An Extensive Review on Image Retrieval System

Monika Rathore[1], Prof. Ravindra Gupta[2]
[1]M-Tech Research Scholar, [2]Research Guide, Department of Computer Science and Engineering
RKDF Institute of Science and Technology, Bhopal

Abstract- An image is an artistic work that employs a number of smaller images combined together to form another large image. Each building block of image, that has its own distinctive and meaningful content. Many researchers have made huge efforts in recent years. The study of image retrieval systems, but no competent algorithms are accepted broadly that can extract image features similar to human visions. This is constantly the challenge to all researchers. Compared to the researches various methodologies have been studied for example Content-based Image Retrieval systems, the difference can be analyzed in the recent decade of studies. In this review papers a region-based image retrieval system that uses color and textures visual features to describe the content of an image region has been studied in order to contribute better and efficient image retrieval system. Our contribution is to reach and get better literature review.

Keywords - Image Retrieval, Image Mining, Texture, Shape and Content.

I. INTRODUCTION

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 has been digitized and made computer readable. Big archives of films, music, images, satellite pictures, books and magazines have been made accessible for computer users. Internet world 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 may facilitate the ability of users to locate such relevant information. In this computer age, virtually all spheres of human life including commerce, government, academics, crime prevention, surveillance, engineering, journalism, fashion and graphic design, and historical research use images for efficient services. A huge collection of images is referred to as image database. An image record is a system where image data are integrated and stored [1]. Image data contain the raw images and information extracted from images by automated or computer assisted image analysis. The police maintain image database of criminals and stolen items. In the medical profession, X-rays and scanned image database are kept for diagnosis, monitoring, and research works. In architectural and engineering design, image database exists for designing projects, finished projects, and machine parts. In publishing and advertising, journalists create image databases for a variety of events and activities such as sports, buildings, personalities, national and international events, and product advertisement. In historical research work, image databases are created for archives in areas Efficient Content Based Image Retrieval that include arts and medicine. In a small collection of images, simple browsing may identify an image. This is not the case for large and varied collection of images, where the users encounter 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.

Applications

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].

Region Based Image Retrieval

Early CBIR methods used global feature extraction to obtain the image descriptors. For example, QBIC [7] developed at the IBM Almaden Research Center extracts several features from each image, namely color, texture, and shape features. These descriptors are obtained globally by extracting information by means of color histograms for color features; global texture information on coarseness, contrast, and direction; and shape features about the curvature, moments invariants, circularity, and eccentricity. Similarly, the Photobook system [8], Visualseek [9], and VIR [10], use global features to represent image semantics.

II. System Model

The first part of CBIR system, this system defines the similarity between contents of two images based on global features (i.e., features extracted from the whole image). Texture is one of the crucial primitives in human vision and texture features have been used to identify contents of images. Moreover, texture can be used to describe contents of images, such as clouds, bricks, hair. Both identifying and describing contents of an image are strengthened when texture is integrated with color, hence the details of the important features of image objects for human vision can be provided. In this system.

Texture Similarity Measure

To test the similarity between a query image *Q* and a database images *I* based on their texture feature we proposed to use the Euclidian distance for its simplicity. The attributes of the texture features vector may have different ranges (one of very small value and one of very high value), therefore a normalization method should be applied to make all the texture features have the same effect in measuring image similarity. The Min-Max algorithm [31] is employed as a normalization technique.

Color Feature Extraction

In this system we used global color histograms in extracting the color features of images. The main issue regarding the use of color histograms for image retrieval involves the choice of color space, color space quantization into a number of color bins, where each bin represents a number of neighboring colors, and a similarity metric [61].

HSV Color Space

In the literature, there is no optimum color space known for image retrieval, however certain color spaces such as HSV, Lab, and Luv have been found to be well suited for the content based query by color. We adopt to use the HSV (Hue, Saturation, and Value) color space for its simple transform from the RGB (Red, Green, Blue) color space, in which all the existing image formats are represented.

Advantages and disadvantages of CBIR

- *Advantages*
  - Easy to implement
  - Fast retrieval
  - Web image search
- *Disadvantages*

Manual annotation is not always available Manual annotation is impossible for a large DB Manual annotation is not accurate A picture is worth a thousand words Surrounding text may not describe the image

CBIR Applications

- Commerce (fashion, catalogue)
- Biomedicine (X-ray, CT)
- Crime prevention (Security filtering)
- Cultural (Art galleries, museums)
- Military (Radar, Aerial
- Entertainment (Personal Album)

III. Literature Review

Jan-Ming Ho, Shu-Yu Lin and Chi-Wen [1] investigated the Image retrieval that has been popular for several years. There are different system designs for content based image retrieval system. This research work proposes novel system architecture for CBIR system which combines techniques include content-based image and color analysis, as well as data mining techniques. K-means clustering and bring in the neighborhood module to build the CBIR system. Concept of neighborhood color analysis module which also recognizes the side of every grids of image is first contributed in this paper. The result shows the CBIR system performs well in the training and it also indicates there contains many interested issue to be optimized in the query stage of image retrieval.

Wenhui Yang, Guiquan Liu and Lei Zhang [2] studies and analyzed the explosive growth of Internet image data, labeling image data for image retrieval has become an increasingly onerous task. To that end, authors proposed a novel multi-view learning with batch mode active learning

framework, MV-BMAL, for improving the performance of image retrieval. The color, texture and shape features are extracted and considered as un-correlated and sufficient views of an image, then each classifier is trained on these views respectively, and the schema makes full use of the classification results of each unlabeled samples to find out the most informative and representative samples for automatically or manually labeling.

Sergyan. S. [6] worked for Using content based image retrieval systems the images are compared on the grounds of the stored information of images. As simple descriptors the intensity (or color), the texture and the shape is widely used. In this work the effectively usable intensity descriptors are examined, and the two dimensional Cartesian moments - as a descriptor - proved the best choice using thermal images. These feature descriptors can be compared by several distance and similarity measures, and some special weights are necessary in the case of thermal images.

Hai Wang, Shuwu Zhang and Wei Liang [7] investigated the distance measurement is one of the key tasks in content-based image retrieval. This work proposed a new fractional distance metric for CBIR. Researchers conduct extensive experiments on three famous benchmark datasets, using different color, texture and shape features. The result shows that retrieval performance of the new distance metric consistently outperforms the more common City Block and Euclidean distance metrics, as well as several other commonly-used distance functions. The results on three commonly used benchmark datasets show that the new fractional distance metric can be used universally.

Sahu. M., Shrivastava, M. and Rizvi, M.A. [8] presented a research work for studying the Image data mining can be done manually by slicing and dicing the data until a pattern becomes obvious. Colour, texture and shape of an image have been primitive image descriptors in Content Based Image Retrieval (CBIR) system. A novel framework for texture information of an image and achieves higher retrieval efficiency than the shape features of an image. There is a trade-off between accuracy and computational cost. The trade-off decreases as more efficient algorithm is used to solve the problem and increases the computational power and will decreases the cost of the whole system as well.

Buch, A., Jain, S. and Pradhan, S.N., [9] have done research on the content based Image Retrieval systems have become a reliable tool for many image database applications. The goal of author is to develop a prototype of Content based Image Retrieval Systems for Satellite Images, wherein the user enters query in form of a sample image and the relevant similar images based on image content are displayed as an output. The Work presented a novel approach of satellite image search and retrieval system based on colour and texture. Texture information is obtained using co-occurrence matrix of the gray scaled images. The properties like correlation and homogeneity are good parameters to measure the textural property; these are evaluated at proper offset and angle. Colour is inevitable feature that dominates the human perception most.

Qirong Bo; Jinye Peng, [10] estimated the repaid development of internet technology, image documents have become an important information source. An interactive image recommendation system, which firstly uses color histogram feature or Gabor texture feature to express image contents, then a kernel based K-meanse is utilized to cluster images into multiple classes by their visual features, based on a sample and hyperbolic techniques, images are recommended and displayed. The results demonstrate that the proposed system may recommend and display the similar images from the same class efficiently.

IV. Problem Formulation

The proposed system is designed to operate the content based image retrieval system. Our Methodology may be given to extract the images from large set of images having different features like colors, shapes and textures. System architecture not only works well for image retrieval, but also improves its precision. In our knowledge, this study first describes the segmentation and grid module, feature extraction module, K-means clustering and neighborhood module to build the CBIR system. Even, the concept of neighborhood module which recognizes the side of every grids of image is first contributed in this study.

V. Conclusion

After having review we can propose new method that is given to extract the images from large set of images having different features like colours, shapes and textures. To overcome those drawbacks of text-based image retrieval, content-based images retrieval was introduced. With extracting the images features, CBIR perform well than other methods in searching, browsing and content mining etc. The need to extract useful information from the raw data becomes important and widely discussed. Furthermore, classification technique is usually introduced into CBIR to perform well

and easy retrieval. Although many research improve and discuss about those issues, still many difficulties hasn't been solved. The rapid growing images information and complex diversity has build up the bottle neck.

REFERENCES

[1]. Jan-Ming Ho; Shu-Yu Lin; Chi-Wen Fann; Yu-Chun Wang; Ray-I Chang, "A novel content based image retrieval system using K-means with feature extraction," *Systems and Informatics (ICSAI), 2012 International Conference on* , vol., no., pp.785,790, 19-20 May 2012.

[2]. Wenhui Yang; Guiquan Liu; Lei Zhang; Enhong Chen, "Multi-view learning with batch mode active selection for image retrieval," *Pattern Recognition (ICPR), 2012 21st International Conference on* , vol., no., pp.979,982, 11-15 Nov. 2012.

[3]. Bhagat, A.P.; Atique, M., "Design and development of systems for image segmentation and content based image retrieval," *Computational Intelligence and Signal Processing (CISP), 2012 2nd National Conference on* , vol., no., pp.109,113, 2-3 March 2012.

[4]. Gali, R.; Dewal, M.L.; Anand, R.S., "Genetic Algorithm for Content Based Image Retrieval," *Computational Intelligence, Communication Systems and Networks (CICSyN), 2012 Fourth International Conference on* , vol., no., pp.243,247, 24-26 July 2012.

[5]. Gurve, M.K.; Sarup, J., "Satellite cloud image processing and information retrieval system," *Information and Communication Technologies (WICT), 2012 World Congress on* , vol., no., pp.292,296, Oct. 30 2012-Nov. 2 2012.

[6]. Sergyan, S., "Useful and effective feature descriptors in content-based image retrieval of thermal images," *Logistics and Industrial Informatics (LINDI), 2012 4th IEEE International Symposium on* , vol., no., pp.55,58, 5-7 Sept. 2012.

[7]. Hai Wang; Shuwu Zhang; Wei Liang; Fangyuan Wang; Yingying Yao, "Content-based image retrieval using fractional distance metric," *Image Analysis and Signal Processing (IASP), 2012 International Conference on* , vol., no., pp.1,5, 9-11 Nov. 2012.

[8]. Sahu, M.; Shrivastava, M.; Rizvi, M.A., "Image Mining: A New Approach for Data Mining Based on Texture," *Computer and Communication Technology (ICCCT), 2012 Third International Conference on* , vol., no., pp.7,9, 23-25 Nov. 2012.

[9]. Buch, A.; Jain, S.; Pradhan, S.N., "Exploring a new direction in colour and texture based satellite image search and retrieval system," *Engineering (NUiCONE), 2011 Nirma University International Conference on* , vol., no., pp.1,5, 8-10 Dec. 2011.

[10]. Qirong Bo; Jinye Peng, "A Novel Interactive Image Recommendation System," *Information Technology and Applications (IFITA), 2010 International Forum on* , vol.2, no., pp.248,251, 16-18 July 2010.

[11] H. Tamura, and N. Yokoya, "Image Database Systems: A Survey," Pattern Recognition, vol. 17, no 1, pp.29–49, Sep. 1984.

[12] S. Gerard, C. Buckely, "Term-Weighting Approaches in Automatic Text Retrieval," Information Processing and Management, vol. 24, no.5, pp. 513-523, Jan. 1988.

[13] Y. Chen, J. Wang, "Image Categorization by Learning and Reasoning with Regions," Journal of Machine Learning Research, vol. 5, pp. 913–939, May 2004.

[14] F. Long, H. Zhang, H. Dagan, and D. Feng, "Fundamentals of content based image retrieval," in D. Feng, W. Siu, H. Zhang (Eds.): "Multimedia Information Retrieval and Management. Technological Fundamentals and Applications," Multimedia Signal Processing Book, Chapter 1, Springer-Verlag, Berlin Heidelberg New York, 2003, pp.1-26.

[15] V. Gudivada and V. Raghavan, "Content-based image retrieval systems," IEEE Computer, vol. 28, no 9, pp18-22, Sep. 1995.

[16] M. Kherfi, D. Ziou, and A. Bernardi, " Image Retrieval From the World Wide Web: Issues, Techniques, and Systems," ACM Computing Surveys, vol. 36, no. 1, pp. 35–67, March 2004.

[17] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, and P. Yanker, "Query by image and video content: The QBIC system," IEEE Computer, vol. 28, no 9, pp.23-32, Sep. 1995.

[18] A. Pentland, R. Picard, and S. Sclaroff , "Photobook: Content based manipulation of image databases," International Journal of Computer Vision, vol.18, no 3, pp.233–254, June 1997.

[19] J. Smith and S. Chang, "Visualseek: A Fully Automated Content-Based Image Query System," Proceedings of the 4th ACM international conference on Multimedia table of contents, Boston, Massachusetts, United States, Nov. 1996, pp. 87-98.

[20] A. Gupta, and R. Jain, "Visual information retrieval," Comm. Assoc. Comp. Mach., vol. 40, no. 5, pp. 70–79, May. 1997.

[21] J. Li, J. Wang, and G. Wiederhold, "Integrated Region Matching for Image Retrieval," In Proceedings of the 2000 ACM Multimedia Conference, Los Angeles, October 2000, pp. 147-156.

[22] Y. Deng, B. Manjunath, "Unsupervised Segmentation of Color -Texture Regions in Images and Video," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 8, pp. 800-810, Aug. 2001.

[23] A. Smeulders, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349- 1380, May. 2000.

[24] J. Caicedo, F. Gonzalez, E. Romero, E. triana, "Design of a Medical Image Database with Content-Based Retrieval Capabilities," In Proceedings of the 2nd Pacific Rim conference on Advances in image and video technology, Santiago, Chile, December 17-19, 2007.

[25] B. Manjunath and W. Ma, "Texture features for Browsing and retrieval of image data," IEEE transactions on pattern analysis and machine intelligence, vol. 18. No. 8, pp. 837-842, August 1996.