An Extensive Review on CBIR Based on Combined Edge Detection, Color and Discrete Wavelet Transform

Jyoti Sharma[1], Dr. Vineet Richhariya[2]

[1]Research Scholar, [2]Guide & HOD

Department of Computer Science and Engineering, LNCT, Bhopal

Abstract: - In the recent past the advancement in computer and multimedia technologies has led to the production of digital images and cheap large image repositories. The size of image collections has increased rapidly due to this, including digital libraries, medical images etc. To tackle this rapid growth it is required to develop image retrieval systems which operate on a large scale. The primary aim is to build a robust system that creates, manages and query image databases in an accurate manner. A content-based image retrieval (CBIR) system works on the low-level visual features of a user input query image, which makes it difficult for the users to formulate the query and also does not give satisfactory retrieval results. In the past image annotation was proposed as the best possible system for CBIR which works on the principle of automatically assigning keywords to images that help image retrieval users to query images based on these keywords. There has been a great amount of research work done in the Content-Based Image Retrieval field, the majority of which covers issues such as Feature Representation, User Relevance Feedback and Indexing Structures. There has not been much investigation of the sample size and its effects on the accuracy of a retrieval system.

Keywords- YCbCr; HSI; RGB; Canny Edge Detection; prewitt edge detection, Color Edge Detection; Wavelet Transform; Haar; Precision; Recall.

I. INTRODUCTION

With the advancement in internet and multimedia technologies, a huge amount of multimedia data in the form of audio, video and images has been used in many fields like medical treatment, satellite data, video and still images repositories, digital forensics and surveillance system. This has created an ongoing demand of systems that can store and retrieve multimedia data in an effective way. Many multimedia information storage and retrieval systems have been developed till now for catering these demands. The most common retrieval systems are Text Based Image Retrieval (TBIR) systems, where the search is based on automatic or manual annotation of images. A conventional TBIR searches the database for the similar text surrounding the image as given in the query string. The commonly used TBIR system is Google Images. The text based systems are fast as the string matching is computationally less time consuming process. However, it is sometimes difficult to express the whole visual content of images in words and TBIR may end up in producing irrelevant results. In addition annotation of images is not always correct and consumes a lot of time. For finding the alternative way of searching and overcoming the limitations imposed by TBIR systems more intuitive and user friendly content based image retrieval systems (CBIR) were developed. A CBIR system uses visual contents of the images described in the form of low level features like color, texture, shape and spatial locations to represent the images in the databases. The system retrieves similar images when an example image or sketch is presented as input to the system. Querying in this way eliminates the need of describing the visual content of images in words and is close to human perception of visual data.

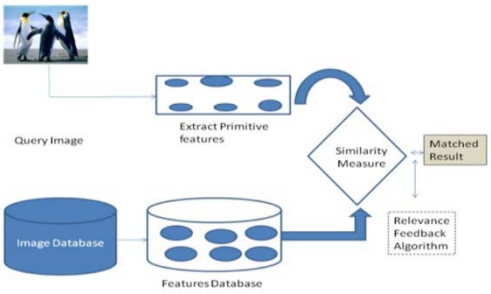

Fig. 1.1: Architecture of a typical CBIR system

In a typical CBIR system (Fig. 1.1), image low level features like color, texture, shape and spatial locations are represented in the form of a multidimensional feature vector. The feature vectors of images in the database form a feature database. The retrieval process is initiated when a user query the system using an example image or sketch of the object. The query image is converted into the internal representation of feature vector using the same feature extraction routine that was used for building the feature database. The similarity measure is employed to calculate the distance between the feature vectors of query image and those of the target images in the feature database. Finally, the retrieval is performed using an indexing

scheme which facilitates the efficient searching of the image database. Recently, user's relevance feedback is also incorporated to further improve the retrieval process in order to produce perceptually and semantically more meaningful retrieval results.

Visual Content Descriptor

Natural images depicting a complex scene may contain a variety of visual artifacts. CBIR systems represent the visual contents of images in the form of a feature descriptor. A good descriptor should not only be invariant to rotation, scaling and illumination variations but also has high discriminating capability. There is a trade-off between invariance and discriminating power of visual features. Employing features having wide variety of invariance may in losing the capability to discriminate between most essential properties. Study of invariance is largely investigated in the field of computer vision but is relatively new in image retrieval.

A feature descriptor may be local or global. Local descriptors are extracted using a part or region of an image while a global feature uses the visual content of the whole image. A CBIR system which uses region features to represent images is known as Region Based Image Retrieval systems (RBIR). On the other hand CBIR systems utilizing global features for describing images are classified as Global CBIR systems. Local and Global features of an image largely represent color, texture, shape and spatial relationships of different objects in the image. Some widely used color, texture, shape and spatial relationship features are discussed in the following subsections.

Color Features

Color is the most commonly used feature of an image. The perceived color at any pixel of an image is obtained by mixing three preliminary colors in appropriate proportion. The three dimensional color provides more discriminating information than the single dimensional gray level values. Before extracting color descriptor a proper color space must be determined first. Commonly used color spaces for image retrieval application are RGB, CIE L*a*b*, CIE L*u*v*, HSV and opponent color space. There is no agreement over which color space is best but one of the desirable characteristic of color space for image retrieval task is its uniformity. Uniformity means that the physical distance between any two color pair in the color space must be equal to the perceived distance between them. Some commonly used color descriptors are color moments, color histogram, color coherence vector and color correlogram.

Texture Features

There is no specific definition of texture one can define texture as the visual pattern that has properties of homogeneity not resulting from the presence of only a single color or intensity. Various techniques for texture analysis have been investigated in the field of computer vision and pattern recognition. The texture extraction techniques can be classified into two categories: statistical and structural. Statistical approaches use intensity distribution of image to extract statistical parameters representing texture of image. Commonly used statistical methods include Fourier power spectra, Co-occurrence matrices, Shift-invariant principal component analysis (SPCA), Tamura feature, Wold decomposition, Markov random field, Fractal model, and Multi-resolution filtering techniques such as Gabor and wavelet transform. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular.

Shape Features

Shape feature provides the most important semantic information about an image. Shape features are usually described using part or region of an image. The accuracy of shape features largely depends upon the segmentation scheme used to divide an image into meaningful objects. Fast and robust segmentation is difficult to achieve. This limits the shape features only to those retrieval applications where objects or region of images are readily available. The shape descriptors are categorized into two classes: boundary based descriptor and region based descriptor. Some boundary based representative shape description techniques are chain codes, polygonal approximations, Fourier descriptor and finite element model. On the other hand state of the art region based descriptors are statistical moment and area. A good shape feature should be invariant to translation, rotation and scaling. A detail review of shape matching techniques used in image retrieval application can be found in[Mingqiang et al.(2008)].

Spatial Information

The performance of an image retrieval system can be improved by considering spatial locations of different objects in the image. The spatial location of objects and their relationship can provide useful discriminating information in image retrieval applications. For instance, parts of blue sky and ocean may have similar color histograms, but their spatial locations in images are different. The spatial location matching can be implemented by matching the images based on fixed

location similarity. In this approach a similar object lying in different regions of an image cannot be detected. For instance; image having tiger in the left part may not get similarity with images having tiger in the right part of images. To overcome this problem systems compare all region of image with the query object or region. This may be in the increase of response time of the system. The most commonly used techniques for finding spatial location similarity includes 2D strings, spatial quad-tree and symbolic images.

Similarity Measure

The degree of similarity between query and target images is calculated based on the value of similarity measure. The images are ranked according to their similarity value and presented as output of CBIR system. Often, the choice of similarity measure affects the performance of retrieval system. Many similarity measures have been developed over the years based on the quantitative estimates of the distribution of features in the image. Some of the most commonly used similarity measures employed in CBIR are Euclidean distance , Minkowski- form distance, Histogram intersection distance, Quadratic- form distance, Mahalanobis distance and Kullback Leibler (KL) divergence distance.

Performance Evaluation

The performance of a retrieval system is evaluated based on several criteria. Some of the commonly used performance measures are average precision, average recall, average retrieval rate and Average Normalized modified Retrieval Rate (ANMRR). All these parameter are computed using precision and recall values computed for each query image. The precision of the retrieval is defined as the fraction of the retrieved images that are indeed relevant for the query:

$$Precision = \frac{No.of\ relevant\ images\ retrieved}{Total\ No.of\ images\ retrieved\ from\ the\ database}$$

The recall is the fraction of relevant images that is returned by the query:

$$Recall = \frac{No.of\ relevant\ images\ retrieved}{Total\ No.of\ relevant\ images\ in\ the\ database}$$

A good retrieval system should have high values for precision and recall.

II. SYSTEM MODEL

Convolutional neural network (CNN) is a type of feed-forward artificial neural network where the individual neurons are tiled in such a way that they respond to overlapping regions in the visual field". They are biologically-inspired invariant of Multilayer Perceptrons (MLP) which are designed for the purpose of minimal preprocessing. These models are widely used in image and video recognition. When CNNs are used for image recognition, they look at small portions of the input image called receptive fields with the help of multiple layers of small neuron collections which the model contains.

Sparse Connectivity

CNNs enforce a local connectivity pattern between neurons of adjacent layers to exploit spatially-local correlation [6]. Here illustrated in fig.1.2 that in layer m the inputs of hidden units are from a subset of units in layer m-1, units containing spatially adjoining receptive fields.

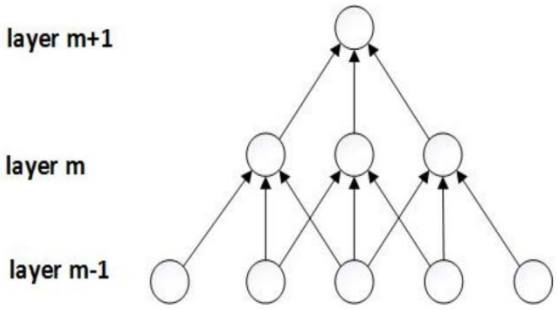

Fig. 1.2: Sparse Connectivity

Let us consider layer m-1 as an input retina. It can be seen in the figure that the layer m have receptive fields of width 3 in the input retina and are thus connected only to 3 adjacent neurons in the retina layer [6]. There is similar connectivity between the units in layer m+1 and the layer below. It can be said that their with respect to the input receptive field is larger whereas with respect to the layer below their receptive field is 3. There is no response in the each unit to variations which are outside their receptive fields with respect to the retina 18 thus ensuring that the strongest response to a spatially local input pattern is produced by the learnt filter.

III. LITERATURE SURVEY

N. Mariam and Rejiram R, [1] Content based image retrieval systems use the contents of the images to represent and access the images. Content basically refers to the image descriptors like color, texture and shape of the image. Among the different image features, edges are the important one as edges represent mainly the local intensity variations. But in the case of color images in order to obtain satisfactory results, authors must consider the color of the image during retrieval processes. This paper describes a new method in which both edge and color features of the images are considered for generation of

feature vectors. Discrete Wavelet Transform is used to preserve the detailed contents of the images along with the reduction of the size of the feature vector. A comparison study on the effect of the proposed method on YCbCr and HSI color spaces is also presented in this paper.

G. Rafiee, S. S. Dlay and W. L. Woo, [2] A comprehensive survey on patch recognition, which is a crucial part of content-based image retrieval (CBIR), is presented. CBIR can be viewed as a methodology in which three correlated modules including patch sampling, characterizing, and recognizing are employed. This paper aims to evaluate meaningful models for one of the most challenging problems in image understanding, specifically, for the effective and efficient mapping between image visual features and high-level semantic concepts. To achieve this, the latest classification, clustering, and interactive methods have been meticulously discussed. Finally, several recommendations for future research issues have been suggested based on the weaknesses of recent technologies.

B. L. Deekshatulu,[3] Image retrieval techniques are text based or content based. In the Text based image retrieval, the images are annotated and the database management system retrieves them. Adding annotations and textual attributes through human intervention, is difficult and time consuming. Content Based Image Retrieval (CBIR) is a better alternative to this traditional text based image retrieval.

A. Mumtaz, S. A. M. Gilani and T. Jameel,[4] This paper presents a novel texture image retrieval system (SVMBIR) based on dual tree complex wavelet transform (CWT) and support vector machines (SVM). Authors have shown that how one can improve the performance of image retrieval systems by assuming two attributes. Firstly, images that user needs through query image are similar to a group of images with same conception. Secondly, there exists non-linear relationship between feature vectors of different images and can be exploited very efficiently with the use of support vector machines. At first level, for low level feature extraction authors have used dual tree complex wavelet transform because recently it is proven to be one of the best for texture based features. At second level to extract semantic concepts, authors grouped images of typical texture classes with the use of one against all support vector machines. Authors have also shown how one can use a correlation based distance metric for comparison of SVM distance vectors. The experimental results show that the proposed approach has superior retrieval performance over the existing linear feature combining techniques.

B. Xue and L. Wanjun, [5] Research on image retrieval technology based on color feature, for the color histogram with a rotation, translation invariance of the advantages and disadvantages of lack of space, a color histogram and color moment combination image retrieval. The theory is a separate color images and color histogram moment of extraction, and then two methods of extracting color feature vector weighted to achieve similar distance, similar to the last distance based on the size of the return search results, based on the realization of the characteristics of the color image retrieval system. The results show that the method is rotation, translation invariance, a single method of extracting color features, enhanced image search and improve the accuracy of the sort.

IV. PROBLEM IDENTIFICATION

In the previous work, a content based image retrieval system has been presented which relies on the image features like color and edge. Authors has used discrete wavelet transform to extract the feature vector having smaller dimension. By using wavelet transform, they are able to preserve the finer details of the images. The presented method extract edges from Y matrix of the YCbCr using Canny and Prewitt edge detectors. Color is preserved by using the RGB histogram as the global statistical descriptor. Manhattan distance has been used as a similarity measure to detect the final image rank. The same procedure has been performed by using HIS color space, in which edges are extracted from the HSI image. Also the computational complexity may further reduced by a large amount by using the different methods to obtain feature vector with smaller dimension.

V. CONCLUSION

Image retrieval is a field of study that is concerned with searching and browsing digital images from a database collection. This area of research is very active since the 1970s. At that time, keyword-based image retrieval systems were used that utilized keywords as descriptors to index images Content based sub-image retrieval is a challenging system of methods of capturing relevant images from a large storage space. Individual methods or specialized approaches have generated a significant amount of published research work. The effectiveness and efficiency of most existing systems is rather poor especially when applied to general image databases. No specialized techniques limited to a highly-specific domain or a limited database size can claim victory over the query by content problem in such databases. Without adoption of meaningful methods of data compression and/or indexation retrieval methods retrieval systems are computationally very expensive.

REFERENCES

[1] N. Mariam and Rejiram R, "A modified approach in CBIR based on combined edge detection, color and discrete Wavelet Transform," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, 2015, pp. 2201-2205.

[2] G. Rafiee, S. S. Dlay and W. L. Woo, "A review of content-based image retrieval," 2010 7th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2010), Newcastle upon Tyne, 2010, pp. 775-779.

[3] B. L. Deekshatulu, "Learning Semantics in Content Based Image Retrieval (CBIR) - A Brief Review," 2010 Second Vaagdevi International Conference on Information Technology for Real World Problems, Warangal, 2010, pp. 76-78.

[4] A. Mumtaz, S. A. M. Gilani and T. Jameel, "A Novel Texture Image Retrieval System Based On Dual Tree Complex Wavelet Transform and Support Vector Machines," 2006 International Conference on Emerging Technologies, Peshawar, 2006, pp. 108-114.

[5] B. Xue and L. Wanjun, "Research of Image Retrieval Based on Color," 2009 International Forum on Computer Science-Technology and Applications, Chongqing, 2009, pp. 283-286.

[6] Wu Siqing; Xiong Gang; Wang Geng; Zhang Guoping, ”Digital-image retrieval based on shape-feature and MPEG-7 standard,” International Conference on Consumer Electronics, Communications and Networks (CECNet), 2011 , vol., no., pp.3095,3097, 16-18 April 2011.

[7] Castellano, G.; Castiello, C.; Fanelli, A.M., ”Content-based image retrieval by shape matching,” Annual meeting of the North American Fuzzy Information Processing Society, (NAFIPS) 2006.,vol.,no., pp.114,119, 3- 6 June 2006

[8] Hongyan Sun; Shuxue Tian, ”Image retrieval based on blocked histogram and Sobel edge detection algorithm,” International Conference onComputer Science and Service System (CSSS), 2011, vol., no., pp.3277,3281, 27-29 June 2011 .

[9] Ramamurthy, K. R. Chandran, ”Content Based Image Retrieval for Medical Images Using canny Edge Detection Algorithm”, International Journal of Computer Applications (0975-8887), Vol. 17, No. 6, Marcb 2012

[10] Lamard, M.; Cazuguel, G.; Quellec, G.; Bekri, L.; Roux, C.; Cochener, B., ”Content Based Image Retrieval based on Wavelet Transform coefficients distribution,” Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE , vol., no., pp.4532,4535, 22-26 Aug. 2007

[11] Agarwal, S.; Verma, A.K.; Singh, P., ”Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor,” International Conference on Information Systems and Computer Networks (ISCON), 2013 , vol., no., pp.19,23, 9-10 March 2013

[12] Agarwal, S.; Verma, A.K.; Dixit, N., ”Content Based Image Retrieval using Color Edge Detection and Discrete Wavelet Transform,” International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 , vol., no., pp.368,372, 7-8 Feb. 201

[13] Ashok Kumar, J. Estber, ”Comparative Study on CBIR based by Color Histogram, Gabor and Wavelet Transform”, International Journal of Computer Applications (0975 - 8887) Vol. 17, No.3, March 2011