Content Based Image Retrieval Using Classifiers and Zernike Moments

Monika Dhiman
Asst Professor
Electronics & Communication Engineering. Deptt.
Global Institute of Technology & Management, Gurgaon

***Abstract*— In this work, it proposes A Novel Content based Image Retrieval (CBIR) System using Zernike Moments and SVM Classifier. It observes that the average retrieval efficiency will be increased as the moment order increases. It also observes that the classification efficiency of the proposed CBIR system is increased with the increase in the number of training samples. The Zernike moments based features are quite unique and exhibit fair image retrieval performance when tested on real time data base of images. The features are normalized with respect to rotation and size so that the image when rotated at some angle appears same as at the original angle and size. The support vector machine & KNN classifier classify the retrieved images further into two categories: one with most similar and others. All simulations are done in MATLAB.**

***Keywords*- CBIR, Image Retrieval, Precision, Zernike Moments etc.**

I. INTRODUCTION

As with the fast development in Internet & decreasing rates of various storage devices, there is requirement to store image, text, audio & video in digital format. It causes a challenge for the various designing methods that provide useful search. It provides a way through contents of digital records. Due to this, the retrieval and indexing of image has become an important research area. Recent developments in Information Technology modernize many disciplines of health care especially Biomedicine.

Image recovery procedures mainly belong to two categories based on the query formats: keyword-based and content-based approaches. In the keyword based approaches, it is based on scheme to store a keyword explanation of content of image. It was formed by a user on contribution, as well as a pointer to the unprocessed image data. After this, retrieval of image is shifted to typical database organization capability that shared with information recovery methods. Some viable search engines, such as Lycos Multimedia Search and Google Image Search are keyword-based image recovery systems.

It is not always available the Manual explanation for a large gathering of images. Sometimes it may be extremely hard to annotate a picture using several keywords. This may help to research on content-based image retrieval: recovery of pictures by image example where a query image or sketch is given as input by a user. Generally speaking, CBIR aims to produce different methods that maintain and help in efficient search and image digital libraries browsing. It is based on imagery characteristics which are automatically derived. Content-based image retrieval is also called as content-based visual information retrieval (CBVIR) and query by image content (QBIC).It is used as the application of computer vision method that helps in image recovery problem as shown in figure 1.1. The quick development of digital image databases has motivated the CBIR that requires proficient investigating schemes. Content-based image recovery is a method that uses visual contents for finding images from large scale databases.

It has been an energetic and fast advancing research area since the 1990 based on user interest. "Content-based" means that the search analyses the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. While image libraries are growing at a rapid rate (personal image collections may contain thousands, commercial image repositories millions of images, most images remain un-annotated, preventing the application of a typical text-based search.

The early time of CBIR research is from 1994 to 2000. In 2000, history introduced a survey of the advancement made in CBIR during this period. The work examined the examples of utilization, sort of pictures, holes in picture understanding and the various kinds of highlights utilized for portraying the picture properties. They have arranged the picture area into thin and wide. The inconstancy in restricted area is constrained and unsurprising while expansive space has boundless and flighty fluctuation in its appearance. As indicated by creator, the three classes of picture search were: search by affiliation, pointed inquiry and classification search. Highlight portrayal is separated into two sections, picture preparing and include development.

From a computational point of view, an ordinary CBIR framework sees the question picture and pictures in the database as an accumulation of highlights, and positions the significance between the inquiry picture and any objective picture in extent to a comparability measure determined from the highlights. In this sense, these highlights, or marks of pictures, describe the substance of pictures. As indicated by the extent of portrayal, highlights fall generally into two classifications: worldwide highlights and neighbourhood highlights.

The previous class incorporates surface histogram, shading histogram, shading format of the entire picture, and highlights chose from multidimensional segregates investigation of an accumulation of pictures. CBIR doesn't require any additional information, as it concentrates picture includes legitimately from the picture information and utilizations these, combined with a likeness measure, to inquiry picture accumulations. Customary techniques for recovering pictures isn't agreeable or may not fulfil client need E.g. in Google picture composing 'Apple' restores the Apple items just as the apple organic product.

From the current overview, it is seen that picture recovery dependent on Zernike minutes is a significant calculation when the picture information base is enormous and the pictures differ in nature in a wide range for example the information base may contain remote detected pictures, hyper phantom pictures, scenes, restorative pictures, photos and street scenes. The Zernike minutes in mix with head part calculation play an extremely viable device in face acknowledgment space. The face picture information base issue ordinarily require a fast reaction, the Zernike minute adapt up to the issue. The primary burden with the Zernike minutes is that it requires a ton of computational endeavours that lessens the proficiency of the calculation when the picture size just as the information base size increments. These minutes have been widely utilized as trait descriptors in picture examination. Obviously, the highlights as acquired by utilizing the Zernike minutes has a decent goals and are reasonable enough in recovering the pictures from their information base precisely.

The paper is organized as follows. In Section II, It describes introduction of CBIR system. In Section III, it describes the proposed system with some introduction of three step search and four step search algorithm. Section IV defines the results of proposed system. Finally, conclusion is given in Section V.

II. CBIR AND ITS FEATURE COMPONENTS

Content based systems usually contain lower-level features like color, texture and shape. Texture is basically the trends in design a picture of information generally follows. Each information does have different textures information. Color is very basic information regarding any picture or video and lies under the category of low level information. Shape distinguishes the important information assigned in a given picture or video with the help of shape the principle information can be classified first and can be used for very constructive purpose. The brief description of CBIR features are classified below:

1. Texture

The ability to retrieve images on the basis of texture similarity may not seem very useful, but can often be important in distinguishing between areas of images with similar colour histograms (such as sky and sea, or leaves and grass). A variety of techniques have been used for measuring texture similarity. The most established ones rely on comparing values of what are known as second-order statistics calculated from the query and stored images. Texture deals with visual patterns in images and describe how they are spatially defined. They are represented by texels which are then positioned into a number of sets. It depends on no. of textures are detected in image. These sets describe the texture as well as location of texture. It is very hard to explain. Texture identification in images is done by modelling texture as a 2-D gray level variation. After this, the parameters related to brightness of pixels are calculated such as regularity, coarseness and degree of contrast etc. Though, the difficulty is to identify patterns of co-pixel deviation and associating them with particular classes of textures such as silky, or rough.

2. Colour

The shading histogram for each picture is put away in the database. At pursuit time, the client can either indicate the ideal extent of each shading (75% olive green and 25% red, for instance), or present a model picture from which a shading histogram is determined. In any case, the coordinating procedure recovers those pictures whose shading histograms coordinate those of the inquiry to inside indicated limits. Varieties of this method are presently utilized in a high extent of current CBIR frameworks. Strategies for enhancing Swain and Ballard's unique method incorporate the utilization of total shading histograms Computing separation estimates dependent on shading similitude is accomplished by figuring a shading histogram for each picture that recognizes the extent of pixels inside a picture holding explicit qualities (that people express as hues).

3. Shape

In CBIR applications, shape highlights feature neighbourhood and worldwide spatial appropriations of the picture designs. Those shapes are characterized by 2-D areas got from low-level pixel shading and circulation highlights, which are gatherings of associated picture pixels having comparative hues or surfaces. As a rule, picture shapes depends on pictures seeming to have similar properties in reality picture scene characterized by human vision frameworks, which is made a decision by human minds as geometric/ relative invariant, commotion/ impediment safe and movement free Shape doesn't allude to the state of a picture however to the state of a specific locale that is being searched out. Shape is one of the essential visual highlights in CBIR. Shape descriptors fall into two classes i.e., form based and district based.

III. DESCRIPTION OF PROPOSED SYSTEM

Colour-based CBIR is one of the most widely used features of CBIR. To improve data handling efficiency, color has been used as a feature vector to represent the content characteristics within an image. Although color can be an efficient representation for digital images, it carries little information about the spatial structures and shape features within the image. In existing work, it presented an effective image retrieval method by combining high-level features from Convolutional Neural Network model and low-level features from Dot Diffused Block Truncation Coding.

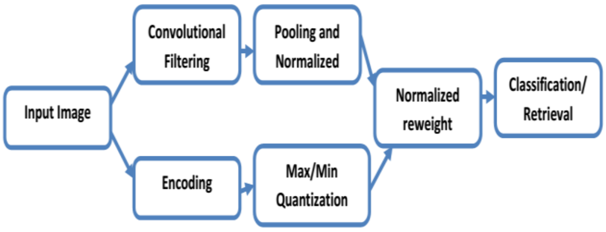

Figure 1: General System Model

As the feature extraction are all image size based. And for even a single feature extraction, the entire image has to be scanned. In the presented work, a modified Zernike moments based image retrieval system is proposed. The limit of operation of image size has been taken care off to a great extent. In the presented work, it has been observed that the Zernike moments are rotational and translational invariant or are made invariant by using image orientation angle. But how the image retrieval accuracy and speed will be affected with increase or decrease in image size has to be examined. It also uses a deep training algorithm for improving the performance of system.

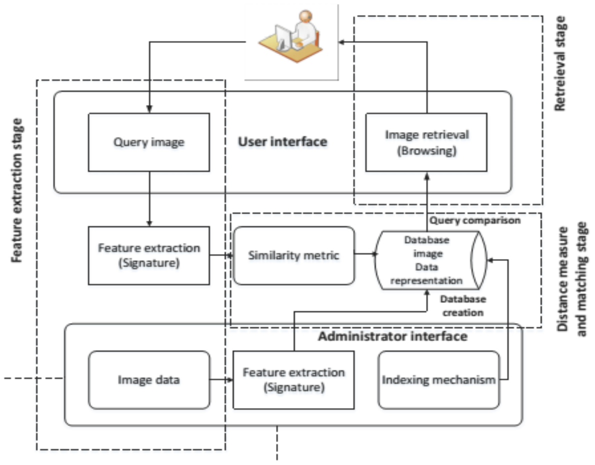

Figure 2: Composition of CBIR System

In the presented work, a modified Zernike moments based image retrieval system is proposed. The main objective is to improve accuracy of system using deep training algorithm KNN and also uses SVM as a classifier. In this work, a database is extracted and then complete database is processed through image processing steps. It is a Offline process. The features of each image on database are extracted, through the visual content descriptors, for a multidimensional vector. Then, the feature vectors are stored on a new database. This stage processes the image query. The system uses the query for example (QBE) paradigm. This means that the user employs one or more sample images as the starting point for a search of visual information. The descriptors and vector format are the same as used in Image Database. This comparison module computes the similarities between the features vectors of database and the query feature vector. After that, this component ranks the images of database

The user can affine the results by providing interactive information. In other words, the user indicates if a resulting image is a positive example (relevant) or a negative example (irrelevant). Then, the process of information needed is refined and started over again. Relevance Feedback is an optional tool in a CBIR system. The gray image is converted into binary image by threshold decomposition, thereby isolating the foreground comprising the yarn from the background. This binary image is further used for feature extraction. Most of the histogram based thresholding techniques work on the assumption that foreground area is comparable to that of background area.

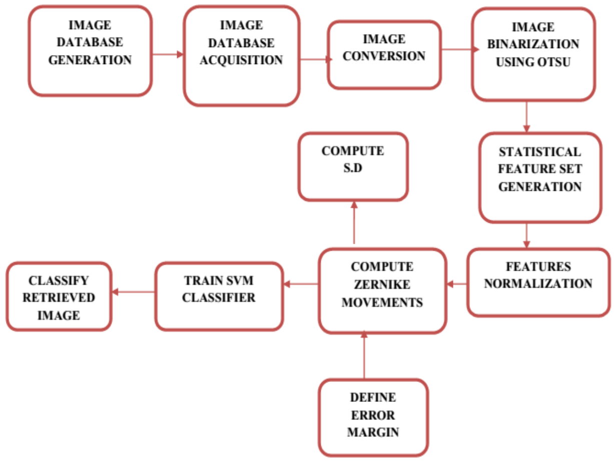

Figure 3: Proposed System Model

In Otsu's method, we exhaustively search for the threshold that minimizes the intra-class variance(Within Class variance), defined as a weighted sum of variances of the two classes:

$$\sigma_w{}^2(t) = \omega_1(t)\sigma_1{}^2(t) + \omega_2(t)\sigma_2{}^2(t) \qquad (1)$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and $\sigma_i{}^2$ variances of these classes. Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance,

$$\sigma_b{}^2(t) = \sigma^2 - \sigma_w{}^2(t) = \omega_1(t)\omega_2(t)[\mu1(t) - \mu2(t)]^2 \qquad (2)$$

which is expressed in terms of class probabilities ω_i and class means μi.

The class probability $\omega_1(t)$ is computed from the histogram as t:

$$\omega_1(t) = \sum_o^t p(i) \qquad (3)$$

The next stage of composition of the system was extracting the Zernike moments (ZM) descriptor feature and the orthogonal moments. ZM is used in this research because it has been successful applied to computer vision and pattern recognition. The Zernike moments of an image have been observed to be rotation and translation invariant. However, prove to be poor on account of scale i.e. size of the image. In case of content based image recovery, it is not necessary that data base images are of same size. In-fact, there may be same image at different scales. For an ideal content based image recovery procedure, the features must be scale/size, orientation and Translation invariant. In the presented work, Zernike moments are computed for the images that are rotation and translation independent and radial features are computed that are scale independent or size invariant. Further, the standard deviation, variance and figure aspect may also be computed to enrich the radial feature set.

After this, it uses SVM and KNN as a classifier. The support vector machine (SVM) is a supervised learning method that generates input-output mapping functions from a set of labelled training data. The mapping function can be either a classification function, i.e., the category of the input data, or a regression function. For classification, nonlinear kernel functions are often used to transform input data to a high-dimensional feature space in which the input data become more separable compared to the original input space.

SVM is an AI technique that deals with the guideline of auxiliary hazard minimization so as to locate the best hyper plane that isolates two classes (ordinary and unusual). The information utilized for this SVM is preparing information and testing information. In this examination, testing information are separated into 3 gatherings. The main gathering, testing information were taken inside from preparing information. The subsequent gathering, testing information were taken outside from preparing information. Also, the third gathering, testing information were taken inside and outside from preparing information.

KNN order strategy is a least difficult procedure theoretically and computationally, however despite everything it gives great arrangement precision. The KNN characterization depends on a lion's share vote of k-closest neighbour classes. To see how the KNN functions, first characterize a point which speaks to include vectors of a picture in an element space. At that point, decide the separation between the point and the focuses in preparing informational collection.

The steps of this technique are:

- *Compute distances of the query to all training examples.*
- *If the k neighbours have all the same labels, the query is labelled and exit; else, compute the par-wise distances between the k neighbours;*
- *Convert the distance matrix to a kernel matrix and apply multiclass SVM;*
- *Use the resulting classifier to label the query.*

IV. RESULTS & DISCUSSION

In this work, a database is created and used for CBIR system. The images retrieved are at affair rate of accuracy. A data base of approximately 100 images is created. The data base is bifurcated based on shapes, color, size, texture and texts. Similar sort of query images are used to test and validate the system. The validation process is under way. The program results as obtained from the algorithm implemented in MATLAB environment. The program asks to input the query image, threshold value and error margin as input variables.

In this, it consists of a LOAD DATASET button that is used to load dataset of images in workspace environment or it can add image directory by providing the path to it. After loading the database, it provides a BROWSE button for selecting the input imag. It can be of any format .jpg, .tif etc. After loading the input image, it provides the Zernike moments of image after following the basic process of image processing. The number of images in this work may vary from one to 10. It uses various similarity metrics like cosine transform, chebyshev, L1, L2 etc. It can use any one by selecting from popup menu. The major justification for introducing orthogonal Zernike Moments is the speed and stability of their numerical implementation.

Figure 4: Input Image Selected

Figure 5: Zernike Moments Output

Figure 6: Image Return After Query

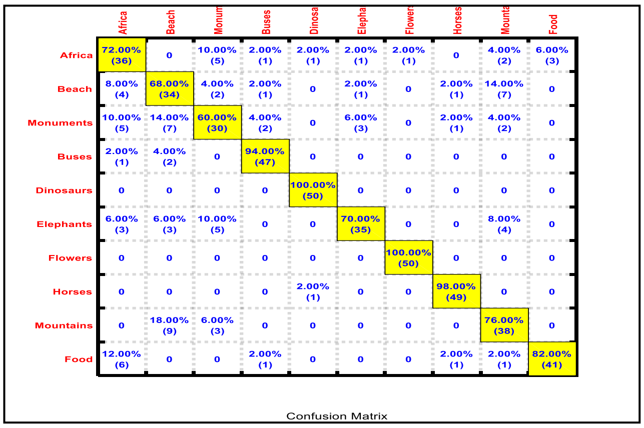

Figure 7: Confusion Matrix after SVM Classifier

In the labelled step all training images of the database compute and add descriptors to algorithm. The classification results are shown in figure respectively. In the test level, the algorithm computes descriptors d of the query image. Figure 8 shows the accuracy results using different Distance metrics. In this, the accuracy is the proportion of the total number of predictions that were correct. In this, the method L1 and cosine metrics shows better accuracy as compared to other metrics performance. It compares various classification techniques on basis of precision and recall parameters. The KNN classifier shows better performance in terms of accuracy and precision values.

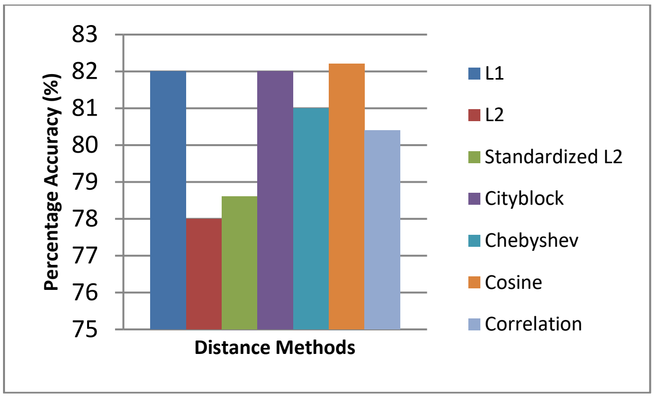

Figure 8: Accuracy using Different Distance Methods

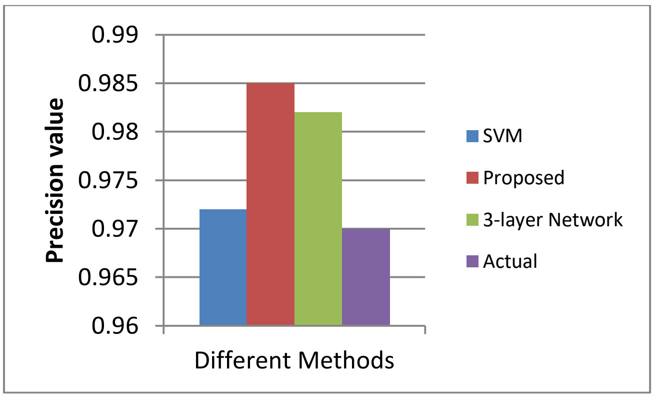

Figure 9: Proposed Precision using Different Method

V. CONCLUSION

It has been concluded on the basis of results that came across that an improved Content based Image recovery (CBIR) System using Zernike Moments has been proposed. It uses classifiers for image classification. It observed that the average efficiency of retrieval is increased when the order of moments increases. The Zernike moments based features are quite unique and exhibit fair image retrieval performance when tested on real time data base of images. The features have been normalized with respect to rotation and size so that the image when rotated at some angle appears same as at the original angle and size. . In this, the accuracy is the proportion of the total number of predictions that were correct. In this, the method L1 and cosine metrics shows better accuracy as compared to other metrics performance. The result shows the performance comparison of system. It compares various classification techniques on basis of precision and recall parameters.

REFFERENCES

[1]. David Edmundson and Gerald Schaefer, "Fast Mobile Image Retrieval", *IEEE International Conference on Systems, Man, and Cybernetics* 2013, pp. 882-887.

[2]. Suraya Abu Bakar, Muhammad Suzuri Hitam, "Investigating the Properties of Zernike Moments for Robust Content Based Image Retrieval", *IEEE International Conference on Computer Applications Technology (ICCAT)* 2013, pp. 5285-5290.

[3]. S. Pradeep, L.Malliga "Content Based Image Retrieval and Segmentation of Medical Image Database With Fuzzy Values", *IEEE International Conference on Computational Science and Computational Intelligence*, 2014, pp. 3834-3840.

[4]. G. Nandhakumar, V. Saranya, "IRMA-Improvisation of Image Retrieval with Markov chain based on Annotation", *IEEE International Conference on Computational Science*, 2014, pp.3841-3847 .

[5]. R. Grace, Dr. R. Manimegalai, "Medical Image Retrieval System in Grid using Hadoop Framework", *IEEE International Conference on Computational Science and Computational Intelligence*, pp. 144-148, 2014.

[6]. M. Dass, M. Ali, "Image Retrieval Using Interactive Genetic Algorithm", *IEEE International Conference on Computational Science and Computational Intelligence*, pp. 215-220, 2014.

[7]. K. Belloulata, L. Belallouche, "Region Based Image Retrieval Using Shape-Adaptive DCT", *IEEE Summit & International Conference on Signal and Information Processing (ChinaSIP)*, pp. 470-474, July 2014.

[8]. X. Yang, X. Qian, "Scalable Mobile Image Retrieval by Exploring Contextual Saliency", *IEEE Transactions on Image Processing*, Vol. 24, No. 6, pp. 1709-1721, June 2015.

[9]. Peizhong Liu, Jing-Ming Guo, "Fusion of Deep Learning & Compressed Domain Features for Content Based Image Retrieval", *IEEE Transactions of Image Processing*, Volume: 26, Issue: 12, pp. 5706-5717, Dec. 2017.

[10]. YanshengLi ,Yongjun Zhang, "Large-Scale Remote Sensing Image Retrieval by Deep Hashing Neural Networks ", *IEEE Transactions On Geo-Science and Remote Sensing*, Vol. 56, No. 2, pp. 950-965, February 2018.

[11]. J. M. Guo, H. Prasetyo, and N. J. Wang, "Effective image retrieval system using dot-diffused block truncation coding features," *IEEE Transactions on multimedia*, vol. 17, no. 9, pp. 1576-1590, Jun. 2015.

[12]. J. M. Guo, and Y. F. Liu, "Improved block truncation coding using optimized dot diffusion," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1269-1275, Mar. 2010.

[13]. Mohamed Elhoseiny, Sheng Huang, "Weather Classification With Deep Convolutional Neural Networks", *IEEE Conference on Computer Vision and Pattern Recognition* 2015, pp. 2249-2253.

[14]. C. Szegedy, et al., “Going deeper with convolutions,” *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9, 2015.

[15]. J. Wan, D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, and J. Li, “Deep Learning For Content-Based Image Retrieval: A Comprehensive Study,” *In Proceedings in Journal of Association for Computing Machinery*, pp. 157-166, 2014.

[16]. G. J. Burghouts, and J. M. Geusebroek, “Material-Specific Adaptation of Color Invariant Features,” *IEEE Pattern Recognition Letters*, vol. 30, pp. 306-313, 2009.