Reduction in Retrieval Time in CBIR using MapReduce

Krishna N. Parekh[1] Mr. Narendra Limbad[2]
[2]PG Student
[1]Department of Computer Engineering
[1,2]L.J.I.E.T, Ahmadabad, India

Abstract— In CBIR main problem is to extract the image features that effectively represent the contents of an image in a database which may be difficult to use, using single feature. Whereas with combination of different features it would give efficient results. The different combination used maybe color-edge; color-shape-texture; color-texture; shape-texture. With more features different aspects of an image could be represented. These type of combination along with Hadoop MapReduce on distributed database less retrieval time can be achieved. Image retrieval based on color, texture and shape is an emerging and wide area of research scope. In this paper we present a novel framework for combining all the three i.e. color, texture and shape information along with MapReduce and achieve higher retrieval efficiency.

Key words: CBIR, Feature Extraction, Precision, Recall

I. INTRODUCTION

Information Retrieval is the field of knowledge that deals with the representation, storage, and access to information items. Particularly, when the retrieved information is a image or collection of images, this term is known as Image Retrieval. Image Retrieval Techniques started in 1979 when a conference on Database Techniques for Pictorial Applications was held in Florence [1]. Since then, the application potential of image database management techniques has attracted the attention of researchers. Early techniques were based on the textual annotation of images not on visual features. However, there were two significant limitations in the retrieval of images in text based systems. The first was in conjunction to the volume of the database. Manual annotation was such a time-taking and expensive task and it was unsuccessful with large image databases. The second limitation affecting the performance of the system was that the description of the images was found to be a highly subjective task that could generate different text labels to the same image. Such problematic image indexing methods have led to the rise of techniques for retrieving images on the basis of automatically-derived features such as color, texture and shape, generally referred to as content based image retrieval. [2]

A. Keyword Based Image Retrieval or Text Based Approach:

In the 1970s, the Keyword Based Image Retrieval system used keyword as descriptors to index an image [1]. In Fig.1 General Framework of keyword based image retrieval is shown.

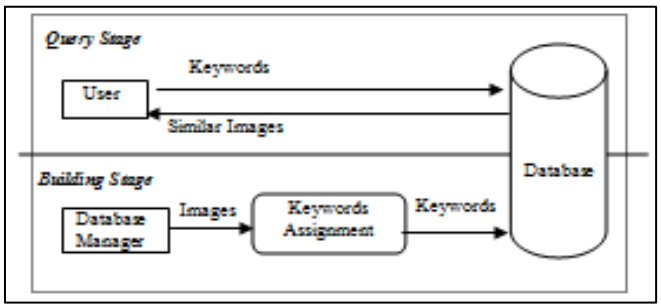

Fig. 1: General Framework of Keyword Based Image Retrieval. [1]

Images stored in the database are examined manually in this technique and most appropriate keywords are assigned to describe their contents. The keywords are stored as the part of the attributes associated to the image. During query stage, user will fulfill the search criteria by providing one or many keywords. Then keyword matching process is performed to retrieve images associated with the keywords provided by the user that matches the search criteria.

B. Content Based Image Retrieval:

In 1980s, Content-based image retrieval (CBIR) was introduced as an alternative to text based image retrieval [1]. Content-Based Image Retrieval is also known as Query by Image Content and Content-based Visual Information Retrieval. In CBIR, retrieval of image is based on similarities in their contents, like, textures, colors, and shapes etc., which are considered as the lower level features of an image. These conventional approaches for image retrieval are based on the computation of the similarity between the user's query and images. In CBIR each image stored in the database, has its features extracted and compared to the features of the query image. Thus, broadly, it involves two processes, viz. feature extraction and feature matching [7].

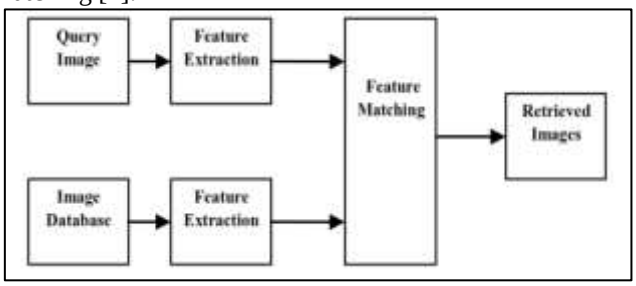

Fig. 2: General Framework of Content Based Image Retrieval [7]

II. FEATURE EXTRACTION TECHNIQUES

Feature Extraction Techniques may include both text based features and visual features. Within CBIR visual feature are required to extract. In the visual features scope it can be classified as low level and high level features. The selection

of the features to represent an image is one of the keys of a CBIR system. Because of perception subjectivity and the complex composition of visual data, there does not exist a single best representation for any given visual feature [4]. Multiple approaches have been introduced for each of these visual features and each of them characterizes the feature from a different perspective [4].Main three low level features are the Color, Texture and Shape.

A. Color Feature:

Color features are the basic characteristics for the content of images. With the color feature human can identifies and distinguish between object and images. Colors are used in image retrieval because they are powerful descriptors and sometimes provide powerful information about images. To extract the color features from the content of an image, we need to select a color space and use its properties in the extraction. In common, colors are defined in three dimensional color spaces. In digital image purposes, RGB color space is the most prevalent choice [2].

Low level feature	Methods
Color	Color histogram
	Conventional color histogram
	Invariant color histogram
	Fuzzy color histogram
	Geometric moments
	Average RGB
	Color moments
	Color correlogram
	Color coherence vector

Table 1: Methods of Color Feature

B. Texture Feature:

There is no formal definition for texture, but it can say that it provides the measure of properties such as smoothness, coarseness, and regularity. In addition, texture can be expressed 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 [2]. Because of there is no formal mathematical definition for texture, many different methods are proposed for computing texture but among those methods, no single method works best with all types of texture.

Low level feature	Methods
Texture	Discrete wavelet transform
	Gabor Wavelet Transform
	Haar Discrete Wavelet Transforms
	Ranklet Transform
	Discrete Fourier Transform
	Discrete cosine transform
	Hadamard Transform
	Gaussian Pyramid
	Laplacian Pyramid
	Steerable Pyramid
	Gabor Filter

Table 2: Methods of Texture Feature

C. Shape Feature:

Another important visual feature is Shape. Shape is the basic features used to describe image content. Shape's representation and description is a difficult task because when a 3-D real world object is projected onto a 2-D image plane, one dimension of object information is lost. As a result, the extracted image is only partially represents the projected object. To make the problem even to difficult, shape is often corrupted with noise, defects, arbitrary distortion and occlusion. As a result, shape properties play an important role in content based image database systems devised by computer vision researchers [14].

The reason for choosing shape feature for describing an object is because of its inherent properties such as identifiability, affine invariance, and reliability and occlusion invariance, thus shape of the object has proved to be a promising feature based on which several image classification and retrieval operations can be performed [14]. The shape descriptor are classified into two major kind namely Contour-based shape representation and description techniques and Region-based shape representation and description techniques.

Low level feature	Methods
Shape	Fourier descriptors
	CSS descriptors
	Zernike moments
	Grid descriptors

Table 3: Methods for Shape Feature

III. PROPOSED ALGORITHM

By using combination of all three feature extraction techniques along with hadoop mapReduce can reduce retrieval time as well as increases precision for retrieved images.

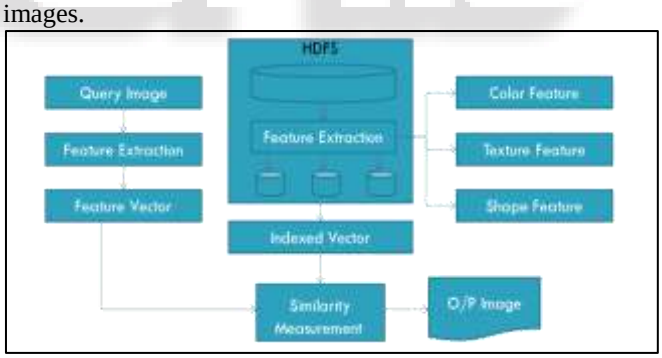

Fig. 3: Flow Diagram of Proposed Model

The proposed system will also tend to decrease the computational time as well as increase the precision and recall.

Input: An image, WANG image dataset
Output: n similar images to the input image
Algorithm:
Offline phase:

1) Centralized database is coverted into distributed database using MapReduce.
2) Feature extractions of each image in database.

A. Extract Color Feature by Using Color Moment

– Construct The Color Feature Vector.

B. Extract texture features using Discrete Fourier Tranform technique.

– Construct The Texture Feature Vector.

C. Extract Shape Features Using Zernike Moment

– Construct The Shape Feature Vector.

D. Real-Time Phase:

1) Extract the features of input image using step 2 procedure.
2) Compare the color feature of input image and database images and store result in the feature vector FC.
3) Compare the texture feature of input image and database images and store result in the feature vector FT.
4) Compare the shape feature of input image and database image and store result in the feature vector FS.
5) In Euclidean three-space, the distance between points (x1,y1,z1) and (x2,y2,z2) is

$$\sqrt{(x2-x1)^2+(y2-y1)^2+(z2-z1)^2}$$

6) Give the rank to each retrieved image
7) As per predefined relevance rank criteria display result.

IV. IMPLEMENTATION RESULTS

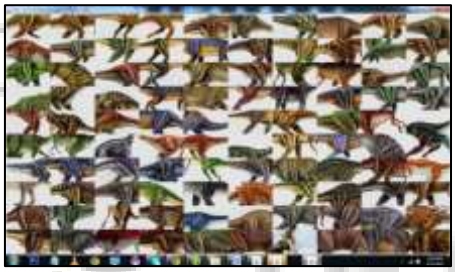

Fig. 4: Retrieved Images Using Combination of All Three Methods

Figure shows all the relavent images retrived for the query image using combination of all three methods of feature extraction ie color moment for color feature extraction , DFT fot texute feature extraction and zernike moments for shape feature extraction by extracting all three features of the query image as well as database images.

Wang Categories	Color	Texture	Shape	Combined
Africans	0.39	0.28	0.31	0.52
Beaches	0.37	0.29	0.32	0.48
Buildings	0.42	0.32	0.38	0.53
Buses	0.48	0.38	0.43	0.58
Dinosaurs	0.98	0.87	0.94	0.99
Elephants	0.54	0.48	0.59	0.63
Flowers	0.62	0.55	0.58	0.68
Horses	0.53	0.39	0.45	0.59
Mountains	0.43	0.32	0.41	0.49
Foods	0.39	0.29	0.33	0.48

Table 4: Comparison Table

Above table shows precision for different methods as well as combination of all three methods.

REFERENCES

[1] Hui Hui Wang, Dzulkifli Mohamad and N.A. Ismail "Approaches, Challenges and Future Direction of Image Retrieval " journal of computing, volume 2, issue 6, june 2010, issn 2151-9617,pp. 193-199.
[2] Ahmed J. Afifi and Wesam M. Ashour "Content-Based Image Retrieval Using Invariant Color and Texture Features" 978-1-4673-2181-5/12/$31.00 ©2012 IEEE.
[3] T. Dharani and I. Laurence Aroquiaraj "A Survey on Content Based Image Retrieval " 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, February 21-22 978-1-4673-5845-3/13/$31.00©2013 IEEE.
[4] Amandeep Khokher, Rajneesh Talwar, "Content-based Image Retrieval:Feature Extraction Techniques and Applications" International Conference on Recent Advances and Future Trends in Information Technology (iRAFIT2012),pp.9-14.
[5] S. Mangijao Singh and K. Hemachandran, "Content based Image Retrieval based on the Integration of Color Histogram, Color Moment and Gabor Texture" International Journal of Computer Applications (0975 – 8887) , Volume 59– No.17, December 2012.
[6] P.S.Suhasini , Dr. K.Sri Rama Krishna, Dr. I. V. Murali Krishna "Cbir Using Color Histogram Processing " Journal of Theoretical and Applied Information Technology, Vol6. No1. pp 116 – 122.
[7] Aman Chadha,,Sushmit Mallik and Ravdeep Johar "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval". International Journal of Computer Applications (0975 – 8887), Volume 52– No.20, August 2012
[8] Nadia Baaziz, Omar Abahmane and Rokia Missaoui, "Texture Feature Extraction In The Spatial-Frequency Domain For Content-Based Image Retrieval,"
[9] Mangimala Singha and K.Hemachandran, "Content Based Image Retrieval using Color and Texture," Signal & Image Processing : An International Journal (SIPIJ) Vol.3, No.1, February 2012, pp.39-57
[10] Vibha Bhandari1, Sandeep B.Patil ,"Comparison of CBIR Techniques using DCT and FFT for Feature Vector Generation, " International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 1, Issue 4, November – December 2012, pp. 90-96.
[11] Ahmed J. Afifi and Wesam M. Ashour, "Image Retrieval Based on Content Using Color Feature," International Scholarly Research Network, ISRN Computer Graphics, Volume 2012, Article ID 248285, 11 pages.
[12] Swapna Borde and Udhav Bhosle, PhD. "Feature Vectors based CBIR in Spatial and Transform Domain," International Journal of Computer Applications (0975 – 8887), Volume 60-No.19, December 2012, pp. 34-45.
[13] Shereena V.B.and Julie M. David "Content Based Image Retrieval : A Review" pp. 65-77.
[14] Sai Anand.C, Tamilarasan.M and Arjun.P "A Study on Curvature Scale Space " International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Special Issue 3, July 2014,pp.168-174.

