RESEARCH ARTICLE OPEN ACCESS

# Content Based Image Retrieval Using SVM

Priyanka Sharma
Department *Computer Science and Engineering*
*Advanced Institute of Technology and Management*
*Palwal - India*

## ABSTRACT

Content based image retrieval has been emerged as the most innovative area for the research work from the last few years. Image retrieval is basically a method for the extraction of most crucial characteristics of an image. The primary aim of this content based image retrieval (CBIR) is to have the exact result within the short period of time. Content-based image retrieval is based on the concept of using the depiction of characteristics which are retrieved from the images themselves. Maximum of the present CBIR systems works on querying-by-example, an approach in which a portion of the image or the whole image is chosen by the user in the form of query. The technique works by retrieving the feature of the image which is given as a query then going through the database for the image having similar features as the image in the query and extracts the appropriate images to the user which matches most appropriately to the query. In this area, content in itself constitutes the other features also like the texture of the image, color of the image, kinesthetic relationships and shape of the image. A number of CBIR systems have been studied earlier which makes the use of these features for the purpose of comparison, analyzing and extraction of image. There are also some systems that have also been successful by integrating the both content-based retrieval as well as the text-based retrieval technique. In all situations, however, it has been not possible to conclude that which particular feature is evolved as the most important for imparting the best extraction. Many researchers used various techniques for the classification of huge amount of data. In this paper we emphasize on a SVM (Support Vector Machine) technique for the purpose of retrieval of images similar to the image in the query. In this paper we primarily present the importance or necessity of support vector machine. The traditional content –based image retrieval techniques may not work as required because of some problems in real world application. For example all users would like to get their desired result in the first time on the web, possibility of consuming a lot of time while searching for the results. In this paper we contemplate over the mechanism of support vector machine (SVM) and also explain some other methods for extracting images.

***Keywords:-*** CBIR (Content Based Image Retrieval), feature extraction, SVM (Support vector Machine), image retrieval, image classification.

## I. INTRODUCTION

Content-based retrieval working is based on the usage of contents for the retrieval of the image as well as for its representation. A traditional CBIR system works by using the technique of online image retrieval and off-line trait mining. The off-line feature extraction method works by retrieving the visual characteristics like colour, shape and texture of the image by calculating its pixel value and on the basis of that pixel value, these visual attributes are kept in a separate database which is called the feature database or trait database. This feature database represents only an abstract view of the images which are present in image database. The reason for this is that the part of the visual traits for any image is if we compare it with the image data. But in the case of other method which is known to be the on-line image retrieval, the system works for searching an appropriate image by firing a query to the system. The systemaddresses this query example as a feature vector in the database. Then dissimilarities or similarities are calculated as well as prioritized. The similarities or distance are calculated between the feature vectors of the query with the data of the whole trait database. For searching an image database in an effective and optimum manner the extraction is done with the usage of any scheme for purpose of indexing. In the last the search outcomes are numbered on the basis of most similar features and the one which are very similar with the image in the query example are retrieved. There is also an option for the relevance feedback to the system which is retrieving the image and this option works by knowing the user information needs. There are mainly three components of any CBIR system in which first is Query image which is defines as an image which can be present in the image database, the similar image can be or cannot be in the database and also the number of images of the similar type are present there or not present. The next

important component is the Image database which is containing a variety of images on the basis of user need. There can be a huge number of images present in the image database. The last but the most crucial component is the Feature extraction. It is the most crucial process for the purpose of classifying the image in CBIR. The information related to appearance of the image is taken and is stored as a feature vector in the feature database. The working of the process of image classification is as shown below [6].

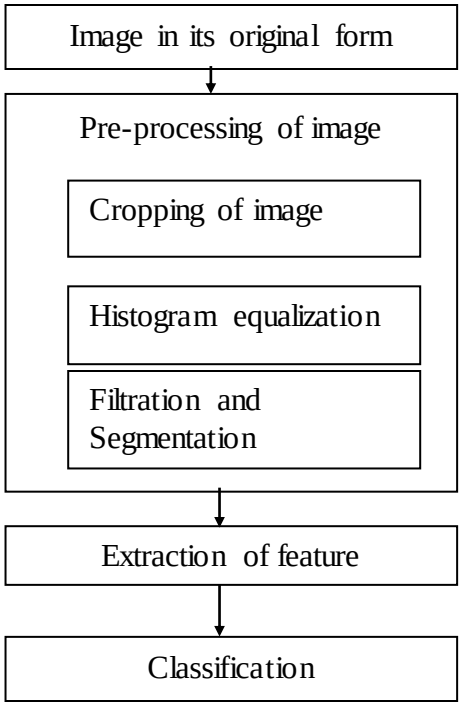


Fig 1 Image Classification

The feature extraction is done for the estimation of traits of the image as some value or a collection of value which is called the feature vector of an image. And these feature vectors are actually matching the image in the query with the other images in feature database. This process works by extracting out all the appropriate or inappropriate traits of the particular image which are driven from the classification. There are generally many features of an image for extracting the features but the most complicated task is to extract the features which are appropriate for the user need. The features of the image are further classified into two parts as the semantic features and the features which can be retrieved by the human eyes that means visual features. Now again these features can be the general features or specific ones. In the general features we consider only the colour, shape, texture for the image. For example, if we are concerned with finding the girl so the domain here is human but on the other side if we are willing to search a monkey then the domain will ne animal. These features are domain specific [7].

There are also some features for which it is very complicated to retrieve as these also contain some important or information related to the meaning of the image. These are called as the semantic features. For these features we use mean value, RGB value, Histogram value, Standard deviation and entropy etc. The componential structure of CBIR is as below.

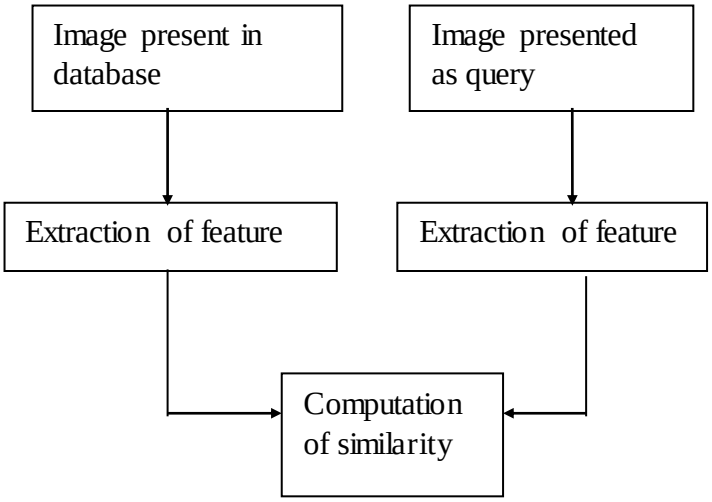


Fig 2 Componential Structure of CBIR

## II. OVERVIEW OF SVM

The concept of SVM was firstly given by Vapnik and since then it has been attaining popularity because of its features which makes it attractive while doing real world performance [2]. It actually provides the remarkable netter outcomes in classifying any image in comparison to other variety of algorithms or presented techniques for the same purpose. It is primarily dealing with the practical application problems normally encountered during voice recognition, recognition of tones, categorisation of text, classification of image as well as data [2]. There are basically two terms present in the concept of machine learning. These are supervised and unsupervised learning. Supervised learning works on the basis of learning by result whereas learn by example is the main feature of unsupervised learning. In case of supervised learning the input is a collection of training data. Support vector machine works on the principle of supervised learning technique by exploring the data as well as recognising the pattern for classification purpose. It works by taking the whole set of input then reading it and then for each of the input, the relevant output is extracted. This whole process is assumed as classification. Now the result can be obtained in two forms either as discrete or continuous. In the case of discrete outcome the process of

classification is performed and in the other one regression is done.

SVM works by the mapping of input space to the feature space. Feature space is defined as the space which is kept for the purpose of calculating similarity by usage of the kernel function. The main feature of this space is that here the linear separation is much easier [7]. Here the transformation of the unprocessed data is done into predetermined extent sample vectors. While discussing about the feature space, the two important concepts named as feature value and the feature vector should also be consider. The main trait of any image is called its feature value and feature vector is vector in which these feature values offers the machine. The aim of the Kernel function in the kernel method is implementing the classification and the clustering of variety of the data like text document, image, graphs, vectors etc. Support vector is basically a point present in the feature space and these are separated by some distance. This point is present between the origin in the feature space and the point which is depicting the position of separator. In the image classification, the images of same kind are grouped into one unit which is called the cluster for that image. So the process of classification can be thought as the stimulating work to retrieve the correct result. For enhancing the performance of this classification it is possible to retrieve the nearby feature of the image as it will be helpful to attain the accuracy in the results [8]. In all the previous systems for image retrieval there were some difficulties when using in real world applications like users always desire for their result in single time on the web, lots of negative results may lead to more time consumption while searching etc [1]. So we are using this CBIR system for retrieving images present in the database which are much similar to the query image and SVM is seen as imparting the best result. The retrieval of all features of image on the basis of visual characteristics of image to the query is the primary aim of CBIR [9].

### *A. Explanation of SVM (Support Vector Machine) in mathematical terms*

In this we are considering few cases which are given as below:

The first one is Optimal separating hyper planes:

For this a collection of points in training data form which is classifying the classes for the pattern is given as

S= {(x1, y1), (x2, y2), (x3, y3) ... (xi, yi)}

here xi represents a p-dimension real vectors, yi is {-1, +1} and n is considered as the total number of sample. According to Vapnik's formula =<w.x>+ b …….. (1)

In this w is assumed as the p-dim vector and b can be thought as scalar or a constant. The margin between hyper planes is increased with the addition of scalar value of b.

The next is Linearly Separable:

It is possible to have the selection of hyper planes in such manner that no points exist between them and by this it will be possible to have the maximum distance it can have. The selection of hyper plane in such a way that its distance with the closest point is the largest from the feature space which contains a number of hyper planes is a most tedious work to be carried out. And this is the optimal separating hyper planes. As the distance with the point that is closest can be 1, so after the subtraction of the two distances the added distance of separating hyper planes to the closest point is obtained. Now we get highest Margin which is given as M= 2/ w where the number 2/w is assumed to be the margin which is used for the estimation of an important term called the generalization ability. It is understood from this that the margin and ability to find the generalization error are directly proportional.

Last is the Non-linear support vector machine: With the non- linear support vector machine, a kind of non-linear function is used for the mapping purpose. The mapping is done from the data into the other space. The function used for the mapping is called the kernel function and the space involved in the mapping is the Euclidean distance.

There were many methods used earlier CBIR. The one method is the extraction of navigation pattern. The images can also be retrieved using Colour Moment and Gabor texture feature. The Histogram Refinement is also a popular method used for the comparison of different images. These are used frequently because of their easy computation and their feature of not being affected by minute changes in the image that can be due to any reason.

## III. CONCLUSION

In this proposed paper this feature extraction was seen as the binary classification problem and SVM (Support Vector Machine)was used for solving this problem This paper examined the primary parts of the content based image retrieval system traits depiction. Many further studies in this area have been recognised and also usage of this technique is shown in the same. It is concluded that to attain the extraction at the high speed as well as for making it so flexible that it can also adjust with the images of large size, it is required to have an impressive multidimensional approach. The primary aim of this paper is to represent the significance of Support Vector Machine in the efficient retrieval of image. In this SVM is used as the classifier which is performing the task of

classifying the image and this process of classification is given to all the traits of the image which are extracted after the feature extraction process. It is mostly used for estimating the highest margin hyper planes within the feature space which is also a high dimensional feature space. The mathematical description of support vector Machine is also shown in this work. The use of kernel based learning method is also shown. One important reason for the development of the highly effective and capable image retrieval system is the drastic or surprising increase in the image database size. The earlier work in this area was based on the extraction of images with the use of textual information and then later it started the retrieval of images on the basis of content. And this is called the term Content Based Image Retrieval or CBIR on which this study has been done. The primary aim of this paper is extraction of images from database in an impressive and effective manner within the short period of time with the use of Support vector method (SVM).

## IV. FUTURE ENHANCEMENT

This system can also be seen as implemented with the attractive feature of image processing in a programmed manner. It will be a great advancement in the field of computer science. The images will be processed from the real world in such a way that it will be possible to infer or retrieve the all information on the instantaneous basis. It may also be possible to have the astronological usage of these type of systems for the study of entities like stars, planets, galaxies etc. And also more importantly it can also be possible to reduce the overall time needed for retrieval.

## REFERENCES


[1] K. Ashok Kumar & Y.V.Bhaskar Reddy, "*Content Based Image Retrieval Using SVM Algorithm*", International Journal of Electrical and Electronics Engineering (IJEEE), ISSN (PRINT): 2231 – 5284, Vol-1, Iss-3, 2012.

[2] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu, "*Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns* " IEEE, and Vincent S. Tseng, Member, Vol. 23, No. 3, March 2011.

[3] Yong Rui and Thomas S. Huang, "*Image Retrieval: Current Techniques*" Promising Directions, and Open Issues January 7, 1999

[4] Martin Solli , "*Content Based Image Retrieval Fonts and Color Emotions*" Copyright c ° 2009 Department of Science and Technology Campus Norrk¨oping, Link¨oping University SE-601 74 Norrk¨oping, Sweden.

[5] Ritendra Datta Jia Li James Z. Wang, "*Content-Based Image Retrieval - Approaches and Trends of the New Age*

[6] Le Hoang Thai, Tran Son Hai and Nguyen ThanhThuy, "*Image Classification using Support VectorMachine and Artificial Neural Network*", I.J. Information Technology and Computer Science, 5, 32-38, 2012.

[7] Le Hoang Thai, Tran Son Hai and Nguyen ThanhThuy, "*Image Classification using Support VectorMachine and Artificial Neural Network*", I.J. Information Technology and Computer Science, 5, 32-38, 2012.

[8] Sandeep Kumar, Zeeshan Khan, Anuragjain, "*A Review of Content Based Image Classification using Machine Learning Approach*", International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-2 Number-3 Issue-5 September-2012.

[9] J Umamaheswari and Dr.G.Radhamani, "*Quadratic Program Optimization using Support Vector Machine for CT Brain Image Classification*", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012.

[10] Ramakrishna Reddy.Eamani and G.V.Hari Prasad, "*Content-Based Image Retrieval UsingSupport Vector Machine in digital imageprocessing techniques*", International Journal of Engineering Science and Technology (IJEST), ISSN: 0975-5462 Vol. 4 No.04 April 2012.