International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 12 200 – 205

Analysis of Image Retrieval using Hybrid Texture and Shape based Feature Vectors

Apoorva Mathur
Masters of Technology, Student
MEC, Bikaner
Apoorva.mathur12@gmail.com

Sunita Chaudhary
Department of Computer Science and Engineering, Guide
Jagannath University, Jaipur
er.sunita03@gmail.com

Dr. Meenu dave
Department of Computer Science and IT
Jagan Nath University, Jaipur
meenu.s.dave@gmail.com

Dr. Amit Sanghi
Department of Computer Science and Engineering
Marudhar Engineering College, Bikaner
dr.amitsanghi@gmail.com

Abstract- Content based image retrieval system is required so as to improve the retrieval of images from large web based and other various types of databases. With the increased importance of digital data, the production of image and video information has resulted in a large amount of images and videos that needs to be properly indexed for retrieval in the future. Various databases contain a large number of images, videos etc for which Content-Base retrieval systems or applications are greatly needed. To date, a number of methods and algorithms have been proposed by researchers from around the world to meet this challenge. This paper presents a literature review of such researches and discussions of a few existing ones.

*****

I. INTRODUCTION

In last few years, the amount of digital data (audiovisual information) has grown exponentially. Gigabytes of new images, audio and video clips are generated and stored every day, which builds up a huge, distributed and mostly unstructured storage of multimedia information, much of which can be accessed and retrieved using the internet. Subsequent retrieval of the digital information however, is not as easy and it may need a lot of additional work so as to be effective and efficient.

Mostly we use these three basic ways to retrieve previously stored digital data are Free browsing, Text-based retrieval and Content-based retrieval.

In free browsing, users search from the collection of images and videos till they get what they are looking for. In text-based retrieval, the textual information added to the data while storing it, is used to guide the requested query and search engines to find the matching data. In content-based retrieval, the search engine converts the requested multimedia information in some way as to query the database' and retrieve the results which are more likely to satisfy the requested query.

These ways have several limitations which includes wastage of considerable amount of time and effort required for manually annotating and storing every file and the imprecision related to the human perception of the contents being annotated. These problems increase when the size of the database gets bigger.

Many research groups are constinously working towards the ultimate goal of making it possible for the users to retrieve the desired multimedia data among massive amounts of storage in an efficient, meaningful, user friendly, and location-independent manner and the search should also be faster.

II. LITERATURE REVIEW

This section includes the work done by various researchers in the area of digital image retrieval.

F.A. Andaló et.al [1] proposed a study in which he used shape based features. In their paper, they exploited the concept for binary images and proposed a shape salience detector and a shape descriptor method as Tensor Scale Descriptor with Influence Zones. They used a robust method which was a graph-based approach called as the Image Foresting Transform to compute the tensor scale. Experimental results were provided, showing the effectiveness of the proposed methods.

H.-W. Yoo et al. [2] suggests two kinds of indexing keys to prune away irrelevant images to a given query image. These two keys are: major colors' set (MCS) signature and distribution block signature (DBS). The first one is related with color information and latter is related with spatial information. After applying these filters successively to a large database, a small amount of high potential candidates were retrieved and they looked somewhat similar to the query image. Then the quad modeling (QM) method was used to set the initial weights of two-dimensional cell in a query image according to each major color. Finally, system retrieves more similar images from the database by comparing the query image and the candidate images by using a similarity measuring function associated with the weights. In that procedure, they also use a relevance feedback mechanism.

Pattanaik , Bhalke [3] proposed a work and prove that Image retrieval system has overcome all the limitation of Semantic Text based Image Retrieval by considering the features of image. Multiple Features like histogram, color structure, color mean, descriptor texture is taken into

consideration so as to extract similar images from the database.

Sajjanhar and Lu [4] proposed a grid-based method. The shape is mapped onto a grid of fixed cell size, and is justified to the top left corner. A value of 1 is assigned to the cells of the grid which are either partially or fully covered by the shape, and the value 0 is assigned to the cells which are not covered by the shape, obtaining a sequence of 0's or 1's. This sequence can be used for shape representation.

In [5], the boundary of a shape is described by a turning function Elmo. The angle of the counter-clockwise tangent is measured as a function of the arc-length s, which is measured in front of a given reference start point 0 on the polygon's boundary. However, this representation is sensitive to small variations of shape.

In [6] K L Tan et al., proposes the centroid radii model so as to estimate the shapes of objects in images. To calculate the shape's descriptor, lengths of a shape's radii from its centroid at regular intervals are captured. If 0 be the regular between radii then, the number of intervals is given by k =360/0. All radii lengths are normalized by dividing with the longest radius length. Thus, the shape descriptor can be represented as a vector radius from the centroid to the boundary of the shape.

III. EXISTING IMAGE RETRIEVAL SYSTEMS

Several retrieval systems are available as commercial packages and their demonstration versions available on the Web. In all that is available, few of the most prominent are described here as:

QBIC : The QBIC system created by IBM is all time famous between all the content based image retrieval systems. It offers retrieval by many combinations of basic features as well as search by text keyword. At search time, the system checks and matches the appropriate features from query and stored images, calculates a similarity value (feature Vector in numerical value) between the query and each images stored in database examined, and the system displays most similar images on the screen.

Virage : VIR Image Engine from Virage, Inc is a well-known commercial system. It is flexible and builds in new types of query interfaces and has an additional customized module to process various classes of images. It is also available for the existing systems such as Oracle or Informix. AltaVista's AV Photo Finder is a high-profile application of Virage technology which allows Web surfers to search for images by content similarity.

Excalibur : Another similar philosophy has been adopted by, a company with a long history of successful database applications, for their Visual Retrieval Ware product known as Excalibur

Technologies [7]. This application offers a various image indexing and matching techniques based on their own proprietary pattern recognition technology.

A few systems have been developed, mainly by academic institutions as experiments, in order to demonstrate the feasibility of new techniques. Some of these are:

Photobook : The Photobook system [8] was created by Massachusetts Institute of Technology (MIT). Unlike other systems, it aims to calculate information preserving features, from which all essential aspects of the original image can in theoretically be reconstructed. This allows features to be computed at search time which are relevant to a particular type of search, giving more flexibility at the expense of speed.

Chabot : This system has provided a combination of text-based and color-based feature vectors developed by California's Department of Water Resources. It is now called Cypress, and incorporated within the Berkeley Digital Library project at the University of California at Berkeley (UCB).

VisualSEEk: The Visual SEEk system is developed by Columbia University. This application offers searching by image region colour, and spatial location, with keyword. Users can build up image queries by defining areas of specified spatial and colour with relative locations within the image.

MARS : The MARS application developed by the University of Illinois has a feedback based module as this is felt to be the only way at present of capturing individual human similarity judgments. The application characterizes every item within an image by a variety of features, and uses a range of different similarity measures to compare query and stored objects.

IV. MODIFIED APPROACH

Fig. 1 shows the architecture of image retrieval system. Core of the system is the image feature database. Image feature extraction can be obtained either from the image itself or can be obtained through user interaction, and is used for computing the similarity between images. The architecture of our system must have two functional components. The first component is the visual content extraction and calculation of the feature vector. Query engine is the second component which consists of a query user interface and a query processing subcomponent.

ISSN: 2321-8169
200 – 205

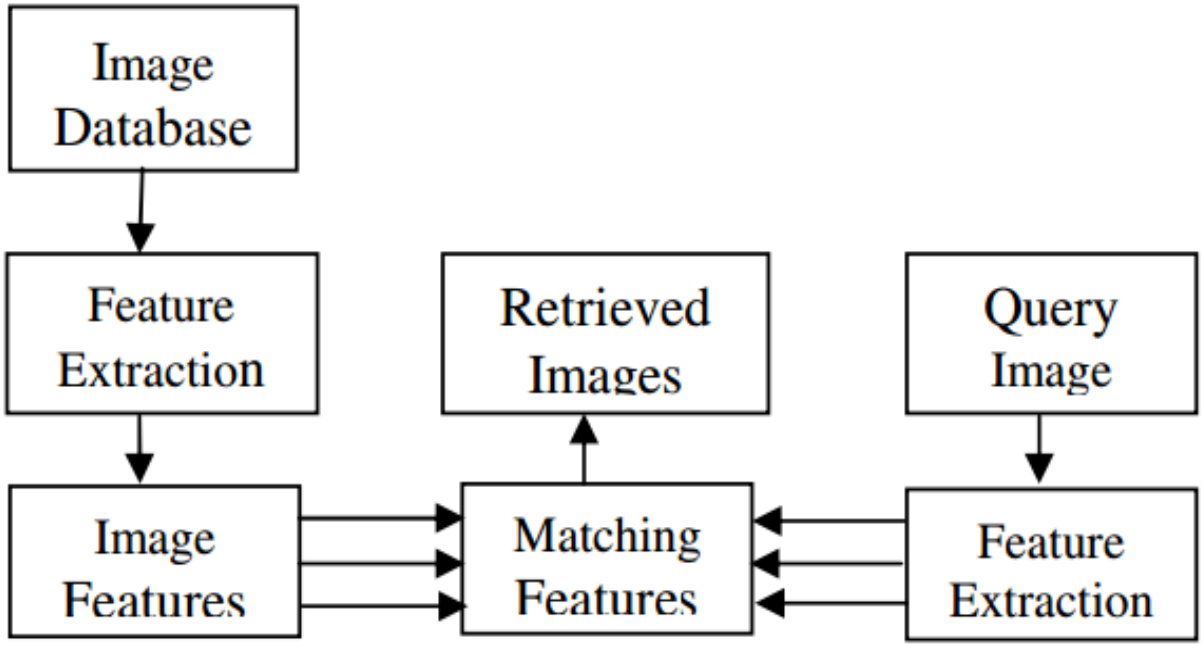

Fig. 1. Proposed Model of Image Retrieval System

The process has the mainly five steps. First, submit the query information by the user. Second, generate the image's feature vector according to the user query information. Third, find the feature similarity value according to the feature vectors. Fourth, find the similar drawings according to the characteristics of similarity values. Fifth, output the detection result of the similar drawings. The similar feature values are obtained by comparing the original feature vectors by the original drawings generated and the image features. The content based image retrieval is based on the image's contents itself which are extracted from the image color, texture, shape, spatial relations and so on. The information is inputted to the image's database, when the user input the query information, according to the image characteristic values, characteristic vectors, find the most similar feature vectors from the original information, then find the similar drawings, at last, output results.

IV. REFERENCES

[1] Andaló, Fernanda A., et al. "Shape feature extraction and description based on tensor scale." *Pattern Recognition* 43.1: 26-36, 2010.

[2] Yoo, Donggeun, et al. "Intra-class key feature weighting method for vocabulary tree based image retrieval." Ubiquitous Robots and Ambient Intelligence (URAI), 2012 9th International Conference on. IEEE, 2012.

[3] Pattanaik, S., & Bhalke, D. G., "Efficient Content based Image Retrieval System using Mpeg-7 Features." International Journal of Computer Applications, 53 (5), 19-24, 2012.

[4] Sajjanhar, Atul, and Guojun Lu. "A grid-based shape indexing and retrieval method." Australian Computer Journal 29.4: 131-140.

[5] Pun, Chi-Man, and Cong Lin. "Shape Classification Using Contour Simplification and Tangent Function." International Journal of Circuits, Systems and Signal Processing 1 : 17-24, 2010.

[6] K-L Tan, B. C. Ooi, & L. F. Thiang , "Retrieving similar shapes effectively and efficiently", Multimedia Tools and Applications, vol. 19, pp. 111-134, 2003

[7] Deb, Sagarmay, and Yanchun Zhang. "An overview of content-based image retrieval techniques." Advanced Information Networking and Applications, 2004. AINA 2004. 18th International Conference on. Vol. 1. IEEE, 2004

[8] Shih, Timothy K., et al. "An intelligent content-based image retrieval system based on color, shape and spatial relations." proceedings-national science council republic of china part a physical science and engineering 25.4: 232-243, 2001.