An Effective Content Based Image Retrieval (CBIR) System based on Model Approach

Ravi Rastogi¹ Rohan Appasaheb Borgalli²
¹Scientist ‘C’ ²Assistant Professor
²Department of Electronics & Telecommunication Engineering
¹NIELIT, Gorakhpur, UP, India ²Shah and Anchor Kutchhi Engineering College, Mumbai, India

Abstract— Image retrieval has been one of the most interesting and vivid research areas in the field of computer vision. Content-based image retrieval (CBIR) systems are used in order to automatically index, search, retrieve and browse image databases. Color and texture features are important properties in content-based image retrieval systems. This paper introduces a Effective Content Based Image Retrieval (CBIR) system based on Model Approach. Initially, the color, shape, edge and texture feature of query image is extracted using different algorithms and also for the database images is extracted in a similar manner. Subsequently, similar images are retrieved utilizing a combination of above features. And finally Model Approach [1] is applied which improved efficiency of system. Thus, by means of the Effective Content Based Image Retrieval (CBIR) system based on Model Approach, the required relevant images are retrieved from a large database based on the given query. The proposed CBIR system is evaluated by querying different images and the efficiency of the proposed system is evaluated by means of calculating different parameters to test the efficiency of different techniques and the combination of them to improve the performance of the retrieved results.

Keywords: Content based image retrieval (CBIR); Human visual system (HSV); Gray level co-occurrence matrix (GLCM); Edge histogram descriptor (EHD); Query by image content (QBIC)

I. INTRODUCTION

As processors become increasingly powerful, fast and memories become increasingly cheaper, the deployment of large image databases for a variety of applications have now become practical. Databases of art works, designs, satellite and medical imagery have been attracting more and more users in various professional fields — for example, geography, medicine, architecture, advertising, design, fashion, and publishing. Effectively and efficiently accessing desired images from large and varied image databases is now a necessity.

CBIR or Content Based Image Retrieval is the retrieval of images based on visual features such as color, edge, texture and shape. Reasons for its development are that in many large image databases are available to use and evaluate results, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete [2]. This is not CBIR. In CBIR, each image that is stored in the database and its features extracted and compared to the features of the query image. It involves two steps:

1) Feature Extraction: The first step in the process is extracting image features of query image and other images from database to a distinguishable extent.
2) Matching: The second step involves matching these features with database images to yield a result that is visually similar.

II. LITERATURE REVIEW

CBIR research started in the early 1990’s and it is likely to continue during the in 21st century [8]. It is a problem of searching for digital images in large database using the visual content of the images.

There is a interest growing in CBIR because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. In order to overcome these limitations, CBIR was first introduced by Kato in 1992 [12]. The term, CBIR, is widely used for retrieving desired images from a large collection, which is based on extracting the features from images themselves. In general, the purpose of CBIR is to present an image conceptually, with a set of low-level visual features such as color, texture, and shape [14].

A large number of general-purpose image search engines have been developed in past. Query by Image Content (QBIC) [4] system was first commercial system for CBIR developed by IBM Almaden Research Center, San Jose in 1995.

NETRA [15] system has been developed by Department electrical and computer engineering at University of California, Santa Barbara in 1997. It uses three feature vectors to represent the image. The first vector is computed from color histogram to represent image color feature. The second vector is the normalized mean and standard deviation, derived from the Gabor Wavelet Transform of the image, to represent image texture feature. The third vector is the curvature function of the contour to represent image shape feature.

Similarity matching is done by the Euclidean distance, Key-point Indexing Web Interface (KIWI) [6], has been developed in France by INSA Lyon in 2001. This system extracts the key points in the query image instead of entire image using wavelet-based salient point detector. Color histograms, are computed from each color component (R, G, and B), and the shape descriptors, computed from Gabor Filter, are used as image features. Euclidean distance is used for similarity matching. Photobook [7] was developed by Vision and Modelling Group at MIT Media Lab in 1997.

However, although these systems formulate the special semantic features for image retrieval, still these are not perfect descriptions for semantic features. This is due to the diversity of visual features, which widely exists in real
applications of image retrieval. Image Miner [18] has been developed by Technology-Zentrum Informatics at University of Bremen in Germany in 1997. In this system similarity measurement special module is developed. Chin-Chin Lai et.al. [9] have proposed an interactive genetic algorithm (IGA) to reduce the gap between the retrieval’s results and the expected results. This method gives 80.6% precision value and 15.8% recall value. Meenakshi Madugunki et.al. [20] have explained detailed classification of CBIR systems. They used color histogram for color, discrete wavelet transform (DWT) for texture feature. Nhu-Van Nguyen et.al [11] have proposed clustering and image mining technique. This method got average precision 20% clustering repeat (CR) and 28% clustering no repeat (CNR). Same technique is used by A. Kannan et.al [10], the main objective is to remove the data loss and extract useful information. Zhang XU-bo et.al. [13] Proposed improved K-means clustering and relevance feedback. Sharadch Ramaswamy et.al. [22] have proposed method on fast clustering-based indexing technique. This method got 100% precision result. OLIVE (2008) provides dual access to web images and PIRIA visual search engines.

III. PROPOSED IMAGE RETRIEVAL SYSTEM

The search for similar images in image databases has been an active research topic in the last few years. A very efficient approach is content based image retrieval (CBIR). In such systems, images are typically represented by their contents. Typical approximations consist of statistics of the raw image data. Aim of feature extraction is extracting information that is semantically meaningful but needs a small amount of storage [29].

A detailed description of some feature extraction techniques is introduced in following section. The information gained by feature extraction is used to measure the similarity between two images. Images are represented by points in the high dimensional feature space. Each extent of the feature corresponds to one dimension in the feature space. A metric is defined to calculate the actual similarity between two of these points. Once all features are extracted and metric is formed then all feature metric is combined to single metric which is used for matching similarity of query image with each of images present in database. For finding similarity distance measurement parameter is used such as Euclidean distance measurement. On the bases of which feature contribute more weightage will be assigned to those features and measured least distance of corresponding images indicate maximum similarity. Based on results top 10 images displayed which have maximum similarity with query image.

IV. FEATURE EXTRACTION

Feature extraction is extracting compact but semantically valuable information from images. This information is used as a indication for the particular image. Similar images should have similar indications. If we look at the image shown in Figure 1, Features such as white color and the texture of the building are characteristic properties. In a similar way, the sky can be indicated by its blue color. Furthermore, we can take the shape of the objects in the image into account. Also the count of vertical, horizontal, and curved edges present in the image is important. Representation of images needs to consider which features are most useful for representing the contents of images extraction of the image in the database is typically conducted off-line so computation complexity is not a significant issue.

This section introduces three features: texture, edge, and color, which are used most often to extract the features of an image.

A. Color Feature

Color is perception caused by light as it interacts with our eyes and brain. The perception of color is greatly influenced by nearby colors in the visual scene. The human eye contains two types of visual receptors: rods and cones. The rods are responsive to faint light and therefore, sensitive to small variations in luminance. The cones are more active in bright light and are responsible for color vision. Ones in the human eye can be divided in three categories, sensitive to long, middle, and short wavelength stimuli. Roughly these divisions give use to the sensations of red, green, and blue. The use of color in image processing is motivated by two principal factors. First, color is a powerful descriptor that facilitates object identification and extraction from a scene. Second, humans can discern thousands of color shades and intensities, compared to about only two dozen shades of gray.

The color histogram describes the color component present in an image; it counts the number of occurrences of each color in an image. The color histogram of an image is rotation, translation, and scale-invariant; therefore, it is very suitable for color-based CBIR: content-based image retrieval using solely global color features of images. However, the main drawback of using the color histogram for CBIR is that, texture and shape-properties are not taken into account. This may lead to unexpected errors.

Color image can be represented using three primaries of color space. But the RGB space does not
correspond to the human way of perceiving the colors and does not separate luminance component from chrominance ones therefore we use HSV color space in our approach. HSV is an intuitive color space in the sense that each component corresponds directly to visual perception. Hue is used to differentiate colors. Saturation indicates percentage of white light added to a pure color. Values give the amount of intensity. The mean of the pixel colors states color of the image and standard deviation can represent variation of a pixel colors.

B. Texture feature

In the field of image processing and computer vision available texture definitions are based on texture analysis methods and the features extracted from the image. However, texture can be thought of as repeated patterns of pixels over a spatial domain. Texture properties are the visual patterns in an image that have properties of homogeneity that do not result from the presence of only a single color or intensity. The different texture properties as perceived by the human eye are regularity, directionality, smoothness, and coarseness. In real world scenes, texture perception can be far more complicated.

The various brightness intensities give rise to a different human perception of texture. Image textures have useful applications in recognition of regions using texture properties, known as texture classification, pattern recognition using texture properties known as texture segmentation, texture synthesis, and generation of texture.

There are two widely used approaches to describe the texture of a region, these are statistical and structural. The statistical approach considers that the intensities are generated by a two-dimensional random field. The methods used are based on spatial frequencies and yield characterizations of textures as smooth, coarse, grainy, etc. Examples of statistical approaches to texture analysis are autocorrelation function, gray-level co-occurrence matrix, Fourier texture analysis, edge frequency, and Law’s texture energy measures. The structural techniques deal with the arrangement of image primitives, such as the description of texture based on regularly spaced, parallel lines. In our research, the co-occurrence matrix was used to perform texture analysis because it is an important grayscale texture analysis method.

C. Edge feature

Edges in images consist of an important feature to represent their content. One way of representing such an important feature is to use a histogram. An edge histogram represents the frequency and directionality of brightness changes in the image. Edge histogram descriptor (EHD) is used to extract edge feature of an image. The EHD is basically the distributions of 5 types of edges in its each sub-image such as vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non-directional edges [24]. The sub-image is obtained by dividing the image space into 4x4 non-overlapping blocks. Then generate a histogram of edge distribution for each sub image. So the histogram for each sub-image represents the relative frequency of occurrences of the five types of edges in the corresponding sub-image. Each histogram bin values are normalized and quantized. For normalization, the number of edge occurrences for each bin is divided by total number of image blocks in the sub-image.

Each local histogram contains 5 bins where each bin corresponds to one of 5 edge types. Since there are 16 sub-images in the image, a total of 5x16=80 histogram bins are required. Note that each of the 80-histogram bins has its own semantics in terms of edge type. The semantics of the one dimension histogram bins starts from the sub-image at (0, 0) and ending at (3, 3), 16 sub-images are visited in the raster scan order and corresponding local histogram bins are arranged in single array. Within each sub-image, the edge types are arranged in the following order: vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non-directional. To extract both directional non-directional edge features, there is need to define a small square image-block. That is, divide an image space into non-overlapping square blocks and then we can extract edge information from each block. Note that, regardless of the image size, divide the sub-image into a fixed number of image-blocks.

D. Shape Feature

Shape feature is obtained by extracting the sketch/outer border of the image queried by the user using Edge detection or Border extraction. It compares the extracted border with the features of the images in the database to get the optimum results Image block.

Shape is defined as the characteristic surface configuration of an object; an outline or contour. It permits an object to be distinguished from its surroundings by its outline. Shape representations can be generally divided into two categories:

1) Boundary-based

Boundary-based shape representation uses the outer boundary of the shape. This is done by describing the considered region using its external characteristics; i.e., the pixels along the object boundary.

2) Region-based

Region-based shape representation uses the entire shape region by describing the considered region using its internal characteristics; i.e., the pixels contained in that region.

There are various boundary-based shape characterizations which perform the task of identifying an image by attempting to identify the boundary of a particular image. Information such as color and texture often help in the identification of the boundaries as well. In our system Region based moment matching technique is used for shape feature extraction. Proposed method uses these features to represent properties of an image.

E. Similarity Matching

The similarity between two images is defined by this step. Selection of similarity metrics is very important as it directly affects performance of content-based image retrieval. The kind of feature vectors selected determines the kind of measurement that will be used to compare their similarity. If the features extracted from the images are presented as multi-dimensional points, the distances between corresponding multi-dimensional points can be calculated. Euclidean distance is the very common metric used to measure the distance between two points in multi-dimensional space. For other kinds of features such as color histogram, Euclidean
distance may not be an ideal similarity metric or may not be compatible with the human perceived similarity. A number of other metrics, such as Mahalanobis distance, City block distance, Canberra distance, Minkowski distance, Earth Mover's distance, and Proportional Transportation distance, have been proposed for specific purposes. Initially three features of a query image i.e. color, texture, edge are extracted by their respective methods. Same features of the database images are extracted and stored. Proposed method uses Euclidian distance for the similarity measurement between query image and database images.

**F. Model approach**

This is the novel approach which can be explained in following steps:

1) Extract the image features of the database image.
2) Retrieve the images based on individual features by comparing.
3) Store all the results obtained from above step in single matrix.
4) Perform zigzag scanning and this will give 3 columned tables. In which first Column represent image, second represent number of time it occurs during zigzag scanning, third column represent weights.
5) Then arrange weight column of table in descending order. Corresponding row of image number and frequency should be shifted according to weight column.
6) Return the images with smallest lower bound distances.

**Fig. 2: Block Diagram of Model Approach**

**V. RESULTS AND DISCUSSIONS**

In this paper for experimental dataset contains 1000 images, divided into 10 categories, each category has 100 images. Experimental images covers variety of content, including people, beach, buildings, buses, dinosaurs, elephant, flower, horses, mountains, food.

Precision and recall rate are used to evaluate the performance of above approach. Proposed method selected top 10 images. And calculated precision and recall by using following formula.

A. Database

The database used in evaluation is SIMPLICITY database. The SIMPLICITY database is a subset of the Corel database of 1000 images, which have been manually selected to be a database of 10 classes of 100 images each. The images are subdivided into 10 classes. The images are of size 384 * 256 or 256 * 384 pixels.

This database was extensively used to test many CBIR systems [17, 29] because the size of the database and the availability of class information allows for performance evaluation as can be seen in the following sections.
B. Implementation Environment

The image retrieval system is implemented using MATLAB image processing tools and statistical tools. During the implementation, we use a platform of Intel Core 2 Processing power of 1.73 GHz CPU with 4GB RAM. 1000 image database went through image feature extraction.

C. Performance Evaluation Metrics of CBIR

The level of retrieval accuracy achieved by a system is important to establish its performance. If the outcome is satisfactory and promising, it can be used as a standard in future research works. In CBIR, precision-recall is the most widely used measurement method to evaluate the retrieval accuracy. Some recent literature uses this pair to measure the retrieval performance. Precision P, is defined as the ratio of the number of retrieved relevant images to the total number of retrieved images.

\[
P = \frac{\text{Number of relevant images retrieved}}{\text{Total number of retrieved images}}
\]

Let the number of all retrieved images be \( n \), and let \( r \) be the number of relevant images according to the query then the precision value is: \( P = \frac{r}{n} \). Precision \( P \) measures the accuracy of the retrieval.

Recall \( R \), is defined as the ratio of the number of retrieved relevant images to the total number of relevant images in the whole database.

\[
R = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}}
\]

Let \( r \) be the number of relevant images according to the query, and \( M \) be the number of all relevant images to the query in the whole database then recall value is: \( R = \frac{r}{M} \).

For comparison initially we found results based on single feature. And then we combine all four features and evaluated precision and recall values.

D. Results using Model approach

<table>
<thead>
<tr>
<th>Category</th>
<th>Precision Value (%)</th>
<th>Recall Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa People</td>
<td>70</td>
<td>7</td>
</tr>
<tr>
<td>Beach</td>
<td>70</td>
<td>7</td>
</tr>
<tr>
<td>Buildings</td>
<td>70</td>
<td>7</td>
</tr>
<tr>
<td>Buses</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>Dinosaurs</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>Elephant</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>Flower</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>House</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>Mountains</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>Food</td>
<td>80</td>
<td>8</td>
</tr>
<tr>
<td>Average</td>
<td>870</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Fig. 3: Horse image as a query

Fig. 4: Bus image as a query
VI. CONCLUSION

Proposed method have designed and implemented a content-based image retrieval system that evaluates the similarity of each image in its data store to a query image in terms of texture, edge, shape and color characteristics, and returns the images within a desired range of similarity. For the color content extraction mean and standard deviation of both the query image and database images is considered. Mean provide color in that image where standard deviation gives variation in color from that particular image. From among the existing approaches to texture analysis within the domain of image processing, statistical approach to extract texture feature. Gray level co-occurrence matrix is used to extract texture of an image. It represents the probability between two pixels in an image. Then we calculated entropy which gives a measure of complexity of the image. Human eye is sensitive to edge features for image perception. Proposed method adopts edge histogram descriptor to represent edge in the image. Edge histogram of the image space represents the frequency and directionality of brightness changes in the image. Shape defines characteristic surface configuration of an object; an outline or contour. Sketch/outer border is extracted using edge detection or border extraction After extracting all features model approach is adopted and relevant images can be obtained depending upon their weights. Better results can be obtained by using multi feature combining method and model approach.

Performance is evaluated by Calculating Precision and Recall parameter for set of images and it is giving much better results compared to other existing methods.

REFERENCES


