An Adequate Approach to Image Retrieval Based on Local Level Feature Extraction

SUMAIRA MUHAMMAD HAYAT KHAN*, AYYAZ HUSSAIN**, AND IMAD FAKHRI TAHA AL SHAIKHLI***

RECEIVED ON 29.12.2014 ACCEPTED ON 17.03.2015

ABSTRACT

Image retrieval based on text annotation has become obsolete and is no longer interesting for scientists because of its high time complexity and low precision in results. Alternatively, increase in the amount of digital images has generated an excessive need for an accurate and efficient retrieval system. This paper proposes content based image retrieval technique at a local level incorporating all the rudimentary features. Image undergoes the segmentation process initially and each segment is then directed to the feature extraction process. The proposed technique is also based on image’s content which primarily includes texture, shape and color. Besides these three basic features, FD (Fourier Descriptors) and edge histogram descriptors are also calculated to enhance the feature extraction process by taking hold of information at the boundary. Performance of the proposed method is found to be quite adequate when compared with the results from one of the best local level CBIR (Content Based Image Retrieval) techniques.

Key Words: Content Based Image Retrieval, Feature Extraction, Feature Vector, Similarity Measure, Edge Histogram Descriptor, Fourier Descriptor.

1. INTRODUCTION

Image retrieval is a challenging task in the area of image processing and it requires an efficient and accurate system for retrieval of the most similar image from a large collection of images. One of the ways to retrieve an image is to associate it with textual description but it is strenuous and time-consuming. Text based image retrieval is also ambiguous because of different views of researchers regarding the content of image. Image retrieval based on the content of image is another approach to image retrieval without associating text with the image and with reduced computational complexity. There are various techniques proposed by the researchers for the effectual CBIR. Visual contents of an image are most widely used methods that give satisfactory results for the retrieval of the image based on its contents. These image contents are indexed by three main features such as texture, color and shape [1].

Image feature extraction and feature vector calculation is the core step of image retrieval based on the content of image. Imperative features of an image include; color, texture and shape. Different scholars have proposed diverse combinations of these features used in different techniques. Feature extraction is one of the crucial phases...
used for CBIR. Feature extraction is basically performed at two levels; global level feature extraction and local level feature extraction. Global level feature extraction extracts image features from the whole image while local techniques extract features from different parts or regions of the same image. Both approaches have their own pros and cons. Global level feature extraction lacks important information about spatial feature distribution. Secondly, it is responsive to distortion, variations and intensity. Spatial limitations have been overcome by spatial chromatic histogram, spatial color histogram, color coherence vector and color correlogram [2]. Region based image retrieval methods are popularly used among various local feature based approaches. In this type of image retrieval every image is first divided into smaller regions and different facets of the region are then extricated. Furthermore, the correspondence between two images is calculated based on the parallel region based features. Segmentation based on texture, color shape and spatial location are further used to explore and retrieve analogous regions from the database [3]. Segmentation algorithm plays a vital role in feature extraction at local level. There are various ways of achieving image segmentation and each method is based on certain properties [4]. Since the proposed technique is based on the local approach, segmentation plays crucial role here. Segmentation method chosen here is the most suitable one as it prevents the loss of information by preventing division of an image into multiple segments. K-means clustering technique is used and clusters are restricted to three. The proposed technique can be divided into the following main steps:

- **Segmentation is performed by K-means clustering technique and the number of clusters is restricted to three to prevent any loss of information. Also it reduces the time complexity of the proposed system.**

- **All the features such as color, texture, shape, FDs and EHDs (Edge Histogram Descriptors) after extraction are combined to form a single collective feature vector. This single feature vector for all the features reduces the space complexity.**

- **Comparison of the feature vectors of two images is found by measuring Euclidean distance among different feature vectors of images.**

- **Results are found at different precisions and then the average of different precision values is compared with the average precision of one of the existing local level CBIR technique. Main contributions of the technique are listed below.**

- A rational segmentation technique is adapted in which segmentation is done by k-means clustering and the number of clusters is limited to 3 to avoid over segmentation.

- All the three vital image features such as color texture and shape are utilized during the calculation of feature vector.

- FDs and EHDs are also added to the feature vector for the comparison of two images to get the effective retrieval results.

Document is organized as: Section 2 explains related work and background of CBIR. Proposed technique is discussed in detail in Section 3. While, Section 4 presents experimentation and comparison of the proposed technique with other existing methods. Conclusion will be made in Section 5 along with some future directions.

## 2. RELATED WORK

A renowned region based image retrieval technique at a local level known as UM (Universal Model) was proposed by Nandadgopalan, et. al. [5], UM technique made use of the color and texture features along with the EHDs. In another well-known method for image retrieval presented by Brandt, et. al. [6] calculation of edge histograms was incorporated to extract information related to the shape of...
An Adequate Approach to Image Retrieval Based on Local Level Feature Extraction

An object. Fuzzy logic has also been added to the process of image feature extraction and region comparison. A novel synthesized framework for retrieval of image had been put forward which is based on texture features, regional color features, the global and semi-global EHDs. Moreover, texture features and region based fuzzy color are integrated for region matching. For image segmentation, color-clustering based method is used since color features are considered more important among all the other features. The image is divided into different regions before feature calculation and hence is one of the local based approaches to CBIR. Feature calculation and distance measurement between the regions were based on the fuzzy logic to avoid crisp decisions [7]. Different scholars have proposed shape feature for image retrieval as well such as Lin, et. al. [8] proposed a MCS (Mountain Climbing Sequence) in this regard which is specified for rotation, translation and scale problems. Torres, et. al. [9] also proposed idea about shape feature of an image and it states that genetic programming approach can be used in the image retrieval process in CBIR [9]. Genetic programming technique has been tried because of its success in other machine learning applications e.g. it can provide better results for pattern recognition than classical techniques such as SVM (Support Vector Machine). It allows nonlinear combination of descriptors. Images are retrieved here on the basis of shape of their objects. GP (Genetic Programming) is used to create a composite descriptor. It combines the predefined descriptor using GP technique which is used to combine the similarity values obtained from each descriptor fusing them into an effective similarity function. BAS (Bean Angle Statistics) and multiscale fractal dimensions are shaped descriptors which can be used in combination with different similarity functions. Among different methods, Hsieh, [10] proposed another technique to characterize every illustration by a number of patterns and their structural relations. Each arrangement of an image is represented by various dominant segmented regions. These patterns reveal diverse arrangements on an object in different states. The joint geometric, texture and color feature divergence among every region and pattern is calculated for finding visual similarity which combines with relation similarity to find the global similarity.

Machine learning can also be incorporated into image retrieval. There are certain situations where CBIR system is applied for a special task. These images are searched from a particular domain and there is a set of queries and known relevant images. In such situations parameter learning for a system is possible to optimize retrieval performance. Relevance feedback is a process where user interacts with the system to tell about the relevant images for optimal retrieval. The machine learning technology used in CBIR [11] indicates a survey of relevance feedback technologies for the retrieval of an image until 2002. A large amount of approaches have characterized the images as individual queries. One of the recent approaches however, is a query-instance-based approach [12] or use support vector machine for two class classifier [13]. This approach is similar to the one presented in [12] as both follows a nearest neighbor approach but instead of using only the best matching query/database image combination, all query images are considered jointly.

In this paper, a local technique for CBIR is proposed in which all the important image features are extracted and utilized for the calculation of feature vector. The proposed technique is based on the idea of feature extraction at local level and therefore the first important step here is image segmentation. Segmentation is performed based on the color feature of image using k-means clustering algorithm. Feature vector is then calculated for each image using all the three basic image features i.e. color, texture and shape. In addition to these basic image features, EHDs and FDs are also incorporated in the calculation of feature vector. Difference between two images is found by calculating the uclidean distance between them.
3. PROPOSED TECHNIQUE

This paper proposes a new technique for CBIR based on the local level feature extraction. Image segmentation is therefore the foremost step for image retrieval like other CBIR techniques based on the local approach for feature extraction. After segmentation, all the vital image features are extracted. The proposed technique extracts all the three important features like color, texture and shape. EHDs and FDs are also calculated to extort information at the edges. Flow chart of the proposed technique is shown in Fig. 1.

**Algorithm**

**Step-1:** Divide the image into 3 clusters by k-means clustering technique.

**Step-2:** Color histogram is constructed to extract color feature \( f \) for each segment.

**Step-3:** Extract texture feature \( f \) by calculating Co-occurrence matrix \( C(i, j) \).

**Step-4:** Let \( x[k] \) and \( y[k] \) be the coordinates of the \( k \)th pixel on the boundary of a given 2D shape containing \( n \) pixels, a complex number can be formed as \( z[k] = x[k] + py[k] \), and the FD of this shape is defined as the DFT of \( Z[k] \).

\[
DFT[Z[k]] = \sum_{k=0}^{N-1} z[k] e^{-\frac{2\pi i}{N} kn} \quad (N = 0, \ldots, n-1)
\]

**Step-5:** EHD is calculated from four directional edges namely vertical, horizontal, 45 degree, 135 degree and one is non-directional edge.

**Step-6:** Euclidean distance is used to calculate the distance between feature vector of database images \( t_d \) and feature vector of query image \( t_q \).

### 3.1 Features of an Image

Images are mainly characterized by features like color, texture and shape [14]. We have used all these features along with the EHD to calculate information at the edges.

---

**FIG. 1. FLOW CHART OF PROPOSED TECHNIQUE**
3.1.1 Color Feature

Color is an image feature represented through some color model and is one of the most important and extensively used visual features. Most frequently used color models are RGB, HSV and YCbCr. Different formulas are applied to the RGB model to create HSV and YCbCr models. Histogram is used for color representation. Histogram does not give semantic information and is invariant to rotation, translation and scaling of an object. Therefore, two images with same histograms may differ in contents. Image retrieval based on the color content is mostly done through color indexing process. Color moments show distribution of color in an image [12,13] (Table 1). Extraction of color, efficient indexing and effective retrieval are required to be satisfied for a good system [3].

Proposed technique extracts color feature by finding mean, mode, median and standard deviation for each color channel. Therefore, twelve color features are extracted for each of the database images and query image. Pixel count and gray levels are extracted from the histogram of the gray image. Each time one color channel is considered for the calculation of the above mentioned features.

3.1.2 Texture Feature

Another important image feature used in the process of image retrieval is texture. Texture mainly uses structural and statistical methods for representation [13]. According to human visual feature, the similarity of most images is distinguished by the relative coarse texture feature mainly. The grey-scale of the initial image will be compressed to reduce calculation before the co-occurrence matrix is formed. 16 compression levels were chosen in the paper to improve the texture feature extracting speed. Co-occurrence matrix texture feature is a technique to extract information from second order texture. It finds out the occurrence probability of a pixel pair with intensity spacing between pixels in two dimensions.

Texture feature calculation is performed by constructing a co-occurrence variance matrix. A gray level image is given as input for calculation of a co-occurrence variance matrix. Ten different features M (Mean), SD (Standard Deviation), SK (Skewness), MAXP (Maximum Probability), UNI (Uniformity), ENT (Entropy), difference moment, IKMOM (Inverse Element Difference Moment of Order 2), HOMO (Homogeneity), CONT (Contrast) and CORR (Correlation) are extracted by using different formulas. All the above mentioned texture features are presented with their formulas in Table 2.

3.1.3 Shape Feature

Shape based image retrieval is measuring of similarity between shapes of query image and database images. Shape is an important visual feature and it is one of the primitive features for image content description. Shape

<table>
<thead>
<tr>
<th>Image Feature</th>
<th>Equation with Symbol Definition</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Median = ( m = \frac{\sum_{i=1}^{n} Y_i}{n} )</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>where ( X ) = gray level, ( i ) = index</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>Median = ( m = \frac{n + 1}{2} )</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>where ( n ) = pixel count</td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>Mode = max(( X_i ))</td>
<td>(5)</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>Standard Deviation = ( \sigma = \sqrt{\frac{\sum(X - m)^2}{n - 1}} )</td>
<td>(6)</td>
</tr>
</tbody>
</table>

**Table 1. Color Features Used in the Proposed Technique**
content description is difficult to calculate because measuring the similarity between shapes is complex. Two main techniques used for shape feature extraction are contour-based and region-based [2]. Former is used to extract information from the boundary of an image and later is used to extract information from the interior. Former uses only the information present in the contour of an object. Later uses the whole area. We have utilized contour

based shape feature extraction methods. Eight different features are extracted here from the contour of an image. They are given in Table 3 with their equations.

### 3.1.4 Fourier Descriptor Method

General FD method mainly consists of computation of boundary pixels, use of shape signature function, and computation of FD. Boundary pixel computation for

<table>
<thead>
<tr>
<th>Image Feature</th>
<th>Equation with Symbol Definition</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>[ SK = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{x_i - m}{\sigma} \right)^3 ]</td>
<td>(7)</td>
</tr>
<tr>
<td>Correlation</td>
<td>[ CORR = r = \frac{\sum_{i,j} (C(i,j) - \mu_i \mu_j)}{\sigma_i \sigma_j} ]</td>
<td>(8)</td>
</tr>
<tr>
<td>Maximum Probability</td>
<td>[ MAXP = \max \left( \frac{n(A)}{n} \right) ]</td>
<td>(9)</td>
</tr>
<tr>
<td>Uniformity</td>
<td>[ UNI = \left( 1 - \frac{\sigma}{m} \right) \times 100 ]</td>
<td>(10)</td>
</tr>
<tr>
<td>Entropy</td>
<td>[ ENT = \sum_{i,j} C(i,j) \log C(i,j) ]</td>
<td>(11)</td>
</tr>
<tr>
<td>Contrast</td>
<td>[ CONT = \sqrt{\frac{1}{MN} \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} (I_{ij} - I)^2} ]</td>
<td>(12)</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>[ HOMO = \sum_{i,j=0}^{N-1} \frac{C(i,j)}{1 + (i-j)^2} ]</td>
<td>(13)</td>
</tr>
<tr>
<td>Difference Moment</td>
<td>[ IKMOM = \sum_{i,j} \frac{1}{1 + (i-j)^2} C(i,j) ]</td>
<td>(14)</td>
</tr>
</tbody>
</table>
boundary coordinates is performed through edge detector and boundary tracing techniques [15]. A pixel set can be formed after the computation of boundary pixels and is represented by the following formula.

\[ P = \{(x(t), y(t))| t \in [1, N]\} \]  \hspace{1cm} (20)

FD method uses pixel coordinates of shape boundary in an image to compute shape signatures. And Fourier transforms are used for these shape signatures to compute Fourier transformed coefficients. The standardized Fourier transformed coefficients are used as FD [16].

3.1.5 Edge Histogram Descriptor

Edge is also an important image feature and can be used to extract information present at the contrast. EHDs represent the spatial distribution of five types of edges.

Among five different types of edges represented by EHD four are called directional edges namely vertical, horizontal, 45 degree, 135 degree and one is non-directional edge. Edge strength is detected by the application of filter coefficients shown in Fig. 2. Edge blocks with a value greater than a given threshold is selected.

MPEG-7 standards revealed that retrieval performance of images can be improved by calculation of EHD. This descriptor shows invariance to scale and rotation[17].

The EHD represents the spatial distribution of five types of edges, namely four directional edges and non-

\[ \begin{array}{cccccc}
1 & 1 & \sqrt{2} & 0 & 2 & -2 \\
-1 & -1 & -1 & 0 & -\sqrt{2} & 2 \\
\end{array} \]

\[ \begin{array}{cccc}
1 & 1 & -1 & -1 \\
0 & \sqrt{2} & 2 & 2 \\
\end{array} \]

**FIG 2. FILTER COEFFICIENTS**

<table>
<thead>
<tr>
<th>TABLE 3. SHAPE FEATURES USED IN THE PROPOSED TECHNIQUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image Feature</strong></td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Circularity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Aspect Ratio</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Complexity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Discontinuity angle irregularity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Solidity</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
directional edge. According to MPEG-7 standard, the performance of image retrieval can be considerably improved by combining the EHD with other features for example the color histogram descriptor. The descriptor is invariant to scale. It also supports rotation invariance and rotation related matching operations [17].

4. EXPERIMENTAL RESULTS

Implementation of the proposed technique utilizes color, texture, shape and edge information for search and retrieval. The system is developed in MATLAB. A database of 1000 natural images is used from Corel image database. It consists of 10 different image categories with different semantics each having 100 images. Images from all the 10 image categories were tested by the proposed system.

4.1 Impact of FD and EHD on Image Retrieval

In addition to the three BF (Basic Features) such as color, texture and shape, FDs and EHD are also integrated to extract maximum statistics of image. FD is invariant to translation, rotation and scaling. FDs correspond to the low frequency components of the boundary to represent 2D shape. The restored shape based on these descriptors estimate the shape without details corresponding to high frequency modules vulnerable to noise. Fig. 3 shows comparison of retrieval results for two feature vectors. Retrieval results of feature vector integrated with FD and EHD are better than the retrieval result of feature vector with BF for five distinct image categories. FDs and EHDs helped in the detection of edges and boundaries in the images. Results show that the retrieval precision of images from beach image category is low compared to other image categories. FDs and EHDs helped to detected edges and boundaries in the images. Results show that the retrieval precision of images in the beach image category is low compared to images in other image categories. Other image categories have more objects resulting in a large number of edges. These edges are detected well by FDs and EHDs while beaches have comparatively less number of edges and therefore the difference between the average precision is less.

4.2 Comparison of Retrieved Results with UM Technique

Local level CBIR technique is proposed where images are first segmented based on the color content and then image features are extracted from different segments of the image. The proposed technique is compared with one of the local level feature extraction techniques in literature i.e. the UM. Results show that the proposed technique has higher precision compared to UM. Comparison of average retrieval precision of proposed technique with UM is given in Table 4 for four distinct image categories. The overall retrieval precision at 10 images returned by using 4 different image categories is compared in Fig. 4.

Fig. 4 shows that the proposed method has better result for all types of images. It is the simple segmentation technique and a larger number of features that have

![FIG. 3. COMPARISON OF RETRIEVAL RESULTS OF FEATURE VECTORS USING FD AND EHD WITH BASIC FEATURES](image-url)
improved retrieval results for the proposed technique. Features used by UM are color, texture and EHD. If shape and spatial features are added to UM they will give more discrimination power to the systems [18]. Combination of all the imperative image features has great impact on the retrieval efficiency. For dinosaur image category we only have one object in the image and for this image category the proposed system has improved result compared to the UM technique. Implication of the proper segmentation technique, addition of shape features and removal of threshold improved retrieval result for the proposed method.

4.3 Variants of UM

UM is a CBIR technique extracting image features at a local level proposed in 2008. Color and texture feature from each segment of an image are extracted for the calculation of feature vector. EHD are also calculated to grasp information at the edges. Matching is based on greedy algorithm where a threshold is set first to compare query image’s segment with the image segments in database. Here fine tuning of threshold is very testing. Threshold value is different for each image category and is inconsistent even for each image within the particular image category [18]. In this study the following two variants of UM are also proposed.

(1) UM with shape feature.
(2) UM without threshold.

After comparing the results of the first variant i.e. UM with shape and original UM we have found that there is no improvement in the results. It means that addition of shape feature does not work well for UM model. However second variant i.e. UM without threshold has shown better results compared to the original UM method. Table 4 illustrates the comparison of UM model with its two variants. Threshold (for greedy algorithm) reduces the retrieval precision and increases the time complexity. Also fine tuning of the threshold is very difficult for UM model. It gives different precision results every time the same threshold is given for the same image. Hence there is no consistency in results. Variants of UM model are applied.

<table>
<thead>
<tr>
<th>No.</th>
<th>Image Category</th>
<th>LM</th>
<th>UM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vehicle</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>2</td>
<td>Dinosaur</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>Horse</td>
<td>81</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Flower</td>
<td>97</td>
<td>92</td>
</tr>
<tr>
<td>Average</td>
<td>-</td>
<td>79</td>
<td>66</td>
</tr>
</tbody>
</table>

**TABLE 4. COMPARISON OF AVERAGE RETRIEVAL PRECISION OF PROPOSED TECHNIQUE WITH UM FOR FOUR DISTINCT IMAGE CATEGORIES**

![Graph of retrieval precision for different image categories](image_url)

**FIG. 4. COMPARISON OF THE OVERALL RETRIEVAL PRECISION OF 10 IMAGES RETURNED BY USING 4 DIFFERENT IMAGE CATEGORIES**
and tested on six different image categories of Corel image database. Table 5 summarizes the results for the two variants of UM in comparison with the original UM technique for different image categories. Comparison of results for UM and its two variants for six distinct image categories is shown in Fig. 5.

4.4 Comparison of the Proposed Technique with Recent Local Level CBIR Techniques

The proposed technique is tested with different number of images retrieved and the results are compared with two recent CBIR techniques from the literature. Fig. 6 shows the comparison of average precision of the proposed technique with FRCE and PRIR at 10, 20, 30, 40 and 50 number of images retrieved. Results show that the proposed technique has better average precision at different number of images retrieved in comparison with FRCE and PRIR.

5. CONCLUSION

Image retrieval is an important area in image processing which needs an efficient and accurate system for the retrieval of the most similar image from the database. The more the number of features extracted the greater will be the information taken out from the image and result is better retrieval. The entire three vital image features need to be counted in the feature vector calculation to avoid any loss of information from different parts of the image. Feature extraction at a local level needs a good segmentation technique to divide the image into meaningful parts. The technique proposed in this paper is based on the local approach for calculation of feature vector. Proposed technique showed better results in comparison with some promising local techniques from the literature. Euclidean distance is used as a distance measure to calculate the similarity between query image and database image. Performance of our proposed technique is measured through precision and recall. Further improvement can be achieved by the application of computational techniques like genetic algorithm to find the best distance measure for a specific image feature.
An Adequate Approach to Image Retrieval Based on Local Level Feature Extraction

ACKNOWLEDGEMENT

Authors are thankful to the reviewers for their intuitive comments and useful recommendations to improve this manuscript.

REFERENCES


An Adequate Approach to Image Retrieval Based on Local Level Feature Extraction


