IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Volume-5, Issue-6 E-ISSN: 2347-2693

A Study and Analysis on Feature Extraction in Content-Based Image Retrieval

Nadira T[1*], Rehna K[2] and Fepslin Athish Mon[3]

[1*] Royal College of Engineering and Technology, Thrissur, Kerala, India
[2] Royal College of Engineering and Technology, Thrissur, Kerala, India
[3] Royal College of Engineering and Technology, Thrissur, Kerala, India

[*]*Corresponding Author: nadira413@gmail.com*

Available online at: www.ijcseonline.org

Received: 25/May/2017, Revised: 02/Jun/2017, Accepted: 20/Jun/2017, Published: 30/Jun/2017

***Abstract*—** The digital image data is rapidly growing in quantity and heterogeneity. The existing information retrieval techniques does not meet the user's demand, so there is need to develop an efficient system for content based image retrieval. 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 annotated image database according to user's requests, in the form of a query image. In this paper we present a study on some technical aspects of current content-based image retrieval systems and feature extraction. Features such as color, shape and texture are analysed to develop a high retrieval accurate cbir system.

***Keywords*-** CBIR, visual database, texture, feature extraction, color correlogram

I. INTRODUCTION

Content Based Image Retrieval (CBIR) is any technology that in principle helps to manage and organize digital image archives by their visual content [1]. The aim of CBIR is to neglect the use of textual descriptions by retrieving images based on similarities in their contents like textures, colors, shapes etc. are considered. These features are usually called as low level features of image. CBIR is defined as an application of computer vision techniques to the image retrieval problem, that is, the problem of finding of images from large data bases. In Content based image retrieval the search aims at the actual contents of an image rather than keywords, tags, and/or descriptions associated with the image. Low level feature extraction is treated as the milestone of CBIR. Feature extraction may be done from region or an entire image [2]. These conventional approaches for image retrieval are based on the computation of the similarity between the users query and database images.

In many areas of commerce, academia, and hospitals, large collections of digital images are being created. Many of these collections are the product of digitized existing collections of analogue photographs, diagrams, drawings, paintings, and prints. Usually, the only method for searching these collections was by keyword indexing, or simply by browsing. The great problem of searching is solved by introducing content-based image retrieval methods. Searching and browsing became an easier task with CBIR. In CBIR each image stored in the large database, its features are extracted, compared to the features of the query image and similar images are retrieved. Thus, broadly, it involves two processes, feature extraction and feature matching [3].

.

The reminder of the paper is organized as follows: section II presents about material and method. Section III gives the concluding remarks and finally the paper ends with few references.

II. MATERIAL AND METOD

A. Content-based Image Retrieval

Content-based image retrieval (CBIR) is a technique for retrieving images from a huge database on the basis of visual features such as color, texture and shape. In typical content-based image retrieval system (Figure 1), the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database. Whenever a query image is given, its features are extracted and similarity measurement is done along with the feature database. Finally most similar images are retrieved from the database.

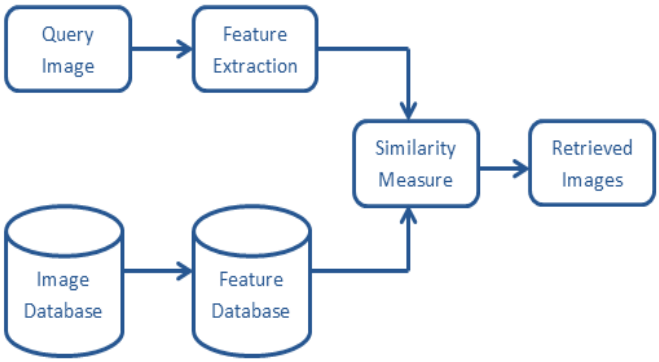

Figure 1: CBIR

B. Feature Extraction

Visual feature extraction is the base of any content-based image retrieval techniques. . In a broad sense, features may include both text-based features (key words, annotations) and visual features (color, texture, shape, etc.). According to the visual feature scope, the features can be further organized as low-level features and high-level features. The selection of features to represent an image is one of the major issue in a CBIR system. Because of perception subjectivity and the complex composition of visual data, representing visual feature using single method is impossible. Multiple approaches are considered for each of these visual features and each of them characterizes the feature from different perspective.

1) Color-Based Retrieval

Color is defined as the property occupied by an object of producing different sensations on the eye as a result of emission or reflection of light on it. Color is one of the major visual features in content-based image retrieval. Usually choosing color features are:

1.1) Color Histogram

The color histogram [4] represents the color distribution using a set of bins. Color histogram are computationally efficient, and generally incurious to small changes in camera position. However, a color histogram provides only a very coarse characterization of an image. Histograms does not involves spatial information. By using color histogram only, images having same color but different texture also retrieved.

1.2) Color Auto Correlogram

Another widely used color feature. A Color Correlogram [5] expresses how the spatial correlation of pairs of colors changes with distance. An *autocorrelogram* captures spatial correlation between identical colors only. As it captures both color and spatial information, retrieval using this feature has more accuracy.

2) Texture based Retrieval

The ability to retrieve images on the basis of texture similarity may not seem very useful. But the ability to match on texture similarity can often be useful in distinguishing between areas of images with similar color (such as sky and sea, or leaves and grass). Texture feature include orientation, directionality, and regularity of an image.

2.1) Bit Pattern Feature

Bit Pattern Feature (BPF), is a texture feature that characterizes the edges, shape, and image contents. BPF is extracted from set of bitmap images by using the advantage of indexing using bit pattern codebook. These feature helps in retrieving images with high accuracy.

3) Shape Based Retrieval

Shape usually related to the specific object in the image, so shape's semantic feature is stronger than texture. By considering texture only retrieval accuracy will be reduced.

III. RESULT AND DISCUSSION

Various methods on feature extraction are analysed based on retrieval accuracy. Table shows the comparison of accuracy by using different features for image indexing.

Features	Accuracy
Color only	low
Texture only	low
Shape only	low
Color, texture and shape	high

IV. CONCLUSION

The purpose of this study is to provide an overview of the functionality and technical aspects of content based image retrieval systems. Most systems use color and texture feature and few systems use shape feature. Studies shows that the retrieved images are closed to the input image from the point of view of color only, when just using the color histogram. When only using texture features, retrieved images closer to input images in shape and edge. By fusing color, texture and shape features, most similar images can be retrieved by inputting an image. Retrieval accuracy is improved by extracting features of color, shape and texture. More low-level features such as spatial location features and clustering techniques will be fused to make the system more robust in the future.

REFERENCES

[1] D. A. Kumar and J. Esther, *"Comparative Study on CBIR based by Color Histogram, Gabor and Wavelet Transform*", Int'l Journal of Computer Applications (0975 – 8887), vol. 17, no. 3, (2011) March, pp. 37.

[2] Khan,W., Kumar,S., Gupta,N., Khan,N., *A Proposed Method for Image Retrieval using Histogram values and Texture Descriptor Analysis*, IJSCE, ISSN: 231-2307, Volume-I Issue-II,(May 2011).

[3] F. Long, H. J. Zhang and D. D. Feng, *"Fundamentals of Content-based Image Retrieval, Multimedia Information Retrieval and Management*", D. Feng Eds, Springer, (2003).

[4] L. Haldurai and V. Vinodhini, "*Parallel Indexing on Color and Texture Feature Extraction using R-Tree for Content Based Image Retrieval",* International Journal of Computer Sciences and Engineering, Vol.3, Issue.11, pp.11-15, 2015.

[5] L. Zheng, S. Wang, and Q. Tian, "*Lp*-norm IDF for scalable image *ns. Image Process.*, vol. 23, no. 8, pp. 3604–3617, Aug. 2014

Authors Profile

NADIRA T. pursued Bachelor of Technology from Calicut University, in 2015. She is currently pursuing Master of Technology under APJ Abdul Kalam Tecnological University, Kerala, India. She has presented 1 paper in international conference. Her main research work focuses on Image Processing

REHNA K. pursued Bachelor of Technology from Calicut University, in 2012. She is currently pursuing Master of Technology under APJ Abdul Kalam Tecnological University, Kerala, India. She has presented 2 papers in international conferences. Her main research work focuses on Image Processing

FEPSLIN ATHISH MON pursued Bachelor of Technology from Anna University, in 2009 and Master of Technology under Manonmaniam Sundarnar University, , in 2011. He has published several papers in reputed journals. His main research work focuses on Digital Image Processing

