IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 2, Ver. I (Mar-Apr. 2016), PP 76-80
www.iosrjournals.org

Sketch Based Image Indexing and Retrieval

Deepali Verma [1], Mohd. Saif Wajid[2]
[1]*Pursuing Master of Technology, Computer Science-Software Engineering, BBDU Lucknow, India.*
[2]*Assistant Professor, Computer Science and Engineering, BBDU Lucknow, India.*

Abstract : *Content based image retrieval (CBIR) is the technology widely used in present era. The main purpose of the CBIR based systems is to excerpt visual features of an image like color, texture, shape or any combination of them. In the previously existing systems, images are manually annotated with keywords and then retrieved using text-based search methods. The proposed system provides a unique scheme for Content based Image Retrieval using sketches. In this system the search was done using the free hand sketches as an input and the desired colored images was retrieved from the database as the output. The existing method specifies the possible solution of how a task specific descriptor, which can handles an information gap between the sketch and the colored images which result an efficient search for the user.*
The Sketch based Image Retrieval system can be used in many areas, some applications of SBIR are social sites, image based digital libraries, and any illiterate person can use this system very efficiently for different purposes.
Keywords: *CEDD, Clustering, K-Mean Clustering, Descriptors, Texture, Quantization, Fuzzy Linking etc.*

I. Introduction

When we talk about digital images over the World Wide Web, it is popular to everyone that there could be hundreds or thousands of users working with digitally used information. This digital information can be in the form of digitalized images as we know that the images are one of the best ways of sharing information, better understanding and memorizing the information more easily.
Text-based image retrieval methods were widely used for conventional database applications in world wide. They were used with business applications purpose and tasks but increasing usage and volume of digital images created accuracy and performance issues generated for text-based image retrieval methods. Thus, a new direction towards better image retrieval with accuracy and performance was followed by researchers from different application domains to take image retrieval technology to the next stage. The new latest generated methods proposed for image retrieval also considered the fundamental basic properties such as color, texture, and shapes of objects in an image. [8]

In mostly cases when we are looking for an efficient searching than we have to be recalled some previous data related to that and as we know that human are a good observer, they are able to recall the visual information more easily and frequently than remembering the whole text as it is but we are able to recall the small things in an image, for example the outlines of an object, or combination of colors and objects formation in an image can easily be remembered by anyone. Since the human is of visual type, they look for images using other images, and follow this approach also at the categorizing. Our purpose is to develop an interactive content based image retrieval system which can retrieve the digital images using the free hand sketches from the database. In this sketch based Image retrieval system the user draws the free hand sketches and blobs on the drawing areas then these sketches are divided into the blocks and the color, texture and shape etc features were determined. The blocks and sub-blocks were also used in the other algorithms like Edge Histogram descriptor, Histogram of orientation etc which was used for matching the images, After matching the system will retrieve the related images stored in database and the clusters of similar images was formed using the appropriate algorithm, from the set of the clustered images the most appropriate digital image was retrieved as an output to user. Using this Sketch based System can play a vital role in many areas of life. The interaction between the user and Sketch based image retrieval system can help in achieving better retrieval results. [1]

II. Literature Review

I go through various research papers for literature review in which the different technique are used for the content based image retrieval that are:

2.1 Sketch4match– Content-Based Image Retrieval System Using Sketches

In this paper the researcher had introduced the problems and challenges which is based and concerned with the design and formation of developed CBIR systems, which is based on a roughly drawn free hand sketch

and with the help of the existing methods, it describe a solution which can handle the informational gap between a handmade sketch and a colored images, making an opportunity for the efficient search hereby. [1]

2.2 Query by Image and Video Content (QBIC)

QBIC (Query by Image Content) System was the first commercial system that offers color, layout, texture, shape and keyword based image retrieval. This system serves as the source for the development and enhancement of new research. To achieve this, QBIC has two provided two main components i.e. Database Population and Database Query. [2]

2.3 Img (Rummager): An Interactive Content Based Image Retrieval System

In this paper researcher describe the application that can execute the image search based on the query image, either from the XML based index file, or directly from the folder containing image files, extracting the comparison feature in real time. In addition this application can execute a hybrid search of images from application server, containing keyword information and visual similarity. [7]

2.4 Img (Anaktisi): A Web Content Based Image Retrieval System

This paper introduces that an img (Anaktisi) is efficient and well developed basic web content based image retrieval system which presents various ways to retrieve images from different databases using a combination of effective descriptors to the users. In order to improve the retrieval results of an image the Img(Anaktisi) employs a new and relevant algorithm i.e. an Automatic Relevance Feedback algorithm.[4]

2.5 MIRROR: An Interactive Content Based Image Retrieval System

In this paper, a content based image retrieval system was described, called MPEG-7 Image Retrieval Refinement based On Relevance feedback (MIRROR), is developed for MPEG-7 visual descriptors evaluation and for developing new retrieval algorithms. This developed system is based on MPEG-7 Experimentation Mode (XM) with the efficient web-based user interface for query by image example retrieval. A new approach was developed for MPEG-7 dominant color descriptor similarity measure which is known as Merged Color Palette and also developed relevance feedback approach in this system. [3]

2.6 Searching Images with MPEG-7 (& MPEG-7-like) Powered Localized descriptors: The SIMPLE Answer to Effective Content Based Image Retrieval.

In this paper researcher propose and evaluate a new technique that confines the description ability of the MPEG-7 like global descriptors. This system employs the SURF detector to identify relevant image patches of blob-like textures and use the MPEG-7 Color Layout (CL), MPEG-7 Scalable Color (SC), and MPEG-7 Edge Histogram (EH) descriptors and the MPEG-7 like Color and Edge Directivity Descriptor, to produce the final local features' vectors. [5]

III. Approaches For Image Retrieval

Most Image Retrieval systems adopt the following two-step approach to search image from databases:

a. ***Indexing:*** For each image in a database, a feature vector is computed for capturing the essential properties of an image and stored them in a feature base.

b. ***Searching:*** The user had to given a query image, its feature vector is computed, compared to the feature vectors generated in the feature base, and those images which are most similar to the query image are shown to the user.

IV. Methodology

The working of the system was described as per the methodology mentioned; this system was started as per the sketch or the query image as the input to the system which gives the reference image to the system i.e. with reference image input the similar images are searched. In this system the Indexing was done using the Document Builder Interface.

The user formulates the query. Then the retrieval system access the index, a data structure for efficient retrieval, where all documents from the corpus are indexed. In the retrieval there are no perfect or all-true results, but so often we always search for adequate search or which is referred best under the circumstances. Or we can say this, being **somewhat relevant** is always better than nothing. This measure of relevance to the user is called relevance function. It decides numerically how relevant a document is for the user judging on the user query. The most relevant results are then presented to the user as the output of the systems in the clustered form of a query images.

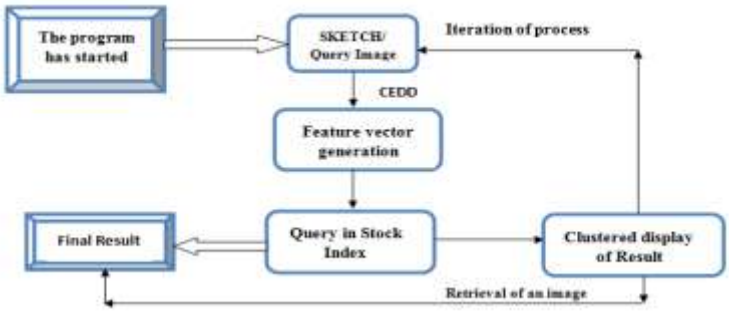

Fig. 1 Methodology of the system

V. Proposed Algorithms:

5.1 K-means Clustering Algorithm

Clustering: The process of organizing objects into groups whose member is similar in some way.
The similarity criterion is Distance: two or more object belongs to same cluster if they are close according to a given distance.

The K-Means algorithm is an unsupervised clustering algorithm that classifies the input data into multiple classes which is based on their inherent distance from each other. The algorithm has some predefined assumptions that the data features form a vector space in which it tries to find natural clustering amongst them. The points are clustered around medians μ_i $\forall$i=1, 2 … k which are obtained by minimizing the objective:

$$V = \sum_{i=1}^{k} \sum_{x_j \in s_i} (x_j - \mu_i)^2 \qquad (1)$$

Where there are k cluster s_i, i=1, 2… k and μ_i is the median or mean point of all the points $x_j \in s_i$.
As a part of this project, an interactive version of this algorithm is implemented. The algorithm takes two dimensional image matrixes as input. Various steps in this algorithm are as follows:

1. Compute the intensity distribution or we can say histogram of the intensities in images.
2. Initialize the median with k random intensities.
3. Repeat the steps until the cluster labels of the image do not change anymore.
4. Cluster the points in an images based on the distance of their intensities from the median intensities.

$$c^{(i)} = \arg\ min_j\ ||x^{(i)} - \mu_j||^2 \qquad (2)$$

5. Compute the new median for each of the clusters.

$$\mu_i := \frac{\sum_{i=1}^{m} 1\{c_{(i)}=j\}x^{(i)}}{\sum_{i=1}^{m} 1\{c_{(i)}=j\}} \qquad (3)$$

Where k is a parameter of the algorithm (the number of cluster to be found), i iterates over the all intensities, iterates over the entire median and μ_i are the medians intensities. [11]

5.2 Color and Edge Directivity Descriptor (CEDD)

Color Unit: Color Unit is the unit associated with the extraction of color information in an image.
Texture Unit: Texture Unit is the unit associated with the extraction of texture information in an image.
Color and Edge Directivity Descriptor: CEDD integrate color and texture information in a single histogram. [10]
CEDD size is limited to 54 byte per image; this descriptor is widely used in large image databases. One of the most important attribute of the Color and Edge Directivity Descriptor is its low computational power needed for its extraction, in comparisons with the needs of the most MPEG-7 descriptors. The working of the CEDD descriptor was described by the following figure.

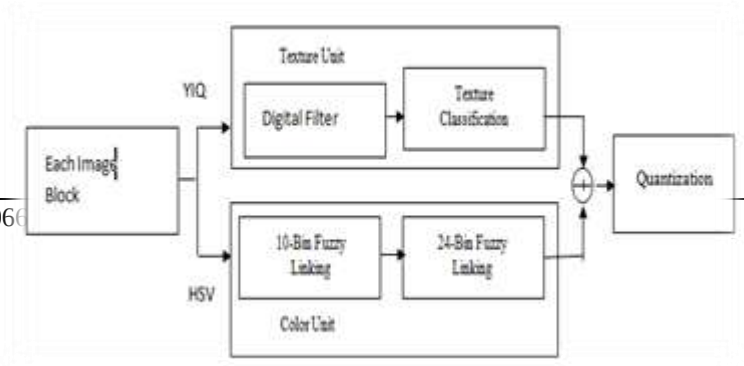

Fig. 2 Implementation of CEDD

5.3 Edge Histogram Descriptor

Histogram: The Histogram is the most commonly used to graphically represent any global feature composition of a digital image. Histogram is useful for indexing and retrieval of digital images.

Edge: Edges in image comprise an important feature in order to represent their content.

For representing an important edge feature in an image, is to use a histogram space which represents the change in the frequency and directionality of brightness in the image, it is called as Edge Histogram. The Edge Histogram Descriptor (EHD) represents the spatial distribution of edges in an image. This mentioned extraction process of the EHD consists of the following stages:

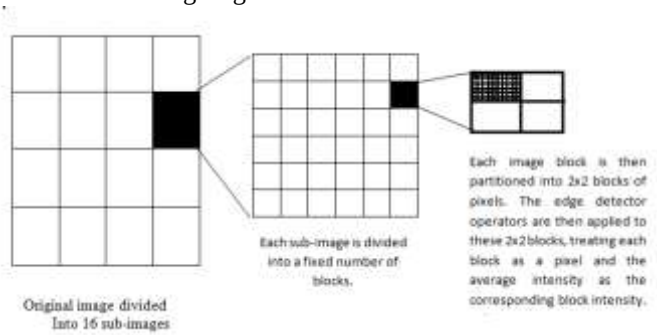

Fig. 3 Edge Histogram Descriptor

- Each image was consisting of an array which is divided into 4x4 sub images.
- Each and every sub image is further divided into non-overlapping matrix image blocks whose size depends on the resolution of the image.
- Edges contain in each image-block are categorized into one of the following six types of histograms which includes: 45± diagonal, 135± diagonal, vertical, horizontal, non-directional edge and no-edge.
- Then a 5-bin edge histogram of each sub image can be obtained.
- From edge histogram each bin value is normalized by the total number of image-blocks in the sub image.
- Then these normalized bin values are nonlinearly quantized.

VI. Conclusion

Developing a practical image retrieval system is still a challenging task. The accuracy and speed are still two key issues in this field. To achieve the goal, we propose a sketch based system for image retrieval.

From among the objectives of this paper performed to design, implement and test a sketch-based image retrieval system. The main aspects were taken that the retrieval process must be highly interactive and unconventional. The actual work with CEDD demonstrates that this standard provides an extensive set of attributes to describe visual content, which is exactly its purpose. In this system CEDD (Color and Edge Directivity Descriptor) uses EHD (Edge Histogram Descriptor) descriptor in order to improve the performance of edge detection based method and makes it more efficient for better performance of the whole system.

The used k-means algorithm is independent on the feature extraction algorithm and is used as a post-processing step in retrieval. The improvements applied in selecting the neighborhood vertices of the retrieval results from tradition image retrieval system in image feature space could also improve the performance of the system.

References

[1] B.Sz´ant´o, P. Pozsegovics, Z. V´amossy, Sz. Sergy´an: "Sketch4Match– Content-based Image Retrieval System Using Sketches", 9th IEEE International Symposium on Applied Machine Intelligence and Informatics, January 27-29, 2011.
[2] M. Flickner, H.S. Sawhney, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," IEEE Computer 28, Sept. 1995.
[3] Wong, K.-M., Cheung, K.-W., Po, L.-M.: "MIRROR: an interactive content based image retrieval system." In: Proceedings of IEEE International Symposium on Circuit and Systems 2005.
[4] Konstantinos Zagoris, Savvas A. Chatzichristofis, Nikos Papamarkos and Yiannis S. Boutalis "Img (Anaktisi): A Web Content Based Image Retrieval System", IEEE 2009.
[5] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan and A. Yamada, "Color and texture descriptors", IEEE Trans. Circuits Syst. Vid. Technol. 11(6) (2001).
[6] R. Datta, D. Joshi, J. Li, J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age", ACM Computing Surveys 40 (2) (2008) 1–60.
[7] C. Iakovidou, N. Anagnostopoulos, A. Ch. Kapoutsis, Y. Boutalis, S. A. Chatzichristofis "Searching Images with MPEG-7 (& MPEG-7-like) Powered Localized descriptors: The SIMPLE answer to effective Content Based Image Retrieval".
[8] S. A. Chatzichristofis, Y. S. Boutalis and Mathias Lux, "IMG (RUMMAGER): An interactive Content based image retrieval system", 2nd Int. Workshop on Similarity Search and Applications (SISAP), Czech Republic 2009.
[9] Babita Singh, Waseem Ahmad: "Content Based Image Retrieval:A Review Paper", IJCSMC,Vol.3, Issue. 5, May 2014, ISSN 2320–088X.
[10] Savvas A. Chatzichristofis and Yiannis S. Boutalis,"CEDD:Color and Edge Directivity Descriptor.A Compact Descriptor for Image Indexing and Retrieval", Democritus University of Thrace, Greece, 2010.
[11] Suman Tatiraju, Avi Mehta, "Image Segmentation using k-means clustering, EM and Normalized Cuts".