Fast Multi-Feature Image Retrieval using Dominant Colour, Logical Edge and Binarization Techniques

Monika Rathore¹, Prof. Ravindra Gupta²

¹M.Tech Research Scholar, ²Research Guide, Department of Computer Science and Engineering, RKDF Institute of Science and Technology, Bhopal, India

Abstract - Modern web era has lots of multimedia contents to share on social networking websites, emails, video and photo sharing portals, chat tools, and these are going faster and faster day by day due to easy availability of internet and high quality digital cameras in every hand. As a result the servers have a large collection of these multimedia contents especially images and to extract out the desired image from such large collection is tough and complex task. Everyone in this world has different need and desire to search images for their need. In this paper a fast multi-feature image retrieval technique is proposed which is based on dominant colour method for colour feature extraction, Logical Edge for texture extraction and binarization for shape feature extraction.

Keywords-- Image Retrieval, Dominant Colour, Logical Edge Method, Binarization and Texture Feature.

I. INTRODUCTION

The search for similar images in large-scale image databases has been an active research area in the last couple of years. A very promising approach is content based image retrieval (CBIR). In such systems, images are typically represented by approximations of their contents. Typical approximations consist of statistics, and Fourier or wavelet transformations of the raw image data. This so-called feature extraction aims at extracting information that is semantically meaningful but needs a small amount of storage.

In the past decade, more and more information has been published in computer readable formats. In the meanwhile, much of the information in older books, journals and newspapers has been digitized and made computer readable. Big archives of films, music, images, satellite pictures, books, newspapers, and magazines have been made accessible for computer users. Internet makes it possible for the human to access this huge amount of information. The greatest challenge of the World Wide Web is that the more information available about a given topic, the more difficult it is to locate accurate and relevant information. Most users know what information they need, but are unsure where to find it. Search engines can facilitate the ability of users to locate such relevant information.

Content based image retrieval for general-purpose image databases is a highly challenging problem because of the large size of the database, the difficulty of understanding images, both by people and computers, the difficulty of formulating a query, and the issue of evaluating results properly. A number of general-purpose image search engines have been developed.

Content-based retrieval uses the contents of images to represent and access the images. A typical content-based retrieval system is divided into off-line feature extraction and online image retrieval. A conceptual framework for content-based image retrieval is illustrated in Fig. 1.1 [4].

Fig. 1.1 A Conceptual Framework for Image Retrieval System

In off-line stage, the system automatically extracts visual attributes (color, shape, texture, and spatial information) of each image in the database based on its pixel values and stores them in a different database within the system called a feature database. The feature data (also known as image signature) for each of the visual attributes of each image is very much smaller in size compared to the image data, thus the feature database contains an abstraction (compact form) of the images in the image database.
One advantage of a signature over the original pixel values is the significant compression of image representation. However, a more important reason for using the signature is to gain an improved correlation between image representation and visual semantics [4].

II. IMAGE RETRIEVAL SYSTEM

In on-line image retrieval, the user can submit a query example to the retrieval system in search of desired images. The system represents this example with a feature vector. The distances (i.e., similarities) between the feature vectors of the query example and those of the media in the feature database are then computed and ranked. Retrieval is conducted by applying an indexing scheme to provide an efficient way of searching the image database. Finally, the system ranks the search results and then returns the results that are most similar to the query examples. If the user is not satisfied with the search results, he can provide relevance feedback to the retrieval system, which contains a mechanism to learn the user’s information needs.

Image Retrieval Problem:

In this computer age, virtually all spheres of human life including commerce, government, academics, hospitals, crime prevention, surveillance, engineering, architecture, journalism, fashion and graphic design, and historical research use images for efficient services. A large collection of images is referred to as image database. An image database is a system where image data are integrated and stored [1].

The police maintain image database of criminals, crime scenes, and stolen items. In the medical profession, X-rays and scanned image database are kept for diagnosis, monitoring, and research purposes. In architectural and engineering design, image database exists for design projects, finished projects, and machine parts. In publishing and advertising, journalists create image databases for various events and activities such as sports, buildings, personalities, national and international events, and product advertisements. In historical research, image databases are created for archives in areas that include arts, sociology, and medicine. In a small collection of images, simple browsing can identify an image. This is not the case for large and varied collection of images, where the user encounters the image retrieval problem. An image retrieval problem is the problem encountered when searching and retrieving images that are relevant to a user’s request from a database. To solve this problem, text-based and content-based are the two techniques adopted for search and retrieval in an image database.

Feature Extraction:

Feature extraction is a means of extracting compact but semantically valuable information from images. This information is used as a signature for the image. Similar images should have similar signatures. Representation of images needs to consider which features are most useful for representing the contents of images and which approaches can effectively code the attributes of the images. Feature 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, shape, and color, which are used most often to extract the features of an image.

a. Colour Features:

One of the most important features visually recognized by humans in images is color. Humans tend to distinguish images based mostly on color features. Because of this, color features are the most widely used in CBIR systems and the most studied in literature. Color is a powerful descriptor that simplifies object identification, and is one of the most frequently used visual features for content-based image retrieval. To extract the color features from the content of an image, a proper color space and an effective color descriptor have to be determined.

b. Texture Features

In the field of computer vision and image processing, there is no clear-cut definition of texture. This is because 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, of which the addition of noise to the patterns and their repetition frequencies results in textures that can appear to be random and unstructured. 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, for example, regularity, directionality, smoothness, and coarseness.

c. Shape Features

One of the common used features in CBIR systems is the shape. Shape of an object is the characteristic surface configuration as represented by the outline or contour. Shape recognition is one of the modes through which human perception of the environment is executed. It is important in CBIR because it corresponds to region of interests in images. Shape feature representations are categorized according to the techniques used. They are boundary-based and region-based.
Applications of Image Retrieval Systems:

Image retrieval based on content is extremely useful in a plethora of applications such as publishing and advertising, historical research, fashion and graphic design, architectural and engineering design, crime prevention, medical diagnosis, geographical information and remote sensing systems, etc. [5]. A typical image retrieval application example is a design engineer who needs to search his organization database for design projects similar to that required by his clients, or the police seeking to confirm the face of a suspected criminal among faces in the database of renowned criminals. In the commerce department, before trademark is finally approved for use, there is need to find out if such or similar ones ever existed. In hospitals, some ailments require the medical practitioner to search and review similar X-rays or scanned images of a patient before proffering a solution.

The most important application, however, is the Web, as big fraction of it is devoted to images, and searching for a specific image is indeed a daunting task.

Numerous commercial and experimental CBIR systems are now available, and many web search engines are now equipped with CBIR facilities, as for example Alta Vista, Yahoo and Google [6].

III. PROPOSED METHODOLOGY

Image retrieval is research topic famous among researcher since the internet is the primary source of information and everybody go there and start searching images, videos and contents for his/her use.

In the same context the image retrieval system is come under spotlight to search desired images from huge collection of images over internet. In this section a proposed methodology is given to extract the images from large set of images having different features like colours, shapes and textures. The whole system is divided into two models feature database preparation and image retrieval process.

**Module I: Feature Database Preparation Process**

The block diagram feature extraction process of proposed methodology is shown in Fig. 3.1. The main blocks are loading images from the directory so that the features can be extract out all of them. Then the database of features also need to be loaded to save features of new images in the collection.

The Pre-processing of images is then performed which is resize of images to make all the records similar and median filtering to remove small noises. After this extract the features (texture, colour and shape) of images and store them into the database.
Fig. 3.3 Flow Chart of Features Extraction Process of Proposed Methodology

Module II: Image Retrieval Process

The second section of the image retrieval system is shown in Fig. 3.2, where query image is selected to retrieve the images from the collection of images, the query images is first pre-processed and all the features are extracted to compare with the features of images. During the features comparison a similarity table is prepared as per the features of query image, similarity table is then sort with highest to lowest similarity which is than top 16 results are fetched from directory of images and displayed.

The whole system is simulated and the execution of algorithm is performed and the flowchart is shown in the Fig. 3.3. In the start of flow chart the system loads all the images from directory to extract out the features from it and the database of features also need to be loaded to store the features of new images in the database. If the system is initializing first time the database is created itself. The next process is to extract the features of texture, colour and shape and after done for all the images in directory store it in the database.

The next module is simulated and the flowchart of the this module is shown in the Fig. 3.4. First step is to choose the query image is to search out the similar images of from the collection of images. Then all the features of the query image is then extracted and all the features from the database is also loaded to compare with the features of the query image. After comparison a similarity table is being prepared. After that the similarity table is sorted for highest to lowest features and the top 16 results are displayed on the screen.

Fig. 3.4 Flow Chart of Retrieval Process of Proposed Methodology

IV. SIMULATION RESULTS

The simulation of proposed image retrieval system is explained in the previous section with the details of the execution algorithm. The results of the query images for different feature is shown in the below figure.

In Fig. 4.1, 4.2 and 4.3 the results shows the retrieved images having shape and colours and in Fig. 4.4 the results of texture features are shown.
The colour feature search the colour similarity with the records of images features collection and gives the best matched results, similarly texture feature is also compared by using the edge extraction technique than the edges of all the images in collection is compared with edges and shapes geometry compared with the images collection's geometry and gives best results as per all the features comparison and this can be achieved by sorting the colour, texture and shape features similarity in descending order.

V. CONCLUSION AND FUTURE SCOPE

The image retrieval and its proposed methodology is explained in the previous sections the simulation results of the proposed approach are also explained in the previous section.
From the results it is clear that the proposed approach is giving efficient results for images having all the three kinds of features for example texture features which is having specific geometry in a repeated manner and the colour features which is having visually different colour combinations mainly red, green and bluish shades and the last one shape features which has a object of circular, spherical or of a random shape.

The advancement in the image retrieval system can be done with the make searching algorithm less complex or by making system hardware faster retrieval speed will also be faster.

REFERENCES

AUTHOR’S PROFILE

Monika Rathore is a research scholar at RKDF Institute of Science and Technology under Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal. She is pursuing her Master of Technology (M. Tech.) in Computer Science and Engineering with research domain in image processing. She is currently working on image retrieval methodologies to develop highly efficient image search algorithm in modern web applications.