A Review on Feature Extraction Techniques in Content Based Image Retrieval

Rachna
M.Tech Scholar, N.C.College of Engineering, India.

Jagtar Singh
Assistant Professor & H.O.D., N.C.College of Engineering, India.

Abstract – The Content Based Image retrieval (CBIR) is a very critical for retrieving the most visible applicable images from the huge image database. The numerous low stage functions are extracted primarily based on their visual content which might be shade, shape, texture and so on. The paper gives the review of color and texture function extraction techniques like coloration histogram, color correlogram, colour co-occurrence matrix and tamura texture characteristic, steerable pyramid, wavelet transform, Gabor wavelet remodel respectively and additionally the comparative evaluation of this strategies is proven within the paper.

Index Terms – CBIR; feature extraction; color histogram; color correlogram; DCD; CCM; tamura feature; wavelet transform; GWT.

1. INTRODUCTION

As bulk of digital images quickly will increase continuously, there's a necessary of better access techniques to retrieve image from a huge image database. It is difficult to get entry to this visible information until it's far prepared in a manner which gives efficient browsing, searching and retrieval. Traditional strategies used for photo retrieval are based totally on the annotation of keywords. Images annotation is time eating manner due to the fact its miles tough to describe their contents with words. In CBIR, the photograph annotation isn't always required. As it retrieves the favoured image from the big photo database based on the features which are extracted from the images. CBIR is considered as one of the only manner of retrieving visual statistics [8]. The contents of picture are defined by means of low-degree capabilities like shade, texture, shape and many others. The characteristic vector is generated via this set of low-degree function that is used to symbolize the content material of every image inside the photos database and then retrieval of photograph is based on the similarities of their contents. According to the similarity between the question and target picture the ordered list of the matched snap shots is displayed. The fig. 1 indicates block diagram of simple CBIR machine. In which, the visible contents of the database photos are extracted by the function extraction strategies and represented within the shape of feature vector.

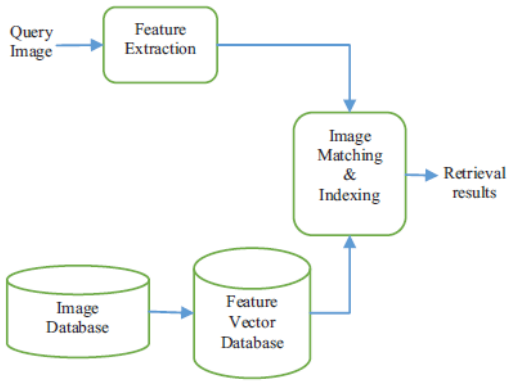

Figure.1 Block diagram of basic CBIR system

The feature vector of photographs in the database forms a characteristic database. The retrieval manner is initiated whilst the person offers the question photo to the machine. The question picture is converted into the inner illustration of function vector through the usage of same feature extraction techniques that changed into used for the image database function vector. The similarity measure is used to calculate the space among the characteristic vector of question image and target snap shots within the function database. The snap shots are ranked in step with their similarity cost and provided as output of CBIR system.

The corporation of the paper is as follows. The segment I give the introduction about the basic CBIR system. A review of the color function extraction strategies like color histogram, colour correlogram, coloration dominant descriptor and shade co-occurrence matrix are supplied in section II. In phase III, an overview tamura function, steerable pyramid, wavelet transform and Gabor wavelet transform are provided. The segment IV gives the comparative analysis of those one of a kind strategies and sooner or later the conclusion is mentioned in section V.

2. LITERATURE REVIEW

In this paper [1], A. Khodaskar and S. A. Ladhake have presented a CBIR technique using colour based feature. Since a colour image, consists of three basic colour components, i.e.

red, green and blue, so in this work, we have given the same importance on all three colour components during image retrieving process. In the presented CBIR technique, initially we have constructed three probability histograms for each colour component and subsequently the histograms are divided into several numbers of significant bins and from each bin, we have computed several statistical values like standard deviation, skewness and kurtosis. The computed statistical values are used as extracted features of the image data. The processing cost of the presented CBIR technique is significantly low. The technique has been tested on standard image databases and satisfactory results have been achieved.

In This paper [2], S. Pradeep and L. Malliga, presented content based image retrieval (CBIR) system using immense power of combination of soft computing techniques such as artificial neural network, fuzzy logic and support vector machine. Traditional content based image retrieval system are retrieved images using low level visual features, hence, suffer from semantic gap. There is need to reduce semantic gap and improve accuracy of CBIR even though size of image databases increase rapidly, because accuracy and efficiency of content-based image retrieval system is extremely important in any real world application. Proposed innovative framework for content-based image retrieval based on combination of three soft computing techniques enormously improves accuracy of image retrieval. A user gives input to the system in the form of specified query image and system return set of relevant images. Proposed system employs relevance feedback based on SVM that intelligently classify images relevant or irreverent to given query image. Performance of proposed content based image retrieval system pragmatically evaluated in term of precision, recall and accuracy.

This paper [3], J. M. Guo and H. Prasetyo proposed a content based image retrieval and segmentation of a medical image. The general purpose of fuzzy database management system (FDBMS) is a powerful tool for retrieval of image based on it's contending like color, shape or texture. This paper focused on the problem of retrieving and segmenting of a patient's medical image. There are two process involved: First image retrieval process is done for the retrieving the medical image from database based upon the CBIR technique. Second image segmentation process is done. After retrieving the medical image, segmentation process is done to obtain the particular region of a medical image on which diseases are affected.

This paper [4], Charmi Katira, Nikunj Vora, Krisha Wali, Abhinav Medhekar presented a technique for content-based image retrieval (CBIR) by exploiting the advantage of low complexity ordered-dither block truncation coding (ODBTC) for the generation of image content descriptor. In the encoding step, ODBTC compresses an image block into corresponding quantizers and bitmap image. Two image features are proposed to index an image, namely, color co-occurrence feature (CCF) and bit pattern features (BPF), which are generated directly from the ODBTC encoded data streams without performing the decoding process. The CCF and BPF of an image are simply derived from the two ODBTC quantizers and bitmap, respectively, by involving the visual codebook. Experimental results show that the proposed method is superior to the block truncation coding image retrieval systems and the other earlier methods, and thus prove that the ODBTC scheme is not only suited for image compression, because of its simplicity, but also offers a simple and effective descriptor to index images in CBIR system.

In This paper[5], Ekta Gupta and Rajendra Singh Kushwah outlines a description of some primitive features of image, which have been utilized in the presented system. These features are extracted and used as the basis for a similarity check between images. The algorithms used to calculate the similarity between extracted features, are then explained. Final result was a MATLAB built software application, with an image database, that utilized different feature of the images in the database as the basis of comparison and retrieval. The structure of the final software application is illustrated. Furthermore, the results of its performance are illustrated.

In This paper[6], Neelima Bagri1 and Punit Kumar Johari proposed the content based image retrieval, using features like colour and texture, called WBCHIR (Wavelet Based Colour Histogram Image Retrieval).The shape and shade features are extracted in the course of wavelet transformation and colour histogram and the arrangement of these features is vigorous to scaling and conversion of objects in an image. It is the first time to present segmentation and grid, feature extraction, K-means module and k-nearest neighbour clustering algorithms and bring in the neighbourhood module to build the CBIR system. It is the hybrid method of global and local features with k-means clustering algorithm.

In this paper [7], Davar Giveki, Ali Soltanshahi, Fatemeh Shiri, Hadis Tarrah comparisons of combination texture and shape features are done with texture Gray Level Co-occurrence Matrix and Hu-moments and the combination of tamura texture and shape invariant Hu-moments. For the performance evaluation of the system we use most commonly used methods namely precision and recall.

This paper [8], P. Selvi Rajendran, R. Sandya, T. Yogeswari and G. Viveka, first experimentally study the effect of choosing color space on the performance of content based image retrieval using Wavelet decomposition of each color channel. To this end, the retrieval results of different color spaces like RGB, YUV, HSV, YCbCr and Lab are analyzed. Then as a result a new Content Based Retrieval model using Wavelet Transform in Lab color space and Color Moments is proposed. In order to increase the efficiency of the proposed model some division schemes are taken into account which improves the performance of the proposed model. The proposed model

tackles one of the important restrictions in content based image retrieval, namely, the challenge between the accuracy of retrieval and its time complexity. The experimental results on two databases demonstrate the superiority of the proposed model compared to existing models.

In this paper [9], S. Sasikala, R. Soniya Gandhi three algorithms have been proposed to improve the performance of such sites (flicker, Photo bucket). To compare the similarity of the images efficiently, HMok-SimRank algorithm is used. It's derived from similarity algorithm. Integrated Weighted Similarity Learning (IWSL) is used to integrate meta information descriptions with image content. Finally, Ranking algorithm is used to rank the images for the order of retrieval. Benefits of our proposed system applied in flicker are experimentally shown in terms of both relevance and speed.

This paper [10], Sandeep Singh , Er.Rachna Rajput aims to introduce the problems and challenges concerned with the design and the creation of CBIR systems, which is based on the accurate image search mechanism. For efficient data management, a system is proposed which generates metadata for image contents. This system is using Content-Based Image Retrieval System (CBIR) based on Mpeg-7 descriptors. First, low-level features are extracted from the query image without metadata and the images with similar low-level features are retrieved from the CBIR system. Metadata of the result images which are similar to the query image are extracted from the metadata database. From the resulting metadata, common keywords are extracted and proposed as the keywords for the query image. The extraction of color features from digital images depends on an understanding of the theory of color and the representation of color in digital images. Color spaces are an important component for relating color to its representation in digital form. The transformations between different color spaces and the quantization of color information are primary determinants of a given feature extraction method. The approach is found to be robust in terms of accuracy and is 92.4% amongst five categories.

In this paper [11], T. Dharani and I. Laurence Aroquiaraj a survey on content based image retrieval presented. Content Based Image Retrieval (CBIR) is a technique which uses visual features of image such as color, shape, texture, etc…to search user required image from large image database according to user's requests in the form of a query image. We consider Content Based Image Retrieval viz. labelled and unlabelled images for analyzing efficient image for different image retrieval process viz. D-EM, SVM, RF, etc. To determining the efficient imaging for Content Based Image Retrieval, We performance literature review by using principles of Content Based Image Retrieval based unlabelled images. And also give some recommendations for improve the CBIR system using unlabelled images.

In this paper [12], A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain discussing the working conditions of content-based retrieval: patterns of use, **types** of pictures, the role of semantics, and the sensory gap. Subsequent sections discuss computational steps for image retrieval systems. Features for retrieval are discussed, sorted by: accumulative and global features, salient points, object and shape features, signs, and structural combinations thereof. Similarity of pictures and objects in pictures is reviewed for each of the feature types, in close connection to the types and means of feedback the user of the systems is capable of giving by interaction. We discussed aspects of system engineering: databases, system architecture, and evaluation. In the concluding section, we present our view on: the driving force of the field, the heritage from computer vision, the influence on computer vision, the role of similarity and of interaction, the need for databases, the problem of evaluation, and the role of the semantic gap.

[13] The numbers of digital images are increasing day by day and mining from large databases is becoming harder & harder. Indexing image data based on text is tiresome and error prone. If the indexing based on low-level feature of the image then it may reduce the workload and mining become faster. In this research paper we propose an indexing technique which indexes the digital images in the database by the highest color percentage. The images will be automatically classified by its own low-level feature i.e. Color. Implementation of this technique will be benefits the image mining.

[14] In order to improve the retrieval accuracy of content-based image retrieval systems, research focus has been shifted from designing sophisticated low-level feature extraction algorithms to reducing the 'semantic gap' between the visual features and the richness of human semantics. This paper attempts to provide a comprehensive survey of the recent technical achievements in high-level semantic-based image retrieval. Major recent publications are included in this survey covering different aspects of the research in this area, including low-level image feature extraction, similarity measurement, and deriving high-level semantic features. We identify five major categories of the state-of-the-art techniques in narrowing down the 'semantic gap': (1) using object ontology to define high-level concepts; (2) using machine learning methods to associate low-level features with query concepts; (3) using relevance feedback to learn users' intention; (4) generating semantic template to support high-level image retrieval; (5) fusing the evidences from HTML text and the visual content of images for WWW image retrieval. In addition, some other related issues such as image test bed and retrieval performance evaluation are also discussed. Finally, based on existing technology and the demand from real-world applications, a few promising future research directions are suggested

[15] Gabor wavelet proves to be very useful texture analysis and is widely adopted in the literature. In this paper, we presented an image retrieval method based on Gabor filter. Texture features are found by calculating the mean and variation of the Gabor filtered image. Rotation normalization is realized by a circular shift of the feature elements so that all images have the same dominant direction. The image indexing and retrieval are conducted on textured images and natural images.

Table 1 Literature Review of CBIR Techniques

Feature Techniques	Author's name	advantages	Disadvantages
CBIR technique	A. Khodaskar and S. A. Ladhake	Low processing cost	Less mobility of management
Soft computing techniques such as artificial neural network, fuzzy logic and support vector machine	S. Pradeep and L. Malliga	Easy to portable	Difficult to maintain
Fuzzy database management system (FDBMS)	J. M. Guo and H. Prasetyo	Less expensive	Maintenances required.
ODBTC techniques	Charmi Katira, Nikunj Vora, Krisha Wali, Abhinav Medhekar	low complexity	Less mobility
Soft computing techniques	Ekta Gupta and Rajendra Singh Kushwah	Better performance	More complex
WBCHIR	Neelima Bagri1 and Punit Kumar Johari	Less expensive	Difficulty occurs during extractions of image features
Gray Level Co-occurrence Matrix and Hu-moments	Davar Giveki, Ali Soltanshahi, Fatemeh Shiri, Hadis Tarrah,	Good precision value	Very expensive
Wavelet decomposition	P. Selvi Rajendran, R. Sandya, T. Yogeswari and G. Viveka	Less time to operate	More complex
Integrated Weighted Similarity Learning(IWSL)	Sandeep Singh , Er.Rachna Rajput	Less flicker	Less mobility
image retrieval process viz. D-EM, SVM, RF	T. Dharani and I. Laurence Aroquiaraj,	Improve the CBIR system	More complex
Accumulative and global features	A. W. M. Smeulders, M. Worring, S. Santini	Better performance	The problem of evaluative
indexing technique	] B. Tarulatha, N. Shroff and M. B. Chaudhary,	Good image mining	Moderate cost

low-level feature extraction algorithm	Ying Liua,∗, Dengsheng Zhanga, Guojun Lua	Improve retrieval accuracy	Less perfor mance
Gabor wavelet	T. Barbu	Good texture analysis	

3. CONCLUSION

Content based photograph retrieval is a hard methodology of shooting applicable photographs from a big area for storing. Although this vicinity has been explored for many years, no technique has finished the accuracy of human visual perception in distinguishing image. Whatever the length and content material of the photograph database is, a individual can honestly well known photographs of equal category. From the very beginning of CBIR analysis, similarity computation between snap shots used either area primarily based primarily or worldwide primarily based functions. Global features extracted from a picture are helpful in supplying textured photos that don't have any certain specific area of interest with relevancy the user. Region based features are extra effective to give an explanation for image that have awesome areas. Retrieval systems based on area capabilities are computationally pricey because of the requirement of segmentation technique within the beginning of a querying technique and were given to remember every image location in similarity computation.

We used a fixed of features vector by way of extracting totally one of a kind quite options that has its very own significance like we generally tend to use texture options extracted by GLCM matrix that considers but normally a pixel with the intensity value i occurs in a specific spatial dating to a pixel with the value j, Gabor filter, that's a effective texture extraction method both in describing the content of image regions or the worldwide content of a photograph. Color histogram as an international colour feature and histogram intersection as shade similarity metric. Except those statistical capabilities and wavelet capabilities also are applied in combination of above functions set. These all capabilities are blended right into a vector with size of 195 for a photograph. These capabilities are than examined with preloaded information in SVM to get the relevant class of check image. Our proposed functions extraction method finished very well in phrases of accuracy and correct class and it's 97.5311% compared to histogram features this is 94.319% and texture function set that is 91.58 %. The precision value for our projected capabilities set is a 100%.

REFERENCES

[1] Khodaskar and S. A. Ladhake, "A novel approach for content based image retrieval in context of combination S C techniques," *Computer Communication and Informatics (ICCCI), 2015 International Conference on*, Coimbatore, 2015, pp. 1-6

[2] S. Pradeep and L. Malliga, "Content based image retrieval and segmentation of medical image database with fuzzy values," *Information Communication and Embedded Systems (ICICES), 2014 International Conference on*, Chennai, 2014, pp. 1-7.

[3] J. M. Guo and H. Prasetyo, "Content-Based Image Retrieval Using Features Extracted From Halftoning-Based Block Truncation Coding," in *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1010-1024, March 2015.

[4] Charmi Katira, Nikunj Vora, Krisha Wali, Abhinav Medhekar," Advanced Content Based Image Retrieval Using Multiple Feature Extraction", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 4, Issue 10, October 2015

[5] Ekta Gupta and Rajendra Singh Kushwah," Combination of Local, Global and k-mean using Wavelet Transform for Content Base Image Retrieval", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.8, No.6 (2015), pp.253-266

[6] Neelima Bagri1 and Punit Kumar Johari," A Comparative Study on Feature Extraction using Texture and Shape for Content Based Image Retrieval", International Journal of Advanced Science and Technology Vol.80 (2015), pp.41-52

[7] Davar Giveki, Ali Soltanshahi, Fatemeh Shiri, Hadis Tarrah," A New Content Based Image Retrieval Model Based on Wavelet Transform", Journal of Computer and Communications, 2015, 3, 66-73

[8] P. Selvi Rajendran, R. Sandya, T. Yogeswari and G. Viveka," Novel Image Retrieval Approach in Similarity Integrated Network Using Combined Ranking Algorithm", World Engineering & Applied Sciences Journal 6 (3): 152-156, 2015.

[9] S. Sasikala, R. Soniya Gandhi," Efficient Content Based Image Retrieval System with Metadata Processing", IJIRST –International Journal for Innovative Research in Science & Technology, Volume 1, Issue 10, March 2015

[10] Sandeep Singh , Er.Rachna Rajput," Content Based Image Retrieval using SVM, NN and KNN Classification", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 6, June 2015.

[11] T. Dharani and I. Laurence Aroquiaraj, "Content Based Image Retrieval", International Conference on Pattern Recognition, Informatics and Mobile Engineering, February 21-22, 2013.

[12] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, Dec 2000.

[13] B. Tarulatha, N. Shroff and M. B. Chaudhary, "VIBGYOR indexing technique for image mining," *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, Ernakulam, 2016, pp. 191-193

[14] Ying Liua,∗, Dengsheng Zhanga, Guojun Lua," A survey of content-based image retrieval with high-level semantics", Wei-Ying Mab aGippsland School of Computing and Information Technology, Monash University, Vic 3842, Australia bMicrosoft Research Asia, No. 49 ZhiChun Road, Beijing 100080, China

[15] T. Barbu, "Content-Based Image Retrieval Using Gabor Filtering," *2009 20th International Workshop on Database and Expert Systems Application*, Linz, 2009, pp. 236-240.