

*A Novel Research in Feature Extraction using Texture and  
Shape for Content Based Image Retrieval*

**S. Sadhana**

Assistant Professor,  
PG & Research Department of Computer Science,  
Hindusthan College of Arts and Science,  
Coimbatore-641028 – India

---

*Abstract: Image Retrieval is the basic requirement of today's life in present scenario. Because of huge amount of different types of images are added in database from different sources for retrieval of the image, different kinds of processing is required to extract the relevant features from them. In this paper, comparisons of combination texture and shape features are done with texture Gray Level Co-occurrence Matrix and Hu-moments and the combination of tamura texture and shape invariant Hu-moments. For the performance evaluation of the system we use most commonly used methods namely precision and recall.*

*Keywords: Content Based Image Retrieval (CBIR), Gray Level Co-occurrence Matrix.*

---

## I. INTRODUCTION

In recent years image processing is used for multimedia database for storing digital images and also based on digital image processing which is the application of computer based algorithms. Image Retrieval aims at developing techniques browsing large image digital libraries to find out whether an image or the image database contains the query pattern given by the user on the basis of similarity measures. Image Retrieval is basically based on two approaches Text Based Image Retrieval (TBIR) [1] and Content Based Image Retrieval (CBIR).

Content Based Image Retrieval (CBIR) [2] is a kind of Image Retrieval technique that could figure out images like sketch similar to querying images from image database. Image features are extracted based on content and these features like color and shape can be used to retrieve the images from the large database based on some similarity measures. In Content Based Image Retrieval (CBIR) image can be retrieved by query known as Query by Image Content and also known as content Based Visual Information (CBVI) which is the application of Computer Vision techniques for the problem of searching digital images from large image database known as image retrieval problem. A typical CBIR system consists of four main steps. Firstly, the most basic process of CBIR system is feature extraction in which extract the features of image like from spatial based on pixels. Secondly, the system will characterize those selected feature spaces and the construct the feature vectors. Thirdly, these feature vectors of query image are compared with images stored in database. Finally, the output queried image results to the similar images to the query images based on computing some similarity measures. Then performance evaluation of the system is done we use most commonly used methods namely, precision and recall. The main issues in CBIR system arises which features should be derived to describe the images better within the database and which data structure be used to store the feature vectors .And most important which learning algorithms should be used in order to make the CBIR efficient. And finally, most important at the retrieval time is how to participate according to user's feedback in order to improve the result.

## II. RELATED WORK IN THIS FIELD

Kato, T. [3] in 1992, National Science Foundation of United States organized a workshop on visual information management system in order to identify a new direction in image Database Management System introduced the term Content Based Image Retrieval, and they emphasized the use of color and shape as the most important criteria of feature extraction for image retrieval system. Since, and then the CBIR has been adopted as to describe an image retrieving process.

Tamura *et al.*, [4] in 1973, proposed texture representations that were based on psychological studies of human perception, and these representations consists of six statistical features, including coarseness, contrast, directionality, regularity, line-likeness, roughness to describe various texture properties. In context to texture features extraction tamura properties are very meaningful and this advantages makes tamura features in Texture Based Image Retrieval. In comparison with psychological measurements for human subjects, the computational measures gave good correspondence in correlation of rank of 16 patterns of typical texture type of patterns. These features were attempted by using similarity measurements.

A. Rosenfield *et al.*, [5] in 1976 applied features and used three standard approaches to automatic texture classification make used these based on Fourier Power spectrum respectively, and these applied features to a set 54 picture samples taken from aerial photographs of nine terrain types (Lakes, Orchard, Railroad and woods *etc.*). Classification results of these types were compared for individual features and pairs of features with each feature class. In general, the Fourier based features performed poorest, while those based on statistics of gray level differences performance's best.

## III. FEATURE EXTRACTION

Feature extraction plays an important role in Image Retrieval system and better selection of feature gives higher accuracy. Feature extraction is basically it separates the visual information from the image and stores them in the form of feature vectors in a feature database. These feature value (or a set of values) called feature vectors of image finds the image information from the feature extraction.

Feature extraction like color, texture and shape. Each feature may have several representations and different representations aspects of the feature when designing an image retrieval system. Texture features like Tamura six features, first-order intensity histogram based features, Haralick features and shape features like Hu-moment.

### 3.1. Feature Extraction Based on Texture

Texture is a very important characteristics for the analysis of many types of images that appears everywhere in nature like natural images, remote sensing images and medical images [15]. Texture can be defined as superficial phenomenon of human visual systems of natural objects. Texture can be attributed to almost everything in nature and also its texture structure of any image is incorporating repeated pattern of all most all of the parts. Texture is commonly known as 'texels'. Texture can be recognized by everyone but it is not easy to define. Texture does not occur over a point but it rather occurs over a region. Texture can be analyzed by quantitative and qualitative analysis.

#### 3.1.1. Tamura Texture Feature

According to quantitative analysis one of the first descriptions given by the Tamura [4] proposed six textural properties and gave descriptions common over all texture patterns in Broadtz's photographic images. These are six different texture features given by tamura Coarseness, Contrast, Directionality, Line-Likeness, Regularity and Roughness.

- Coarseness

Coarseness basically relates to the distance in gray levels of spatial variations, which is implicitly related to the size of primitive elements forming the texture. It has the direct relationship to scale and repetition rates and most fundamental texture

feature. An image will contain repeated textures pattern at different scales, coarseness aims to identify the largest size at which a texture exists, even where a smaller micro texture exists.

The equation calculates the difference between pair of averages corresponding to non-overlapping neighbourhoods.

- Contrast

Contrast measures distribution of gray levels that varies in an image and to what extent its distribution is biased to black or white. The second order and normalized fourth-order central moments of the gray levels are used to define the contrast.

- Directionality

Directionality of an image is measures by the frequency distribution of oriented local edges against their directional angles. It is a global property over a region. This texture feature given by tamura does not differentiate between orientations or patterns but measures the total degree of directionality in an image is given by Directionality. It is the most important feature given by tamura about matrix to distinguish from another image that how much uniform the region is.

- Line-Likeness

Line-Likeness in an image is average coincidence of direction of edges that co-occurred in the pairs of pixels separated by a distance along the edge direction in every pixel.

- Regularity

Regularity measures a regular pattern or similar that occurred in an image.

Roughness is the summation of contrast and coarseness measures.

Roughness= Contrast + Coarseness

In most of the cases, for CBIR system only first three features are used because these features capture the high-level perceptual attributes of a texture and are also useful for browsing of images.

### 3.1.2. Haralick Texture Feature

Gray Level Co-occurrence Matrix (GLCM) a statistical method for examining texture features that consider the spatial relationship of pixels, also known as Gray Level Spatial Dependence. In this a GLCM matrix is created by calculating how often a pixel with the intensity value  $i$  occurs in a specific spatial relationship to a pixel with the value  $j$ . GLCM consists of frequencies at which two pixels are separated by a certain vector occur in the image. GLCM properties by which the distribution in the matrix will depends on the distance and angular or directions like horizontal, vertical, diagonal, anti-diagonal relationship between the pixels. Many statistical features of texture in an image are based on the co-occurrence matrix representing the second order of gray levels pixels relationship in an image. Suppose an input image has  $M$  total number of pixels in horizontal directions and  $M$  total number of pixels in vertical directions. Suppose the gray level that appears at each pixel is quantized to  $z$  number of levels, assume  $N_x = 1, 2, 3, \dots, M$  consists of horizontal space and  $N_y = 1, 2, 3, \dots, N$  consists of vertical space and  $G = 0, 1, 2, 3, \dots, Z$  consists of the set of  $Z$  quantized gray levels. In a given distance  $d$  and direction given by  $\theta$ , the Gray Level Co-occurrence matrix is calculated by using gray scale pixel  $i$  and  $j$ , expressed as the number of co-occurrence matrix in different directions.

Among them five features Contrast, Correlation, Entropy, Energy and Homogeneity.

## Contrast

Contrast measures intensity between a pixel and its neighbor over the whole image and it is considered zero for constant image and it is also known as variance and moment of inertia.

- Correlation

Correlation measures how pixel is correlated to its neighbor over the whole image.

### 3.2. Feature Extraction Based on Shape

#### 3.2.1. Hu-Moment Shape Features

Hu-Moment [17] in 1962 proposed seven properties related to connected region that are invariant to rotation, scaling, and translation (RTS) and are also known as Algebraic Moment Invariants. Moment invariants that are computed from each of the window are used to form feature vectors. They define simply calculated set properties of region that can be used for class identification and also identification of shape, and this classic technique for generating invariants in terms of algebraic.

## IV. ALGORITHM FOR PROPOSED METHODOLOGY

Step1: Input the query image

Step2: Perform the image pre-processing in an input image.

Step3: Convert image from RGBtoGray.

Step4: Extract the texture and shape features form the feature vector.

Feature vector = [Contrast, Energy, Correlation]

Feature vector = [moment1, moment2, moment3]

Step4: Fused the combination of texture and shape feature vector into a single feature vector.

Feature vector = [contrast, energy, correlation, moment1]

Step5: Apply similarity measurement algorithm, Euclidian distance to the extracted feature vector and the feature vectors of image Database stored in database.

Step6: Retrieve the relevant images based on some similarity measures from images stored in Database.

### 4.1 Proposed Methodology

Before analyzing images based on their feature extraction from databases of images, pre-processing methods in images are performed in all types of images. Like, firstly, the images resize according to the region of interest for the faster retrieval of images. Deleting and removing complicated background will speed up further image processing.

Texture [18], very strong discriminative power feature makes an essential component in image and video retrieval. Tamura six texture features in which directionality most important feature based on human perception

**Table 1. Performance Evaluation Method**

Method	Euclidian Distance			
	Tamura+ Hu-Moment		GLCM+ Hu-Moment	
Type of Textures	Precision	Recall	Precision	Recall
Bricks	.78	.88	.67	.75
Stone	.78	.87	.67	.75
Grass	.68	.87	.66	.76
Grills	.78	.87	.67	.75

In the Table the performance evaluation method using precision and recall is shown The result comprises of the comparison of combination of texture and shape. The table consists of four different types of images and the precision and recall value are shown. The comparison of tamura and Hu-moment gives better result than GLCM and Hu-moment is shown.

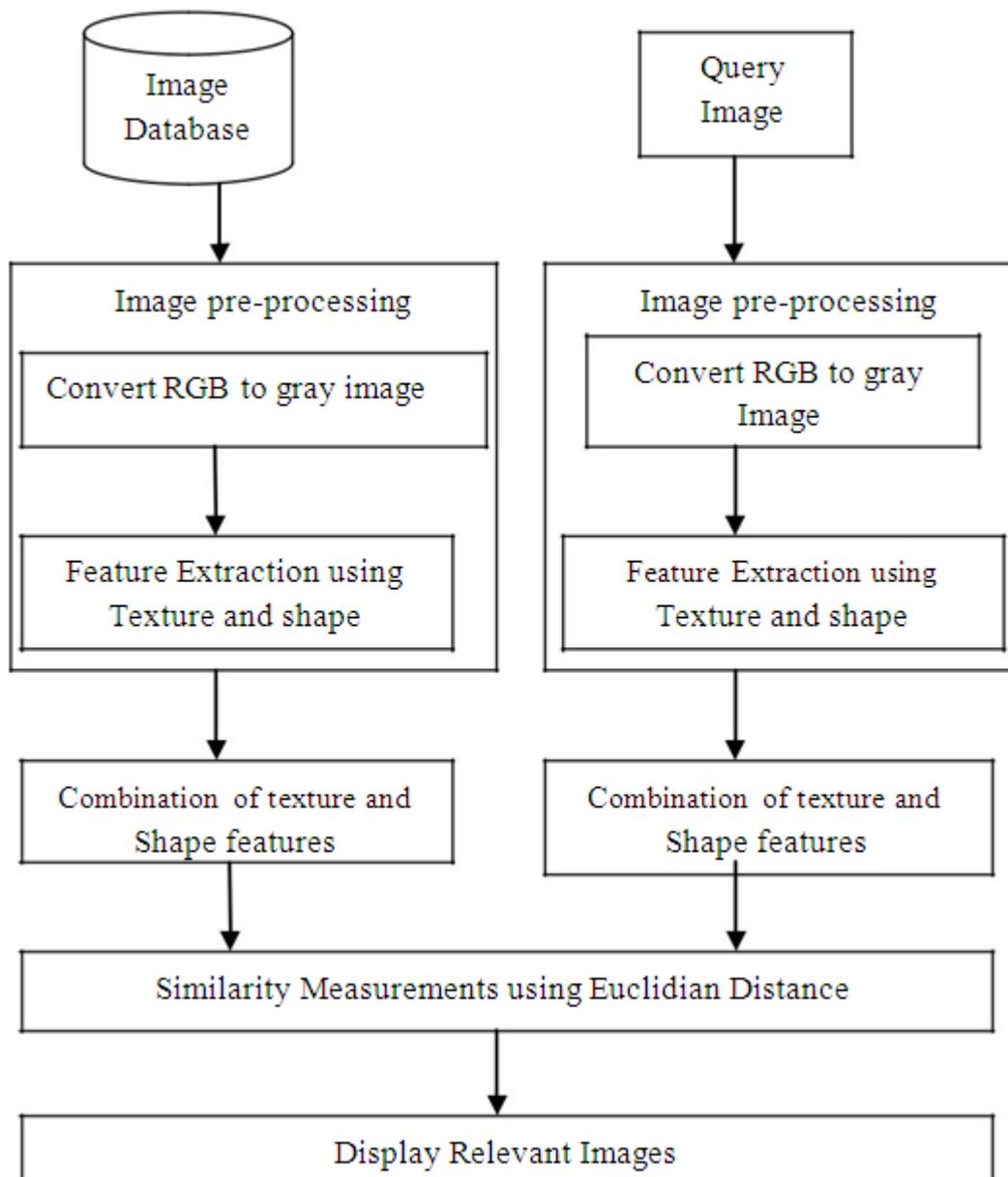


Figure 1. Proposed Methodology of CBIR

## V. SIMILARITY MEASUREMENT

### A. Euclidian Distance

Euclidian Distance [19] matrix is mostly used for similarity measurement in context retrieval of image from database because of its higher accuracy and effectiveness. It measures the distance between the two feature vectors of images by calculating the square root of the sum of the squared absolute differences and is calculated and is denoted by ED.

## VI. EXPERIMENTAL RESULTS

The proposed method is tested by using the Broadtz [20] database of images, which is freely available for researchers. Broadtz database consists of categories like bricks, stones, grass etc. of grayscale images. Retrieval has been classified as accurate if, for a given query image the system retrieved according to human perception the most similar images from the

database are obtained. According to human perception in image retrieval the more the precision value the higher the performance.

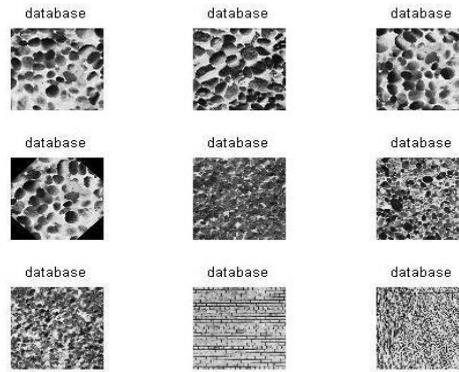


Figure 3. Comparison between Combination Texture and Shape Method

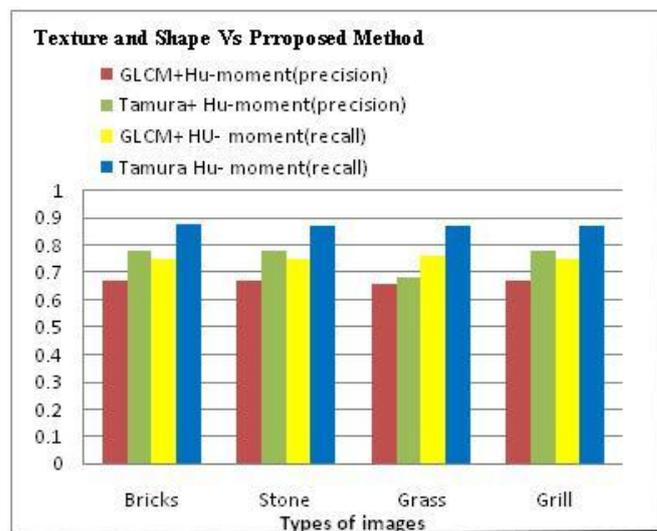


Figure 4. Performance Evaluation Method

## VII. CONCLUSION

In this paper there are different textures and shape feature extraction techniques have been discussed. The main objective of content based image retrieval is to develop is an efficient image retrieval scheme. A comparison is performed between different textural and shapes features are combination of GLCM textural properties and Hu-moment, and tamura textural properties and shape Hu-moment. The combination of tamura textural and shape Hu-moment feature vectors perform better than combination of GLCM Textural and shape moment invariant given by Hu. The result can be further improved by using KNN K-Nearest Neighbor classifier at preprocessing step which gives better result at the retrieval time. Support Vector Machine, Soft computing techniques like neural network can be also applied as a classifier to improve the retrieval time. Evolutionary algorithms can also be used to optimize the result for the better feature selection process

## References

1. N. S. Chang and K. S. Fu, "Query By Pictorial example", IEEE Trans .on Software Engineering, vol. 6, (1980) November, pp. 519-524.
2. T. Kato, T. Kurita and N. Otsu, "A sketch retrieval method for full color image database-query visual example", Pattern Recognition, (1992) January, pp. 530-533.
3. T. Kato, "Database architecture for content- based image retrieval", Image Storage and Retrieval Systems, Proc SPIE 1662, (1992), pp. 112-123.
4. S. Mori Tamura and T. Yamawaki, "Textural Features Corresponding to Visual Perception", IEEE Transactions on Systems, Man and Cybernetics, vol. smc-8, no. 6, (1978) June, pp. 460-473.
5. J. S. Weszka, C. R. Dyer and A. Rosenfeld, "A comparative study of texture measures for terrain classification", IEEE Trans., Syst., Man, Cybern., vol. SMC-6, (1976) April, pp. 269-285.
6. H. Zhao, Z. Xu and P. Hong, "Performance evaluation of three classes of textural coarseness", IEEE, (2009).

7. H. Yao and B. Li, "An efficient approach for texture-based image retrieval", *Neural Networks and Signal Processing*, vol. 2, (2003), pp. 1039-1043.
8. H.-C. Lin, C.-Y. Chiu and S.-N. Yang, "Finding textures by textual descriptions, visual examples, and relevance feedbacks", *Pattern Recognition Letters*, vol. 24, no. 12, (2003) October, pp. 2255-2267.
9. J. Zhang, G.-L. Li and S.-Wun, "Texture-Based Image Retrieval by Edge detection Matching GLCM", the 10<sup>th</sup> international Conference on High Performance computing and Communications, (2008), pp. 782-786.
10. P. Gangadhara Reddy, "Extraction of Image features for an Effective CBIR System", *IEEE*, (2010), pp. 138-142.
11. N. Chaturvedi, S. Agrawal and P. Kumar Johari, "A novel Approach of image retrieval based on texture", *International Journal of Enhanced Research in Management & Computer Applications*, ISSN: 2319-7471, vol. 3, no. 1, (2014) January, pp. 42-48.
12. N. Puviarasan, Dr. R. Bhavani and A. Vasanthi, "Image Retrieval Using Combination of Texture and Shape Features", *International Journal of Advanced Research in Computer and Communication Engineering*, ISSN: 2319-5940, vol. 3, (2014) March, pp. 5873-5877.
13. P. Howarth and S. Ruger, "Evaluation of Texture Features for Content Based Image Retrieval", *Springer-Verlag Berlin Heidelberg LNCS 3115*, (2004), pp. 326-334.