Content Based Image Retrieval: Survey on the Combination of Colour Features and Texture Features

Sonam
PG Scholar, Computer Science and Engineering, Galgotias University, Greater Noida, U.P, India

Pranjul Mishra
PG Scholar, Computer Science and Engineering, Galgotias University, Greater Noida, U.P, India

***Abstract*--Digital image processing encompasses processes whose inputs and outputs are images and also encompasses process that extract attributes from images and including the recognition of individual objects. With the increase in popularity of the network and development of multimedia technologies, users are not satisfied with the traditional information retrieval techniques so there should be an appropriate system to efficiently manage these collections. The most common approach used is Content Based Image Retrieval . Goal of Content Based Image Retrieval system is to support image retrieval based on visual content of image. There are many algorithms for efficient image retrieval. In this survey paper the techniques used for content based image retrieval are discussed. It also introduced the combination of features like color, texture for accurate and effective Content Based Image Retrieval System.**

Keywords: Content Based Image Retrieval, Colour, Texture.

I. INTRODUCTION

Due to the repaid development of internet technology, image documents have become an important information source. It is hard to retrieve certain images from all available images. For this purpose, many general purpose image retrieval systems have been developed.

Firstly text base retrieval system came but due to some disadvantages of that system content-based image retrieval (CBIR) was introduced in the early 1980s.

CBIR is an automated system that searches for a query image in an image database and retrieves the relevant images [1,2].

Image retrieval based on content is extremely useful in a many applications such as publishing and advertising, historical research, fashion and graphic design, architectural and engineering design, crime prevention, medical diagnosis, geographical information and remote sensing systems, etc [3].

A typical CBIR system uses the contents of an image to represent and access. Colour, texture, shape are the visual contents used for retrieving a particular image from the database[4]. Extraction of good features which compactly represent a query image is one of the important tasks in CBIR . Colour is the most commonly used visual feature and is invariant to image size and orientation, because we can simply extract colour information from images. Extracting information about shape and texture feature are much more complex tasks, so usually performed after extracting color features from image[2,3,4].

Shape is a visual feature that describes the contours of objects in an image, which are usually extracted from segmenting the image into meaningful regions or objects [4]. However, since it is difficult to achieve such image segmentation for natural images, the use of shape features in image retrieval has been limited to special applications [5].

Texture is also a visual feature that refers to surface properties of an object and their relationship to the surrounding environment. It is intended to capture the granularity and repetitive patterns of surfaces within a picture [4, 5].

Various image retrieval systems has been built, based on the low-level features for general or specific image retrieval tasks. Some of the existing CBIR systems are :

- *Query by Image Content (QBIC)* - It is the first commercial content based retrieval system [6]. This system allows users to graphically pose and refine queries based on multiple visual properties such as color, texture and shape. It supports queries based on input images, user-constructed sketches, and selected colour and texture patterns.
- *WebSEEK and VisualSEEK* - WebSEEk is a World Wide Web oriented text/image search engine and VisualSEEk is a visual feature

search engine, both are developed at Columbia University.

- *Virage* – It is content based image search engine developed at Virage Inc. It supports colour and spatial location matching as well as texture matching.
- *NeTra* - This system uses colour, shape, spatial layout and texture matching, as well as image segmentation [7].
- *Visual Information Processing for Enhanced Retrieval* (*Viper*) - This system retrieves images based on color and texture matching.

II. CLASSIFICATION OF IMAGES

(i) Intensity Images

An intensity image is a data matrix, whose values represent intensities within some range. An intensity image is represented as a single matrix, with each element of the matrix corresponding to one pixel of an image. The matrix can be of class double, uint8, or uint16. Class "double" assigns a floating number ("a number with decimals") between 0 and 1 to each pixel. The other class "uint8" assigns an integer value between 0 and 255 to represent the brightness of a pixel [8, 9].

(ii) Indexed Images

In an indexed image, the image matrix values do not determine the pixel colours directly. Instead, MATLAB uses the matrix values as indices for looking up colors in the figure's colormap. This is a practical way of representing colour images. An indexed image stores an image as two matrices. The first matrix has the same size as the image and a value for each pixel. The second matrix is called the color map and its size may be different from the image [8, 9].

(iii) Binary Images

This image format also stores an image as a matrix but can only colour a pixel black or white (and nothing in between). It assigns a 0 for black and a 1 for white [8, 9].

(iv) RGB images

RGB image is stored as an m-by-n-by-3 data array that defines red, green, and blue colour components for each individual pixel. The colour of each pixel is determined by the combination of the red, green, and blue intensities stored in each color plane at the pixel's location. Graphics file formats store RGB images as 24 bit images, where the red, green, and blue components are 8 bits each. This yields a potential of 16 million colors. [8,9].

III. General Image Retrieval System

The overall structure of a general image retrieval system is shown in Fig. 1. Basic idea behind CBIR is that, when building an image database, feature vectors from images (color, texture, shape, etc.) are to be extracted and then store the vectors in another database for future use.

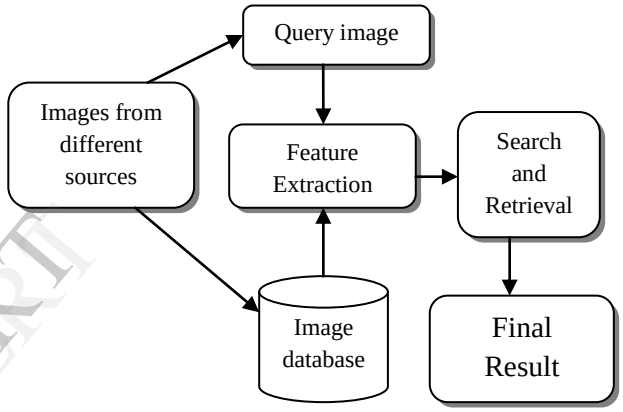

Fig.1

IV. CURRENT IMAGE RETRIEVAL APPROACHES

CBIR is basically a two step process which is *Feature Extraction* and *Feature Matching*. Feature Extraction is the process to extract image features to a distinguishable extent [10, 28]. Information extracted from images such as colour, texture and shape are known as feature vectors. The extraction process is done on both query images and images in the database. Image matching [25] involves using the features of both images and comparing them to search for similar features of the images in the database [26] . Using multiple feature vectors to describe an image during retrieval process increases the accuracy when compared to the retrieval using single feature vector [3]. For example, searching of image based on its colour and texture provides a better result than using a single colour feature.

V. FEATURE EXTRACTION

A. *Colour Feature*

Human eyes are sensitive to colours, human can distinguish objects in the images due to colour feature. Colour feature provide sometimes powerful information about images, and they are very useful in image retrieval. Many methods can be used to describe colour feature. There are colour histogram [11, 27], colour correlation [12], colour moments [13], colour structure descriptor (CSD), and scalable colour descriptor (SCD) [14].

B. *Texture Feature*

Quantifying texture content of an image is the most important method to image region description. No formal definition for texture, but we can say that it provides the measures of properties such as smoothness, coarseness, and regularity [15,16]. Furthermore, texture can be thought as repeated patterns of pixels over a spatial domain. If the texture has exposed to some noise, the patterns and their repetition in the texture can be random and unstructured. Some common methods are used for texture feature extraction such as statistical, model-based, and transform-based methods. Most existing approaches to texture feature extraction use *statistical methods.* Some methods for extracting Texture features are Gray Level Co-occurrence matrix (GLCM), Gabor Transform and Tamura Features.

VI. EVOLUTION

Early work on image retrieval can be traced back to the late 1970s. In 1979, a conference on Database Techniques for Pictorial Applications was held in Florence. Since then, the application potential of image database management techniques has attracted the attention of researchers. In the early 1990s, as a result of advances in the Internet and new digital image sensor technologies, the volume of digital images produced by scientific, educational, medical, industrial, and other applications available to users increased dramatically. The difficulties faced by text based retrieval became more and more severe. The efficient management of the rapidly expanding visual information became an urgent problem. Since the 1990s, content based image retrieval has become an active and fast advancing research area in image retrieval [30]. Early studies on CBIR used a single visual content such as colour, texture, or shape to describe the image. The drawback of this method is that using one feature is not enough to describe the image since the image contains various visual characteristics. So then it started to combine colour feature with other feature . But to extract shape feature from images was very complex task and time consuming [35]. But Colour and texture feature extraction are simpler then compared to other feature. So combination of colour and texture feature became popular because it enhance the performance of CBIR system [26, 37].

In 2003 *H.* Permuter, J. Francos and I. H. Jermyn [17] described a Gaussian mixture model, in which the mean and covariance of RGB values were used to extract colour feature and the energies of DCT coefficient were used to extract texture feature. Then A Vadivel, A K Majumdar and Shamik Sural in 2004 [16] proposed a combination of colour and texture features for CBIR. For extracting colour feature they generated colour histogram and for extraction of texture feature Haar Wavelet or Daubechies' wavelet was used. During retrieval both colour and texture feature of query image was combined, weighted and compared with feature vector of database images using Manhattan distance [30] metric. The retrieved result is dependent on the weight given to each of the feature vector. By combining these two features with weight have enhanced the performance of system. In oct, 2004 Spyros Liapis and Georgios Tziritas [18] explored a mechanism in which for texture feature extraction they used Discrete Wavelet Frames(DWF). For extraction of colour feature Two-dimensional (2-D) or one-dimensional (1-D) histograms of the *CIE Lab* chromaticity coordinates are used. The similarity measure [25] defined on the feature distribution was based on the Bhattacharya distance. This approach was as effective as other methods while computationally more tractable. In the above methods the feature vector dimension is not considered as an important factor in combining multiple features. Then Young Deok Chun, Nam Chul Kim and Ick Hoon Jang [19] in 2008 proposed a CBIR method based on the efficient combination of multiresolution colour and texture features. The method used the combination of colour autocorrrelograms of hue and saturation component images and BDIP and BVLC moments value of component images in the wavelet transform domain. The dimension of the combined feature vector is determined at a saturation point above which the retrieval accuracy nearly increases as the dimension increases. This method yields higher retrieval accuracy then some other conventional methods. In 2008 [20] S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak described a novel approach for generalized image

retrieval. A combination of three feature extraction methods namely colour, texture, and edge histogram descriptor was presented. For colour, the histogram of images are computed, for texture co occurrence matrix (CCM) based entropy, energy, etc, are calculated and for edge density , Edge Histogram Descriptor (EHD) found. This novel idea was developed based on greedy strategy to reduce the computational complexity. But this approach required large storage space and lot of computation time to calculate the matrix of features. Computational complexity and retrieval efficiency are the key objectives in the image retrieval system. Nevertheless it is very difficult to reduce the computations and improve the performance of image retrieval technique. Then in 2010 Dr. H.B.Kekre, Sudeep D. Thepade and Akshay Maloo [21], presented a innovative content based image retrieval (CBIR) technique based on feature vectors as fractional coefficients of transformed images using DCT and Walsh transforms. In that paper the feature vector size per image was greatly reduced by taking fractional coefficients of transformed image. The feature vectors were extracted in fourteen different ways from the transformed image. The performance of image retrieval was improved using fractional coefficients of transformed images at reduced computational complexity. For DCT and Walsh transform the average precision/recall values of CBIR using fractional coefficients were better than using all coefficients of transformed images as feature vectors. Then P. V. N. Reddy and K. Satya Prasad in 2011 [22] proposed a Multi wavelet-based approach for texture feature extraction and colour correlograms was used for colour feature extraction for CBIR. colour and texture features were combined to improve the retrieval efficiency. Multiwavelets have several advantages in comparison with scalar wavelet. The features such as compact support, Orthogonality, symmetry, and high order approximation are the base features for this transform . A Multiwavelet system can simultaneously provide perfect representation while preserving length, good performance at the boundaries, and a high order of approximation . Multiwavelets was initiated by Goodman, Lee and Tang. In 2011 [23] Sagar Soman, Mitali Ghorpade, Vrushali Sonone and Satish Chavan proposed a CBIR system, in which the texture features were extracted from the query image by applying block wise Discrete Cosine Transforms (DCT) [33,35] on the entire image and from the retrieved images the color features were extracted by using moments of colors (Mean, Deviation and Skewness) theory.

VII. RECENT WORK

For extracting colour feature from an image we have to choose a colour space, then we use the properties of chosen colour space for extraction.In Digital image processing , RGB colour space is most prevalent choice, so we were using it till now. But main drawbacks of RGB space is that it is perceptually non-uniform and device dependent system. So nowadays HSV colour space [24] is useful for extraction. The HSV color space is an intuitive system, which describes a specific colour by its hue, saturation, and brightness values [29]. This color system is very useful in interactive colour selection and manipulation.

Some new techniques for texture feature extraction like Gray Level Co-occurrence matrix (GLCM), Gabor Transform[32], Ranklet Transform are using with HSV colour space for better retrieval. Nowadays work on the combination of colour, shape and texture is also going on.

VIII. CONCLUSION

The purpose of this survey is to provide an overview of various methodologies for content based image retrieval systems. Mostly systems are using color and texture features, so this paper provides a detailed review of the works carried out on these two features. Combination of these two feature improve the performance of the system. Different-2 techniques for extracting colour and texture features are combined to get better accuracy and work is still in the progress.

REFERENCES

[1] Gudivada and V. Raghavan, “Content-based image retrieval systems,” in IEEE Computer, vol. 28, no 9, pp18-22, Sep. 1995.
[2] Source http://en.wikipedia.org/wiki/Content-based_image_retrieval.
[3] Sang-Mi Lee, Hee-Jung Bae and Sung-Hwan Jung, “ Efficient Content based Image Retrieval
Methods using Color & Texture,” in ETRI Journal, Vol. 20, No. 3, pp.272-283, 1998.
[4] Remco C. Veltkamp and Mirela Tanase, “Content Based Image Retrieval Systems: A Survey,” in International Journal of Engineering Science and Technology, Vol. 20, No. 5, pp. 1-62, 2002.
[5] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng, "Fundamentals of Content-Based Image Retrieval" - http://research.microsoft.com/asia/dload_files/group/mcomputing/2003 P/ch01_Long_v40-proof.pdf.
[6] W. Niblack, R. Barber,W. Equitz, M. Flickner, E. Glasman, D. Pektovic, P. Yanker, C. Faloutsos, and G. Taubin, “The QBIC Project: Querying Images by Content using Color Texture and Shape”, Proc. SPIE Int. Soc. Opt. Eng., in Storage and Retrieval for Image and Video Databases, vol. 1908, pp. 173-187, 1993.
[7] W.Y. Ma and B.S. Manjunath, “NeTra: A toolbox for navigating large image databases”, in IEEE Int. Conf. on Image Processing (1997) pp. 568-571.
[8] http//www. mathswork.com.
[9] Rafel C. Gonzalez and Richard E. Woods, “Digital Image Processing”, Second Edition, Pearson Education Asia, 2005.
[10] Sagarmay Deb Yanchun Zhang “An Overview of Content-based Image Retrieval Techniques” 2004 IEEE.

[11] R. Brunelli and O. Mich, "Histograms Analysis for Image Retrieval," Pattern Recognition, Vol.34, No.8, pp1625-1637,2001.
[12] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," IEEE Int. Conf. Computer Vision and Pattern Recognition, San Juan, Puerto Rico, Jun. 1997, pp.762–768.
[13] A. Del Bimbo, M. Mugnaini, P. Pala, and F. Turco. Picasso, "Visual querying by color perceptive regions," In Proceedings of the 2nd International Conference on Visual Information Systems, San Diego, December'97, pages 125-131,1997.
[14] ISO/1EC JTC 1/SC29/WG 11/N4062: "MPEG-7 requirements document" 2001.
[15] T. Chang and C-C. J. Kuo, "Texture Analysis and Classification with Tree-structured Wavelet Transform", IEEE Trans. on Image Processing, 2, no. 4, 1993, pp. 429-441.
[16] A. Vadivel, A. K. Majumdar, and S. Sural, "Characteristics of weighted feature vector in content-based image retrieval applications," in Proc. IEEE Int. Conf. Intelligent Sensing and Information Processing, Chennai, India, Jan. 2004, pp. 127–132.
[17] H. Permuter, J. Francos, and I. H. Jermyn, "Gaussian mixture models of texture and colour for image database retrieval," in *Proc.* IEEE Int. Conf. Acoustics, Speech, Signal Processing, Hong Kong, Apr. 2003, vol. 3, pp. 569–572.
[18] S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," IEEE Trans. Multimedia, vol. 6, pp. 676–686, Oct. 2004.
[19] Young Deok Chun, Nam Chul Kim and Ick Hoon Jang, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features," in IEEE transactions on multimedia, vol. 10, no. 6, oct 2008.
[20] S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak, "A Universal Model for Content-Based Image Retrieval ,"in World Academy of Science, Engineering and Technology 46 2008.
[21] Dr. H.B.Kekre, Sudeep D. Thepade, Akshay Maloo, "Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform," in International Journal of Engineering Science and Technology, Vol. 2(4), 2010, 362-371.
[22] P. V. N. Reddy, K. Satya Prasad "Colour and Texture Features for Content Based Image Retrieval," Int. J. Comp. Tech. Appl., Vol 2 (4), 1016-1020.
[23] Sagar Soman, Mitali Ghopade, Vrushali Sonone, Satish Chavan, "Content based image retrieval using advanced color and texture features,"ICCIA, 2011.
[24] Ching-Hung Su, Huang-Sen Chiu, Tsai-Ming Hsiehb, "An efficient image retrieval based on HSV color space".
[25] S.Pavani, Shivani, T.Venkat Narayana Rao, Deva Shekar, " Similarity Analysis Of Images Using Content Based Image Retrieval System," International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 2 Issue 1 Jan 2013 Page No. 251-258.
[26] Faiz Z Rangari, Dr. Uma Ramarao, "Searching by Content based image retrieval through combined features," International Journal of Science and Engineering, ISSN-2347-2200, Volume 1, Number 2 – 2013.
[27] Shamik Sural, Gang Qian and Sakti Pramanik,"Segmentation and histogram generation using the HSV color space for image retrieval," IEEE ICIP 2002.
[28] 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.
[29] Simardeep Kaur, Dr. Vijay Kumar Banga, "Content based image retrieval : survey and comparison between RGB and HSV model," IJETT, vol.4, issue 4, April 2013.
[30] Felci Rajam and S. Valli, "A Survey on Content Based Image Retrieval ," in Life Science Journal 2013; 10(2).
[31] Jing Huang, S Ravi Kumar, Mandar Mitraz Wei-Jing Zhux , Ramin Zabih, "Image Indexing Using Color Correlograms,".
[32] S. Mangijao Singh , K. Hemachandran "ContenContent-Based Image Retrieval using Color Moment and Gabor Texture Feature,"in IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012 ISSN 1694-0814.
[33] Nidhi Singhai, Prof. Shishir K. Shandilya, "A Survey On: Content Based Image Retrieval Systems," in International Journal of Computer Applications (0975 – 8887) Volume 4 – No.2, July 2010.
[34] Tienwei Tsai, Yo- Ping Huang, Te-Wei Chiang, " Fast Image Retrieval Using Low Frequency DCT Coefficients ".
[35] Ahmed J. Afifi, Wesam M. Ashour, "Content-Based Image Retrieval Using Invariant Color and Texture Features,"2012 IEEE.
[36] Reshma Chaudhari, A.M.Patil, "Content based image retrieval using color and shape features," IJAREEIE, vol.1, issue 5, November 2012.
[37] Subrahman yam Murala(2009) . "Color and Texture Features for Image Indexing and Retrieval".