Proc. of Int. Conf. on Emerging Trends in Engineering & Technology, IETET

Content based Image Retrieval System: A Review

Radha Rani[1] and Tarun Kumar[2]
[1] M.Tech. Scholar, Department of CSE, GIMT, Kurukshetra, India
Email: radha.gangus15@gmail.com
[2]Asstt. Prof., Department of CSE, GIMT, Kurukshetra, India
Email: tarun.dhiman@gmail.com

***Abstract*—The paper presents a review on content based image retrieval system. Content Based image retrieval is a system by which various images are retrieved from a large database collection. These image databases are prepared using various visual features like color, texture, shape and spatial layout which are extracted using different techniques. The various classification algorithms like K-NN, Naïve byes are used for better performance. The precision rate can be improved by using similarity metric like Jaro winkler for strings similarity. In this paper we discuss the fundamental aspects, visual features and techniques for fast searching and retrieval of images from the database.**

***Index Terms*— Content Based Image Retrieval, k-NN, String similarity, image database.**

I. INTRODUCTION

A. Data Mining

Data mining [1] is an interdisciplinary subfield of computer science. It is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating. Data mining is the analysis step of the "knowledge discovery in databases" process, or KDD.

B. Content Based Image Retrieval

Content Based Image Retrieval (CBIR) is a technique that enables a user to derive an image based on a query, from a database which has a large amount of images. The two methods by which image can be retrieved are Text based image retrieval and Content based image retrieval. From historical perspective, it is noticed that the former image retrieval systems are rather text-based search since the images are required to be annotated and indexed accordingly. However, with the substantial increase of the size of images as well as the size of image database, the task of user-based annotation becomes very cumbersome, and, at some extent, subjective and therefore, incomplete as the text often fails to express the rich structure of the images. To overcome these difficulties, content-based image retrieval was proposed. In place of manually annotated by text-based key words, images would be listed by their own visual content, such as color and texture.

Grenze ID: 02.IETET.2016.5.26

C. Components of CBIR System

The system consists of mainly 3 steps. These are: database classification , feature extraction and similarity measure.

- **Database classification:** In CBIR system, for a given input query image retrieving the relevant images from the large database. In this process we compare input query image with the each image in given database. It takes large amount of time. So, we need to improve the performance of retrieving the relevant images from the huge database, firstly we need to classify the database. The various classifier used are SVM, K-NN, Naïve bayes, decision tree etc.
- **Feature Extraction:** Feature extraction is a process in which we locate an outstanding part, quality and characteristic in a given image. For classification an image, we must first remove some features out of the image. Features can be categories as application independent features such as shape, texture, and color. On the other hand, all features can be coarsely categorized into two levels that is low-level features and high-level features. Low-level features can be extracted directly from the original images, whereas high-level feature extraction must be based on low-level features.
- **Similarity Matching:** The similarity matching phase is the main decision making phase of CBIR system and uses the features extracted in the earlier stage to identify the image. It has two algorithms Euclidean Distance Algorithm, K-Means Clustering Algorithm.

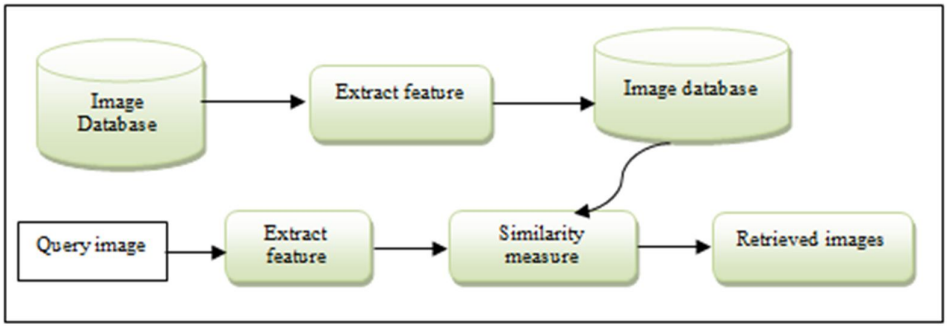

Figure 1: Components of CBIR System

The aim of the retrieval experiments is to measure the retrieval effectiveness between the without database classification and with database classification. The technique is evaluated based on precision and recall. Precision and recall measures have been widely used for evaluating the performance of the CBIR system.
A precision rate can be defined as the number of relevant images retrieved by a search divided by the total number of images retrieved by that search.
A recall rate is described as the number of relevant images retrieved by a search divided by the total number of existing relevant images (which should have been retrieved).

II. Related Work

Hechao Yang, Xuemei Zhou in 2010[4] presented an Image retrieval method based on shape , the description and match based on shape are discussed and compared.
Deepak S. Shete, Dr. M.S. Chavan in 2012 [5] described the fundamental aspects of CBIR. Features for Image Retrieval like color, texture and shape are brought forward. Similarity measures based on which matches are made and images are retrieved. Dimension reduction and indexing schemes are also discussed. Relevance feedback is discussed which helps in improving the performance of a CBIR system.
Sumiti Bansal, Er. Rishamjot Kaur in 2014 [3] presented a content based image retrieval using various techniques such as support vector machine (SVM) that should combine all relevance or irrelevance features such as color, texture, shape, size. SVM is used to find out the maximum result and give fast result as compared to others.
Deepu Rani, Monica Goyal in 2014 [6] presented a content based image retrieval system which uses the contents of the images to represent and access the images. The paper presented a modified SVM technique to derive the images which are similar to the query image.

K. Haridas, Dr. Antony Selvadoss Thanamani in 2014 [7] presented a paper which experiments on various methods available for Content based image retrieval System, they are precision, recall and accuracy value for Bag of Visual words, Color and Edge Directive Descriptors, Fuzzy Color and Texture Histogram. The methods are carried out and tested based on three parameters like precision value, recall value and Accuracy rate. The Experimental results obtained after the result show that FCTH (Fuzzy Color and Texture Histogram) method is more efficient when comparing with other methods.

Kommineni Jenni, Satria Mandala, Mohd Shahrizal Sunar in 2015 [2] presented a new content based image retrieval approach based on the database categories using Support Vector Machine (SVM) and colour string coding feature selection. The image database used in the experiment contains 1800 colour images from Corel photo galleries. This CBIR approach has considerably increased the accuracy in obtaining results for image retrieval. Computational complexity is also decreased in the system.

Ramesh K Lingadalli1, N.Ramesh in 2015 [8] presented a paper which proposed an algorithm which includes all three features such as colour, shape and texture to give the advantages of various other algorithms to enhance the accuracy and performance of retrieval of images.

Miss.Aboli W. Hole1, Prof Prabhakar L. Ramteke in 2015[8] presented a paper which describes research on different feature extraction and matching techniques in designing a Content Based Image Retrieval (CBIR) system. In this paper technique used for retrieving the images based on their content namely dominant color, texture and combination of both color and texture. The technique verifies the superiority of image retrieval using multi-feature than the single feature.

III. Conclusion

Content Based Image Retrieval is an active and rapidly growing research area in retrieving image from a large database. The accuracy of the system can be improved by using better a classification algorithm. General visual features extensively used in content-based image retrieval are color, texture, shape, and spatial information. There are various ways to calculate the similarity measure like Jaro winkler for strings similarity. The effective precision and recall rate is calculated in the system. In future a system should be developed so that it can retrieve the images from a large data set efficiently in minimum amount of time. The authors can conclude on the topic discussed and proposed. Future enhancement can also be briefed here.

References

[1] https://en.wikipedia.org/wiki/Data_mining accessed 5/6/2016.

[2] Kommineni Jenni, Satria Mandala, Mohd Shahrizal Sunar (2015) “Content Based Image Retrieval Using Colour Strings Comparison”, Published by Elsevier B.V., volume 50 on pages 374 to 379.

[3] Sumiti Bansal, Er. Rishamjot Kaur (2014) “ A Review on Content Based Image Retrieval using SVM”, published by International Journal of Advanced Research in Computer Science and Software Engineering.

[4] Hechao Yang, Xuemei Zhou (2010) “Research of Content Based Image Retrieval Technology” , Proceedings of the Third International Symposium on Electronic Commerce and Security Workshops(ISECS ’10) Guangzhou, P. R. China, 29-31,July 2010, pp. 314-316.

[5] Deepak S. Shete, Dr. M.S. Chavan (2012) “Content Based Image Retrieval: Review” , published in International Journal of Emerging Technology and Advanced Engineering.

[6] Deepu Rani, Monica Goyal (2014) “A Research Paper on Content Based Image Retrieval System using Improved SVM Technique” published in International Journal Of Engineering And Computer Science ISSN: 2319-7242.

[7] K. Haridas, Dr. Antony SelvadossThanamani (2014) “Efficient Content Based Image Retrieval System in Visual Words, Color and Edge Directive Descriptors and Fuzzy Color and Texture Histogram”, published by International Journal of Innovative Research in Computer and Communication Engineering.

[8] Ramesh K Lingadalli1, N.Ramesh (2015) “Content Based Image Retrieval using Color, Shape and Texture”, published by International Advanced Research Journal in Science, Engineering and Technology.

[9] Miss.Aboli W. Hole, Prof Prabhakar L. Ramteke (2015) “Content Based Image Retrieval using Dominant Color and Texture features” published by International Journal of Advanced Research in Computer and Communication Engineering.