International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Survey over the Content-Based Image Retrieval Techniques

Suhasini Gunjal[1], S. M. Rokade[2]

[1, 2]Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India

Abstract: ***The content based image retrieval (CBIR) is one of the most popular, rising research areas of the digital image processing. Most of the available image search tools, such as Google Images and Yahoo! Image search, are based on textual annotation of images. In these tools, images are manually annotated with keywords and then retrieved using text-based search methods. The performances of these systems are not satisfactory. The goal of CBIR is to extract visual content of an image automatically, like color, texture, or shape. The CBIR technology can be used in several applications such as digital libraries, crime prevention, and photo sharing sites. Such a system has great value in apprehending suspects and identifying victims in forensics and law enforcement. A possible application is matching a forensic sketch to a gallery of mug shot images. The area of retrieve images based on the visual content of the query picture intensified recently, which demands on the quite wide methodology spectrum on the area of the image processing. In this paper, we will propose a more comprehensive survey on content based image retrieval techniques.***

Keywords: CBIR, image processing, textual annotation, image processing

1. Introduction

As processors become increasingly powerful, and memories become increasingly cheaper, the deployment of large image databases for a variety of applications have now become realizable. Databases of art works, 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 [1].

Effectively accessing desired images from large and varied image databases is now a necessity. In typical CBIR system the visual content of the images in the database are extracted and described by multidimensional future factors. Colors are widely used for image retrieval in an Image Retrieval System.

"Content-based" means that the search will analyze the actual contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. Here the 'content' refers to colors and textures information that can be derived from the image itself. CBIR is desirable because most web based image search engines rely purely on metadata and this produces a lot of false detection in the results [2], [3]. Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Therefore, in this paper we proposed efficient content based image retrieval method using color and texture feature to improve above mentioned problems.

2. Literature Survey

Jaiswal, Kaul [4] concluded that content based image retrieval is not a replacement of, but rather a complementary component to text based image retrieval. Only the integration of the two can result in satisfactory retrieval performance. In this paper they reviewed the main components of a content based image retrieval system, including image feature representation, indexing, and system design, while highlighting the past and current technical achievement.

Ivan Lee, et.al. (1996) [5] have present the analysis of the CBIR system with the human controlled and the machine controlled relevance feedback, over different network topologies including centralized, clustered, and distributed content search. In their experiment for the interactive relevance feed backusing RBF, they observe a higher retrieval precision by introducing the semi-supervision to the non-linear Gaussian-shaped RBF relevance feedback.

Verma, Mahajan, (2012) [6] have used canny and sobel edge detection algorithm for extracting the shape features for the images. After extracting the shape feature, the classified images are indexed and labeled for making easy for applying retrieval algorithm in order to retrieve the relevant images from the database. In their work, retrieval of the images from the huge image database as required by the user can get perfectly by using canny edge detection technique according to results.

Ryszard S. Chora´s (2007) [7] contributes their work for the identification of the problems existing in CBIR and Biometrics systems describing image content and image feature extraction. They have described a possible approach to mapping image content onto low-level features. Their paper investigated the use of a number of different color , texture and shape features for image retrieval in CBIR and Biometrics systems.

Pattanaik , Bhalke (2012) [8] has worked to prove that Content Based Image Retrieval has overcome all the limitation of Text Based Image Retrieval by considering the contents or features of image. A query image can be retrieved efficiently from a large database. A Database consists of different types of images has implemented on the system. Different Features such as histogram, color mean, Color structure descriptor texture is taken into consideration for extracting similar images from the database. From the experimental result it is seen that combined features can give

better performance than the single feature. So selection of feature is one of the important issues in the image retrieval. The system is said to be efficient if semantic gap is minimum .The result can be improved in future by introducing feedback and user ‟s choice in the system.

Zhao,Grosky (2002) [9] view that bridging the semantic gap between the low-level features and the high-level semantics is within the interface between the user and the system, other research direction is towards improving aspects of CBIR systems by finding the latent correlation between low-level visual features and high-level semantics and integrating them into a unified vector space model.

Peter Stanchev, et.al. [10] Proposed that several visual descriptors exist for representing the physical content of images, for instance color histograms, textures, shapes, regions, etc. Depending on the specific characteristics of a data set, some features can be more effective than others when performing similarity search. For instance, descriptors based on color representation might be effective with a data set containing mainly black and white images. Techniques based on statistical analysis of the data set and queries are useful.

From [11] a study conclude that a system based on the fuzzy c-means clustering algorithm, the CBIR system fuses color and texture features in image segmentation. A technique to form compound queries based on the combined features of different images is devised. This technique allows users to have a better control on the search criteria, thus a higher retrieval performance can be achieved.

The Color Selection exploited CBIR system [12], facilitates query-by-color. It is based on 11 color categories, used by all people, while thinking of and perceiving color. Then the low frequency DCT coefficients that are transformed from YUV color space as feature vectors are used for retrieval of images [13]. This system allows users to select its dominant feature of query images so as to improve the retrieval performance. But the technique is sufficient for performing effective retrieval by introducing users' opinions on the query images. In Region of Interest Image Indexing System [14], user can select the region of interest (ROI) and the system will search all the images in the database to find the all related regions among the database. A Universal Model for Content-Based Image Retrieval combine three feature extraction methods namely color, feature and edge histogram descriptor [15]. The image properties analyzed in this work are by using computer vision and image processing algorithms. For color the histogram of images are computed, for texture co-occurrence matrix based entropy, energy, etc., are calculated and for edge density it is Edge Histogram Descriptor (EHD) that is found. For retrieval of images, a novel idea is developed based on greedy strategy to reduce the computational complexity. Such existing approaches required large storage space and lot of computation time to calculate the matrix of features. Therefore, in this paper, the efficient content based image retrieval using advanced color and texture feature extraction is deployed. The color features are extracted using three color moments and texture features are extracted directly from block based DCT coefficients which are in transform domain. Hence it does not need any complex computation for texture feature extraction. The proposed method can be directly applied to image in the compressed domain, this solve the storage space problem.

Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape. Reasons for its development are that in many large image databases, 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. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. Several CBIR systems currently exist, and are being constantly developed.

3. Conclusions

This paper presents a survey of some modern content based image retrieval techniques. The CBIR technology can be used in several applications such as digital libraries, crime prevention, and photo sharing sites. Such a system has great value in apprehending suspects and identifying victims in forensics and law enforcement. A possible application is matching a forensic sketch to a gallery of mug shot images. The common problems have also been identified in the existing content based image retrieval techniques. This paper will help the researchers in identifying the problem areas in the field of content based image retrieval.

References

[1] Sang-Mi Lee, Hee-Jung Bae and Sung-Hwan Jung, “Efficient Content based Image Retrieval Methods using Color & Texture,” ETRI Journal, Vol. 20, No. 3, pp.272-283, 1998.
[2] Remco C. Veltkamp and Mirela Tanase, “Content Based Image Retrieval Systems: A Survey,” International Journal of Engineering Science and Technology, Vol. 20, No. 5, pp. 1-62, 2002.
[3] Dr. H.B.Kekre, Sudeep D. Thepade and Akshay Maloo,“Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform,” International Journal of Engineering Science and Technology, Vol. 2, No. 4, pp. 362-371, 2010.
[4] Gaurav Jaswal Amit Kaul, “Content Based Image Retrieval”, National Conference on Computing, Communication and Control, A Literature Review, National Institute of Technology, Hamirpur-177001, Himachal Pradesh (India).
[5] Ivan Lee, Paisarn Muneesawang, Ling Guan, “Automatic Relevance Feedback for Distributed Content-Based Image Retrieval”, ICGST, ieee.org FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
[6] Pooja Verma, Manish Mahajan, “Retrieval of better results by using shape techniques for content based retrieval”,IJCSC ,Vol. 3, No.2, January-June 2012, pp. 254-257, ISSN: 0973-7391.
[7] Ryszard S. Chora´s, “Image Feature Extraction

Techniques and Their Applications for CBIR and Biometrics Systems", International Journal of Biology and Biomedical Engineering Issue 1, Vol. 1, 2007.
[8] Swapnalini Pattanaik, Prof. D. G. Bhalke, "Beginners to Content Based Image Retrieval", International Journal of Scientific Research Engineering &Technology (IJSRET),Volume 1 Issue2 pp 040-044 May 2012 www. ijsret.org ISSN 2278 – 0882, IJSRET, 2012.
[9] Rong Zhao and William I. Grosky, "Bridging the Semantic Gap in Image Retrieval", Wayne State University, USA, Idea group publishing, 2002.
[10] Peter Stanchev, David Green Jr., and Boyan Dimitrov" MPEG-7: The Multimedia Content Description Interface", International Journal "Information Theories & Applications" Vol.11.
[11] www.scribd.com/doc/59845513/91/Fuzzy-cMeans-clustering
[12] E.L. van den Broek, L.G. Vuurpijl, P. Kisters and J.C.M. von Schmid Nijmegen, "Content Based Image Retrieval: Color Selection exploited," International Journal of Engineering Science and Technology, Vol. 30, No. 3, pp. 456-462, 1997.
[13] Tienwei Tsai, Taiwan, Yo Ping Huang and Te-Wei Chiang, "Fast Image Retrieval Using Low Frequency DCT Coefficients," International Journal of Engineering Science and Technology, Vol. 6, No. 3, pp. 106-120, 2003.
[14] Greg Pazz, Ramin Zabih and Justin Miller, "Region of Image Indexing System by DCT and Entropy," International Journal of Engineering Science and Technology, Vol. 8, No. 2, pp. 93-101, 2002.
[15] S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak, "A Universal Model for content Based Image Retrieval," International Journal of Computer Science, Vol. 4, No. 4, pp. 531-538, 2009.

Author Profile

Suhasini Gunjal has completed her B.E in Computer Engineering from University of Pune in 2011. Currently pursuing Master of Engineering from SVIT Chincholi, Nashik, India

Prof. S. M. Rokade has completed his B.E in Computer Engineering from Pune University and M.E in Computer Science & Engineering from MGM, Nanded. He is presently working as an Associate Professor in SVIT Chincholi, Nashik, India