as repeated patterns of pixels over a spatial domain, of which the addition of noise to the patterns and their repetition frequencies result in textures that can become visible to random and unstructured. Or in other word we can say that texture is that innate property of all surfaces that describes visual patterns, each having properties of homogeneity [9].
2.1 Gray Level Co-occurrence matrix (GLCM):

The architecture composed of two different phases: Phase 1 Preprocessing and Phase 2 Feature extraction phase, phase 1 processes the input data, and phase 2 extracts the features and combines both software and hardware to calculate GLCM features. The GLCM feature vectors are calculated in hardware and the software supports hardware by performing additional computations. The preprocessing phase passes the conversion of the image into an array that is suitable for processing by the feature extraction phase. Each element \( a = [a_0, a_1, a_2, a_4] \) of \( A \) corresponds to each pixel, and it is formed by five integers, \( [a_0] \) is the gray level of the corresponding pixel, and \( [a_1, a_2, a_3, a_4] \) are the gray-level intensities of first neighbors in four directions. The resulting near quantization of the image intensities leads to 16-bit representation for each element [5].

The equations for identifying texture measures can be computed from gray-level co-occurrence matrices as shown below:

\[
\text{Mean} = \sum_i \sum_j P(i,j) \times i \\
\text{Variance} = \sum_i \sum_j P(i,j) \times (i - \mu^2) \\
\text{Homogeneity} = \sum_i \sum_{j \neq i} \frac{P(i,j)}{1 + |i-j|} \\
\text{Contrast} = \sum_i \sum_j P(i,j) \times (i - j)^2 \\
\text{Entropy} = \sum_i \sum_j -P(i,j) \times \log_2 P(i,j) \\
\text{Angular Second Moment} = \sum_i \sum_j P(i,j)^2 \\
\text{Correlation} = \sum_i \sum_j (i - \mu_x) \times (j - \mu_y) \times P(i,j) \\
\text{Dissimilarity} = \sum_i \sum_{j \neq i} \frac{1}{1 + (i-j)^2} \times P(i,j)
\]

3.2.2 Tamura Texture Feature:

Tamura features are based on the human visual recognition. Fundamentally there are six tamura texture features and they are contrast, coarseness, line likeness, directionality, regularity and roughness. These features are used in psychological studies by considering human perception of texture [10].
3.2.3 Steerable Pyramid:
This technique recursively partitions an image into a low pass residual and a set of oriented sub-bands. Decimated low pass sub bands and set of decimated oriented sub bands are obtained by dividing the original image. It is a useful technique which has many advantages as compared to other methods like multi-orientation, multi scale image perishing. Kth- order directional derivative operator is a major function of this technique which is found in many sizes and has orientation of K+1 [13].

3.2.4 Gabor Wavelet Transform:
This transform is used to dilate and rotate the two dimensional Gabor function. Gabor function is basically established by sampling all frequency domains and by finding all parameters of orientation and particularly its center frequency. In this method, both Gabor filter and Gabor wavelets of a particular spatial frequency can be applied to an image. This function denoted by g(x, y) can be expressed by equation [13]:

\[ g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp [-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}) + 2\pi j f_c] \quad (11) \]

Here center frequency is given by \( W \), \( \sigma_x \) and \( \sigma_y \) are scaling factors of the filter.

3.3 Shape Feature Extraction Techniques:

Basically shape describes outlines and entire area of an image. This technique is divided into two parts one is region based (defines whole area of an object) and other is contour based (describes only boundary lines) [25].

3.3.1 Zernike moments (ZMs):
ZMs are basically an overall descriptor therefore doesn’t provide local characteristics of an image. ZMs are a good descriptor in the various fields like pattern recognition, character identification, image fetching, image reconstruction, vehicle recognition etc. [25].

3.3.2 MPEG-7 descriptors:
The MPEG-7 is divided into the following types: RSD (region shape descriptor), CSD (contour shape descriptor), 3D shape descriptor. RSD introduces pixel distribution within a 2D object. It is depends on internal pixels and boundary. CSD is based on representation of the contour in the curvature scale space. Among these, 3D shape descriptors are preferred because of their major application in real world. [26].

4. SIMILARITY MEASUREMENT

To obtain similarity between any query image and database images, various distance metrics are used. These distance metrics computes the difference between database image feature vector and query image feature vector. Smaller the distance, similarity between two images will be more. Various distance performance parameters are used such as, Euclidean distance, Manhattan, KL distance and Mahalanobis distance etc.[7].

A. Euclidean distance: The most popular and widely used similarity measure in image retrieval is Euclidean and is given by:

\[ D_E = \sqrt{\sum_{i=1}^{n}(I_i - D_i)^2} \quad (12) \]

B. City block distance: This method also called by the name of Manhattan distance and is expressed by the following given equation:

\[ D_C = \sum_{i=1}^{n}|I_i - D_i| \quad (13) \]

Performance evaluation metrics: To examine the performance with an image retrieval systems two major evaluation metrics, Precision and Recall are used [7].
Precision: \[ P = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \]  
(14)

Recall: \[ R = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images in the database}} \]  
(15)

APPLICATIONS

In biometric protection, these systems are concentrated on recognizing person depending on their notable physical and behavioral traits. These days biometric protection is used in different fields as per the demand [22]. In medical field, for diagnostics and therapy; images and digitized images are generated in more quantity [20]. In Crime Prevention, to identify the criminals through CCTV camera and find out their complete information from database. In military domain, recognition of target from satellite photographs, identification of enemy aircraft by using radar screen. In Digital Library, various digital libraries are used to support services based on image content like, geographical and historical. In Fashion and interior design, retrieving pattern, specific combination of color and texture provide important aids in designing process.

CONCLUSION

The review paper narrates preliminaries of CBIR system and its prime features distinguished by color, texture and shape. As correlate to ancient methods, the combination of features assigned more correct outcomes. Color is usually described by the color histogram, color co-relogram, DCD dominant color descriptor, and color moment. Texture can be explained by Tamura feature, GLCM grey level co-occurrence and Gabor. Shape can be described by Zernike moment and MPEG-7 descriptor. The advantage of CBIRS techniques is that this is best suitable for large databases. The future scope of the above discussed systems is that issues related optimization and machine learning can be implemented in the basic CBIR systems to make them more relevant in the process of retrieval. It can be applied on various databases such as, COREL (5k, 10k), WANG.

REFERENCES


