INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

REVIEW ON CONTENT BASED IMAGE RETRIEVAL BY USING VISUAL SEARCH RANKING

MS. PRAGATI ASHOK DEOLE, PROF. PRADHNYA KAMBLE

M. Tech. (CSE) in G. H. Raisoni Academy of Engineering and Technology, Nagpur (MS) India.

Abstract: Image Retrieval system is an effective and efficient tool for managing large image databases. Content based image retrieval system allows the user to present a query image in order to retrieve images stored in the database according to their similarity to the query image. 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 image database according to user's requests in the form of a query. In this paper content based image retrieval method is used retrieve query image from large image database using three features such as color, shape, texture etc. The main objective of this paper is classification of image using KNN Algorithm. MKNN is the traditional KNN algorithm. Modified K-Nearest Neighbor (MKNN) can be considered a kind of weighted KNN so that the query label is approximated by weighting the neighbors of the query. Visual content is Method of content based image retrieval. Visual search is search by browsing, User submitting keyword for object of interest. then system return visual result. This paper also concentrates identifying the unlabeled images with help of MKNN algorithm.

Keywords: Image Retrieval, Content based image retrieval, Image Classification, KNN, MKNN, Unlabeled image.

Corresponding Author: MS. PRAGATI ASHOK DEOLE

Access Online On:

www.ijpret.com

How to Cite This Article:

Pragati Deole, IJPRET, 2014; Volume 2 (9): 1130-1137

PAPER-QR CODE

1130

INTRODUCTION

Due to the advances in digital photography, storage capacity and networks speed, storing large amounts of high quality images has been made possible. Digital images are used in a wide range of applications such as medical, virtual museums, military and security purposes, and personal photo albums. However, users have difficulties in organizing and searching large numbers of images in databases, as the current commercial database systems are designed for text data and not well suited for digital images. Therefore, an efficient way for image retrieval is desired (M. Vetterli, Feb. 2002). An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Image retrieval is classified into two types of retrieval are Text Based Image Retrieval and Content Based Image Retrieval. Text Based Image Retrieval is having demerits of efficiency, lose of information, more expensive task and time consuming (T.Dharani, I. Laurence Aroquiaraj,2013) Overcome these problems by using Content Based Image Retrieval (CBIR) system for image retrieval. The image retrieval system acts as a classifier to divide the images in the image database into two classes, either relevant or irrelevant there are two classification stage supervised classification and unsupervised classification The supervised classification is finding the labeled images very well. The unsupervised classification is used to finding the unlabelled image. Visual content is Method of content based image retrieval. Visual search is search by browsing, User submitting keyword for object of interest. then system return visual result (V.Suryanarayana, N. R. Rao,2012).

II. MATERIAL AND METHOD

1) Content Based Image Retrieval

In CBIR systems, image processing techniques are used to extract visual features such as color, texture and shape from images. The system uses a query model to convert the image into an internal representation of query, based on features extracted from input images. A retrieval model performs image retrieval by computing similarities between images in object and the query image. Content Based Image Retrieval (CBIR) is defined as a process to find similar image in the image database when a query image is given.
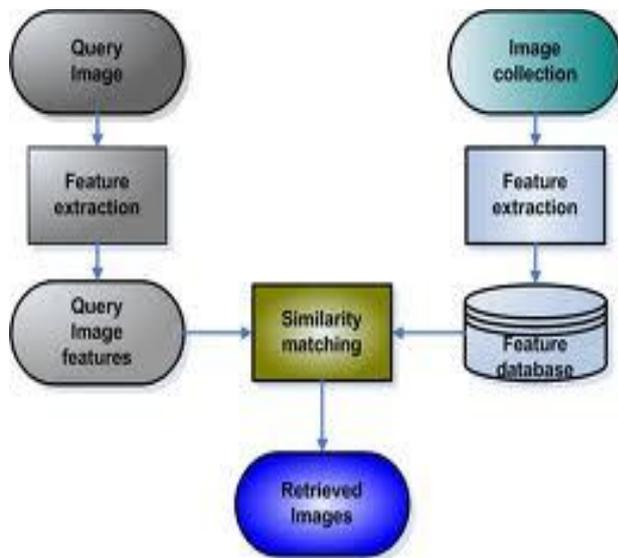
Fig1. Block diagram of content based image retrieval

Figure1 show the basic block diagram of content based image retrieval. Textural features are extracted for both query image and images in the database (C. Singh and Pooja, 2012) The distance between the features vectors of the query image and database are then computed and ranked. The database images that have highest similarity to the query image are retrieved. Then the performance analysis is carried out using precision and recall. Content based image retrieval is working with different types of image database. All databases are having two types of images like labeled and unlabeled. (S. Kumar, N. Gupta and N. Khan 2011)

A. Image Database

Content Based Image Retrieval systems are working with various types of image file formats and also different types of size, themes, features, information, etc. The image file formats are

 TIFF 6.0 (Tagged Image File Format)

 GIF 89a (Graphics Interchange Format)

 JPEG (Joint Photographic Expert Group)

 JFIF (JPEG File Interchange Format)

 JP2-JPX

 JPEG 2000

 Flashpix1.0.2

 Image Pac, Photo CD

PNG 1.2 (Portable Network Graphics)

PDF 1.4 (Portable Document Format)

B. Feature Extraction

Feature means characteristics of object. Feature extraction is refers that dimensionality reduction of that object. It plays an important role in image processing. Features are classified into three types in image processing, that is low, middle and high. Low level features are color, texture and middle level feature is shape and high level feature is semantic gap of objects. These three features are extracted by different techniques. Color feature is extracted by Color Histogram and Color Descriptor. Shape feature is extracted by Edge detection Method. Texture feature is extracted by Gray Level co-occurrence matrix and texture descriptor.

C. Color Feature

Color feature is most common feature of image. The color images are having the standard Color is RGB color. Color histograms are commonly use content based image retrieval.

D. Similarity Measures

CBIR system frequently used similarity measures for an accuracy of image retrieval from large image database. CBIR system frequently used similarity measures for an accuracy of image retrieval from large image database. The measures are sometimes similar or otherwise dissimilar between two images. These distance measures are mainly used for calculate of similarities Distance between images. Content Based Image Retrieval includes the standard measures the measures are Recall, Precision, and Fallout. These measures are used to improve the image retrieval process with various measures. The similarity measure is done by Euclidean distance.

Euclidean distance is measure of distance between two similar images. The Euclidean distance is calculated between the query image and every image in the database. This process is repeated until all the images in the database have been compared with the query image. Similarity measures are very useful to calculate the different object of an image for accuracy and fast image retrieving.

2) Feature Extraction Method

A) Color Histogram

1. RGB histogram

RGB is the acronym for Red, Blue and Green. This histogram is the most used histogram in computer graphics and it uses its red, green and blue components to create a new color. This color histogram combines red, green and blue colors to form a new color. To form a new color it is necessary to increase the values of one or more of the components of the RGB components. The RGB color histogram can be represented on the Cartesian coordinate system, which is illustrated in Figure 1. As it shows a new color can be formed by changing the values of the RGB component. For example, if the magenta color is needed, it is necessary to combine green and blue to create it.
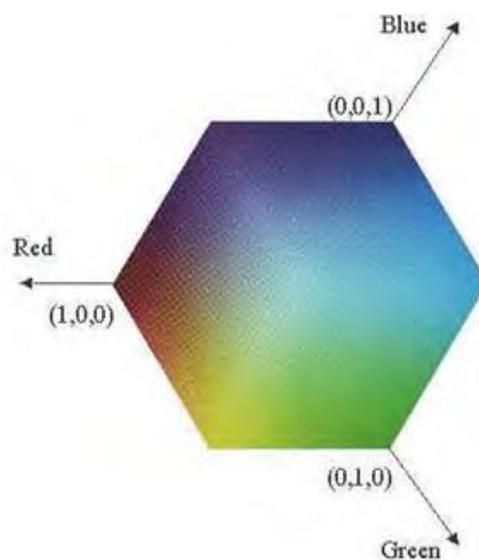

Fig 2. RGB Color Space

Figure 2. RGB coordinate system where the "x" axis represents red colors, "y" axis represents green colors and "z" represents blue colors [10]. As this figure shows changing the values of RGB component can form a new color.

2. HSV histogram

The HSV color space (Hue, Saturation, Value) is often used by people who are selecting colors (e.g., of paints or inks) from a color wheel or palette, because it corresponds better to how people experience color than the RGB color space does. As hue varies from 0 to 1.0, the corresponding colors vary from red through yellow, green, cyan, blue, magenta, and back to red, so that there are actually red values both at 0 and 1.0. As saturation varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value, or brightness, varies from 0 to 1.0, the corresponding colors

become increasingly brighter. The following Fig. 3 illustrates the HSV color space.
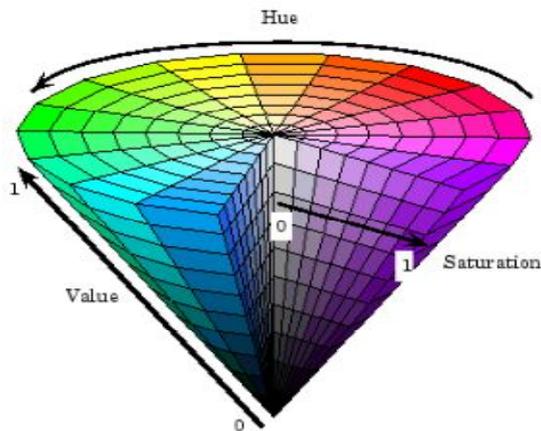

Fig 3. HSV Color Space

B) Edge Detection on Images

Shape features are extracted by Edge detection method .In Edge detection method, we considered two methods that are Sable and canny. Shape is a very powerful feature. Two dimensional Shapes can be described by two different ways such as external representation and internal representation. Boundary based representation is called the External representation which classified into two categories, spatial and transform. Region based representation is also called as Internal representation.

C) Gray Level Co-Occurrence Matrix

A statistical method for examining the texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. Many Statistics such as Contrast, Homogeneity, Correlation and Energy are derived from the GLCM. These characteristics provide the details about the Texture of an image. Contrast measures the local variations in the gray-level co-occurrence matrix. Correlation measures the joint probability occurrence of the specified pixel pairs. Energy provides the sum of squared elements in the GLCM also known as uniformity or the angular second moment. Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Comparing to Texture analysis statistics and Texture filter characteristics, texture analysis method provides the shape feature characteristics.

IV. RESULT & DISCUSSION

K-Nearest Neighbors Algorithm also known as KNN, is a supervised machine learning method, which classifies the data with this classifier a set of data can be classified in order to discover which elements are from the set of data. The purpose of this algorithm is to classify a new object based on attributes and training samples. An existing CBIR system working with various type of image classification methods, they are D-EM(Discriminate EM) Algorithm using the approach Expectation- Maximization (EM) with unlabeled image as data set. Support Vector Machines using classification of image in image database.

In proposed system using Modified K-Nearest Neighbor Algorithm for image classification in Content Based Image Retrieval system. Algorithm consists of two part are Training and validation.

Modified K-Nearest Neighbor Algorithm classified image in two form label image and unlabeled image. Content based image retrieval retrieve image is base on feature extraction such as color, shape, texture. For image classification are used for KNN Algorithm.

V. CONCLUSION

In this paper, CBIR is technique for retrieval query image from large image database, Content based image retrieval retrieve image from large image database using three Method such as color, shape and texture. Color features are extracted using color histogram by calculating the RGB pixel count, HSV pixel count, Shape features are extracted using Edge detection, similarly texture feature which is extracted by texture analysis method, K-Nearest Neighbor Algorithm is used for image classification. In the future we will use MKNN Algorithm.

MKNN Algorithm is use for image classification in Content Based Image Retrieval system. CBIR system used to identifying unlabeled image, interesting visual features of image and providing label with help of MKNN classification by using the user input. We will also improve the performance of CBIR system using MKNN image classification algorithm.

VI. REFERENCES

1. International Journal of Future Computer and Communication, Vol. 1, No. 4, December 2012.

2. T. Dharani, I. Laurence Aroquiaraj, "A Survey on Content Based Image Retrieval," IEEE-2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), PRIME 2013, 978-1-4673-5845-3/13/2013

3. Suchismita Das, Shruti Garg, G. Sahoo, "Comparison of Content Based Image Retrieval Systems Using Wavelet and Curvelet Transform," The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.4, (August 2012).

4. A Vadivel, A K Maunder, ShamikSural C. Singh ,and Pooja , An effective image retrieval using the fusion of global and local transforms based features, Optics & Laser Technology 44 , 2249–2259,2012.

5. V. Suryanarayana, N. R. Rao, P. B. Reddy and G. R.Babu, Image Retrieval System Using Hybrid Feature Extraction Technique, IJCSIT, Vol 4, No 1, and 2012.

6. W. Khan, S. Kumar, N. Gupta and N. Khan, Signature Based Approach For Image Retrieval Using Color Histogram And Wavelet Transform, IJSCE, Vol-1, Issue-1, 2011

7. M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.