Integrating Shape and Edge Histogram Descriptor with Stationary Wavelet Transform for Effective Content Based Image Retrieval

Manoj D. Chaudhary
P.G Student,
Dept. of Electronics & Communication,
L.D. College of Engineering,
Ahmedabad, India
chaudharymanojd@gmail.com

Abhay B. Upadhyay
Asst. Professor,
Dept. of Electronics & Communication
L.D. College of Engineering,
Ahmedabad, India.
abupadhyay1979@gmail.com

***Abstract*— In this paper we propose a hybrid approach for Effective Content Based Image Retrieval based on texture and shape feature. Towards this, first Stationary Wavelet Transform (SWT) is applied on query image to extract horizontal, vertical and diagonal detail matrices. Stationary Wavelet Transform is used because of its translational invariant property. After this Edge Histogram Descriptor (EHD) is used to exploit the absolute location of edges in the image as well their global composition. To aid the retrieval process, five different shape measures have also been included. Finally Euclidean distance is used to retrieve the relevant results. Experimental results show that the combination of SWT and EHD techniques provides significant improvement over existing methods thereby increasing the retrieval efficiency.**

Keywords—Content-Based Image Retrieval; Stationary Wavelet Transform; Time invariant; Edge Histogram Descriptor; Haar wavelet; Horizontal Span; Vertical Span; Euler Number.

I. INTRODUCTION

Due to the rapid development of computing hardware, digital acquisition of information has become one of the popular methods in recent years. Every day, Giga-bytes of images are generated from various sources. Military, Medical, Journalism are few to name among them. Use of digital content necessitates the development of effective ways for management and retrieval of visual information. Basically images can be retrieved in two different ways, firstly, by text and secondly, by content. Text based approach is very well known and widely used. In this approach a user can search a desired image by entering keywords related to the image. Though the method is simple, it suffers from the problem of human perception based image annotation. This means that the keywords used to search the image are language dependent. An alternative to text based approach is Content Based Image Retrieval (CBIR). CBIR is a technique which uses visual features of an image such as color, shape, texture, etc…to search images from a large database similar to user's query [1]-[3]. The CBIR system have been used in variety applications such as Crime prevention, Architectural and Engineering design, Fashion and Interior design, Journalism and Advertising, Medical diagnosis, Home Entertainment and Web searching etc.

The rest of the paper is organized as follows. Section II provides a brief overview of existing methods. Section III describes the proposed methodology in detail. Section IV presents Experiments and Results and Section V concludes the paper by presenting the scope for future work.

II. LITERATURE SURVEY

The three important features based on which the images are compared in CBIR systems are colour, texture and shape. Reference [4] and [5] describe techniques based on colour histogram for retrieving images. In the work proposed in [6], the authors have used a combination of Hadamard matrix and Discrete Wavelet Transform to extract colour and texture feature for retrieving images. Reference [7] presents an image retrieval technique based on Ripplet transform and Neural network based classifier. The authors have used Ripplet transform to represent the image at different directions and scales followed by neural-network for classification. Reference [8] compares several feature extraction techniques to test their effectiveness in retrieving medical images based on the application to CT brain images. The aim of this paper is to help medical experts to retrieve similar cases from medical database for diagnosis. The work reported in [9] introduces a scheme for retrieving images based on geometrical shapes of the objects in an image. First all the objects in the image are segmented. Then the geometrical shapes of these objects are estimated and compared with predefined shapes to retrieve relevant results.

III. PROPOSED METHODOLOGY

Feature extraction is an important step in any CBIR system. The retrieval accuracy of a CBIR system highly depends upon the extent the given feature vector represents an image under test. The details of methods used for extracting image features in this approach are described below:

A. Texture Feature Extraction

As a measure of texture we first apply Stationary Wavelet Transform on query image and extract Horizontal, Vertical and

978-1-4799-2397-7/14/$31.00 ©2014 IEEE

Diagonal detail sub-matrices. After this 85-bin Edge Histogram Descriptor is applied on each of these sub-matrices to generate feature vector containing the global and local information about the different edge orientations in the query image as described below.

I. Stationary Wavelet Transform:

Discrete Wavelet transform has been widely used for feature extraction in the past [10][11]. However DWT suffers from time variant property. This means that DWT of a translated version of a signal X is not equal to the translated version of DWT of X [12]. Time invariance is desired in many applications such as Change detection, De-noising, Pattern Recognition etc. To overcome the limitations of the traditional wavelet transform, we have used a multi-layer Stationary Wavelet Transform (SWT) [13] in our approach.

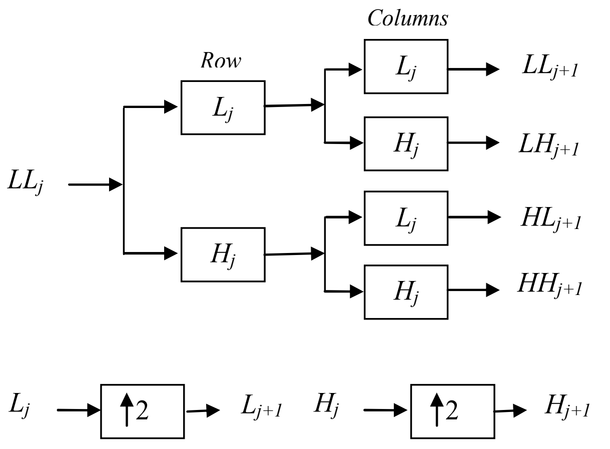


Fig. 1. Stationary Wavelet Decomposition of a two-dimensional image

In Fig. 1 H_j and L_j represent high-pass and low-pass filters at scale j, resulting from interleaved zero padding of filters H_{j-1} and L_{j-1} $(j > 1)$. LL_o is the original image and the output of the scale j, LL_j, would be the input of scale $j+1$. LL_{j+1} denotes the low frequency (LF) estimation after the stationary wavelet decomposition, while LH_{j+1}, HL_{j+1} and HH_{j+1} denote the high frequency (HF) detailed information along the horizontal, vertical and diagonal directions, respectively. These sub-band images will have same size as that of the original image because no down-sampling is performed during the transformation [13]. In our approach we have used "Haar" wavelet to perform multi-layer stationary wavelet transform on the input 2-D image. Mathematically, the wavelet decomposition can be described by (1)-(4).

$$LL_{j+1}(x,y) = \sum_{m,n} L[n]L[m]LL_j(2^{j+1}m - x, 2^{j+1}n - y) \quad (1)$$

$$LH_{j+1}(x,y) = \sum_{m,n} L[n]H[m]LL_j(2^{j+1}m - x, 2^{j+1}n - y) \quad (2)$$

$$HL_{j+1}(x,y) = \sum_{m,n} H[n]L[m]LL_j(2^{j+1}m - x, 2^{j+1}n - y) \quad (3)$$

$$HH_{j+1}(x,y) = \sum_{m,n} H[n]H[m]LL_j(2^{j+1}m - x, 2^{j+1}n - y) \quad (4)$$

Where $L[.]$ and $H[.]$ represent the low-pass and high-pass filters respectively, and $LL_0(x,y) = f(x,y)$. Compared with the traditional wavelet transform, the SWT has several advantages. First, each sub-band has the same size, so it is easier to get the relationship among the sub-bands. Second, the resolution can be retained since the original data is not decimated. Also at the same time the wavelet coefficients contain many redundant information which helps to distinguish the noise from feature. Using SWT we have extracted horizontal, vertical and diagonal detail sub-matrices.

II. Edge Histogram Descriptor

Edges in an image serve as an important feature to represent their content. To explore this feature one way is to use histogram. An edge histogram in the image space represents the frequency and directionality of the brightness changes in the image [14]. To represent this feature we make use of Edge Histogram Descriptor (EHD) [14][15]. It provides information about local distribution of edges over the entire image in form of histogram. EHD basically represents 5 types of edges in a local area which is known as sub-image. The five different types of edges used to compute EHD are shown in Fig. 2. To compute 80-bin local histogram the main image is portioned into 4×4 equal sized non-overlapping blocks. Local histogram is computed for each of these sub-images for every edge type. Thus Local histogram for every sub-image will contain 5 bins representing the relative frequency of occurrence of 5 types of edges in the corresponding sub-image. The total 16 sub-images will provide an 80 (16×5) bin histogram, where bins are arranged serially block by block as shown in Fig. 3.

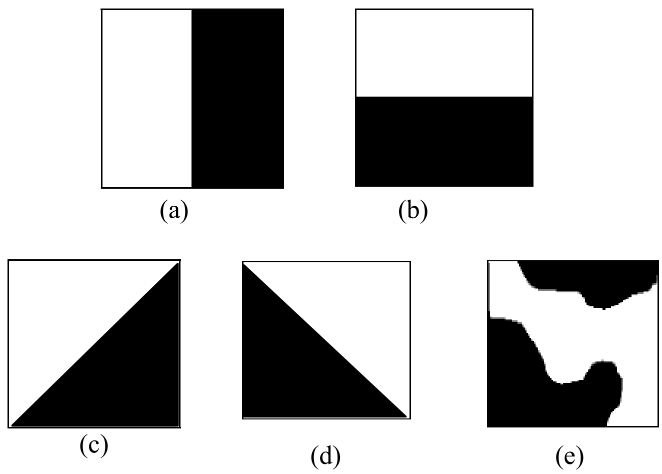


Fig. 2. Five types of edges: (a) Vertical (b) Horizontal (c) 45-degree (d) 135-degree and (e) Non-directional

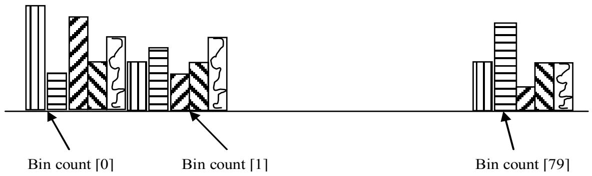


Fig. 3. 1-D array of 80- EHD Bins

Using only local histogram bins may not be sufficient to represent the global features of edge distribution. Hence we consider 5 additional bins as introduced by C. S. Won et al. in [16] as a measure of global edge features. These additional bins are combined with 80 local bins to generate 85-bin EHD. We compute 85-bin EHD for each of the Horizontal, Vertical and Diagonal sub-matrices obtained from previous step. Fig. 4 shows a schematic for generating the feature vector for texture feature.

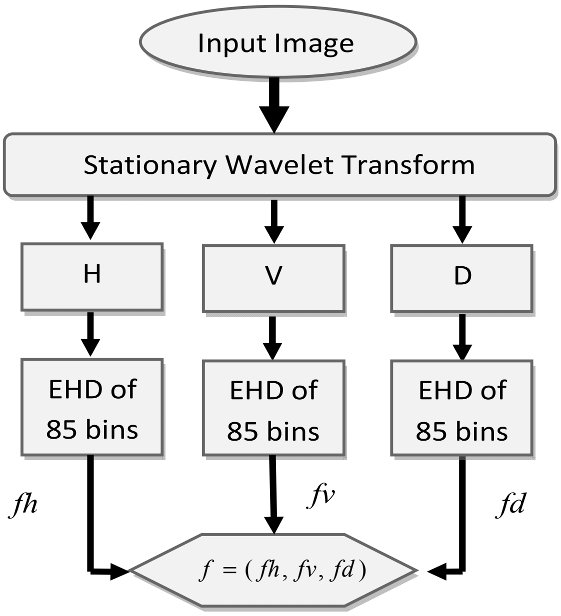


Fig. 4. Schematic for Texture Feature Vector Generation

B. Shape Feature Extraction

Shape is an important visual feature and it is one of the basic features used to describe image content [17]. However representation and description of shape is a difficult task. This is because when 3-D real world object is projected onto a 2-D scene; details of one dimension are discarded. Reference [17] examines and reviews the some important shape representation techniques used in literature. Shape representation and description techniques can be classified into two categories: contour-based methods and region-based methods. The measures of shape used in this paper are described below:

I. Chain Codes: Chain codes are used to represent a boundary by a connected sequence of straight-line segments of specified length and direction [18]. The direction of each line segment is coded by using a numbering scheme as shown in Fig. 5. This representation is based on 4 or 8 connectivity. To obtain a chain code the digital boundary of an image is superimposed with a grid, the boundary points are then approximated to the nearest grid point and a sampled image is obtained.

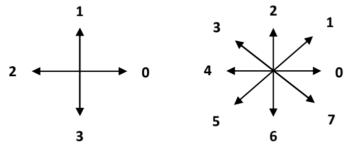
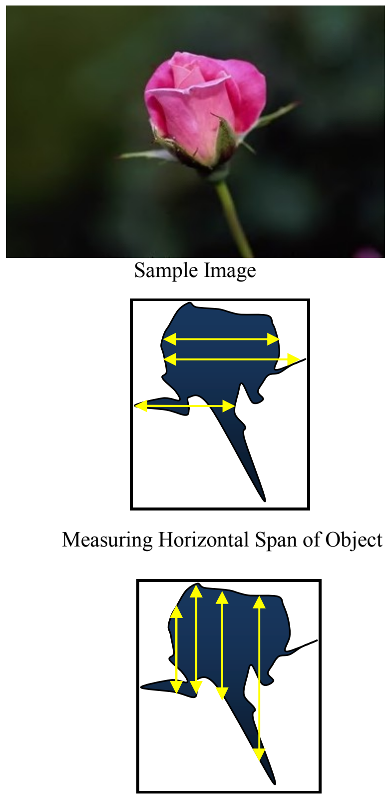


Fig. 5. Directions and Numbering scheme for Chain Codes

II. Area of Object: Total number of pixels lying within a closed boundary. Different shapes will have different area [18].



Sample Image

Measuring Horizontal Span of Object

Measuring Vertical Span of Object

Fig. 6. Computation of Horizontal Span and Vertical span

III. Euler Number: It gives a measure of number of objects in a region minus the number of holes in those objects [18].

IV. Vertical Span: Vertical span describes projection profile of object in vertical direction from top to bottom.

V. Horizontal Span: It describes variance of shape of the object from left to right.

Vertical span and Horizontal span are estimated by finding the extreme boundaries of the object from top to bottom on vertical axis and from left to right on horizontal axis as shown in Fig. 6.

C. Similarity Matching

Once the feature vectors have been generated for Query and Database images, the next step is to compare these vectors to measure the similarity between them. The similarity is measured by computing the distance between the two vectors element wise. The metrics used for similarity matching in our algorithm are *Euclidean* distance and *Manhattan* distance [18][19]. Let the query image be represented by a feature vector $F_{query} = \{Q_i\}_{i=1}^{N}$, where Q_i is the i^{th} component in N dimensional feature space. Similarly let the database image be represented by a vector $F_{database} = \{D_i\}_{i=1}^{N}$. Then the Euclidean distance and Manhattan distance between the two vectors is given by (5) and (6) respectively.

$$ED = \sqrt{\sum_{i=1}^{N}(Q_i - D_i)^2} \tag{5}$$

$$MD = \sum_{i=1}^{N}\left|Q_i - D_i\right| \tag{6}$$

After computing the distance the database images are stored in increasing order of their distance from the query image. The images having similar features to that of query will have smaller values for distance measures. Amongst these top 30 images are returned to the user as relevant images.

IV. Experiments And Results

The proposed method is implemented on MATLAB 7.8.0 (R2009a) on a PC with Intel Dual-Core 3rd Generation Processor having 4 GB of RAM capacity. The Database used for testing the performance of our method is *Wang's Database* [20]. This database consists of 1000 images belonging to 10 different classes. Fig. 7 shows some sample images from each class in Wang's Database. The images are of size 384 × 256 or 256 × 384 pixels. Out of these 100 images from each class, 20 images are used as query, while rest 80 images serve as database from which the similar images are retrieved. Apart from 20 images from Wang's database, 10 additional images from Corel's Database [21] are also used as query. Thus in all 80 images from each class are tested against 30 different queries and the performance of the algorithm is analyzed.

Fig. 7. Sample Images from Wang's Database [20]

A. Performance Evaluation

The retrieval accuracy is evaluated using commonly used measures of Precision and Recall [22]. Precision is a measure of ability of CBIR algorithm to retrieve only relevant images, while Recall decides the ability of CBIR algorithm to retrieve all relevant images as defined by (7) and (8) respectively.

$$P = \frac{No.\,of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \tag{7}$$

$$R = \frac{No.\,of\ relevant\ images\ retrieved}{No.\,of\ relevant\ images\ in\ the\ database} \tag{8}$$

Fig. 8 and Fig. 9 show some retrieval results. The images used as query are "251.jpg" from Class of Buildings and "520.jpg" from Class of Elephants.

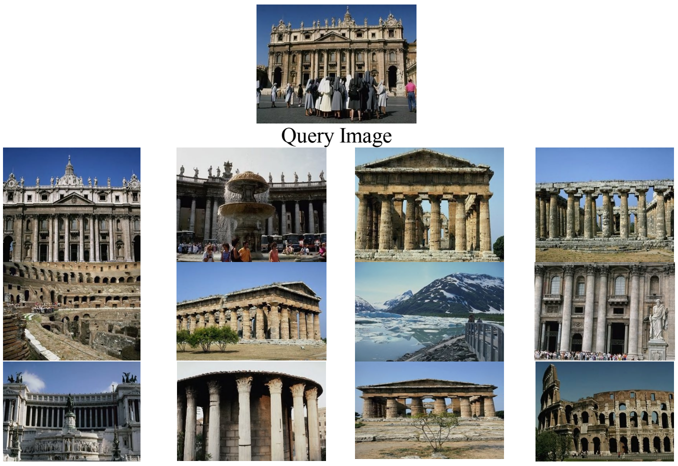


Fig. 8. Query Image (top) and top 12 Retrieved Results

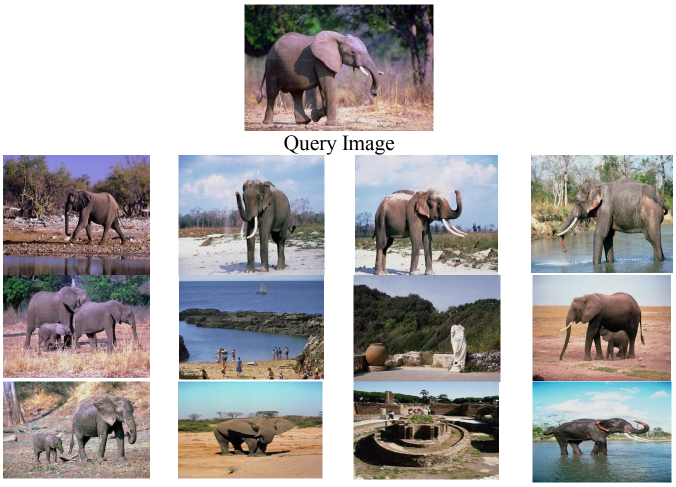


Fig. 9. Query Image (top) and top 12 Retrieved Results

Table I shows average precision values for the proposed method. The precision values are computed at three different retrieval rates. As indicated very high precision value is obtained for the class of "Dinosaur". This is because the regions surrounding the object contain almost no edge information making the comparison task easy. Fig. 10 shows a plot of average precision vs. recall for the proposed method. Table II shows the comparison of proposed scheme with three different existing techniques. Comparison is made on the basis of average precision values computed by retrieving top 30 images at a time. From Table II it can be easily observed that the proposed technique provides improved retrieval performance over other existing CBIR algorithms namely, Pujari et. all's [23] (based on salient points detected by Harris corner detector), Duc et. all's [24] (based on Contourlet Harris detector), and Das et. all's [25] (based on Ripplet transform and Neuro Fuzzy technique).

TABLE I. AVERAGE PRECISION AT DIFFERENT RETRIEVAL RATES

Average Precision			
Image Class	**No. of Retrieved Images**		
	10	***20***	***30***
African	0.69	0.63	0.59
Sea	0.82	0.79	0.75
Building	0.84	0.74	0.73
Bus	1	0.97	0.96
Dinosaur	1	0.99	0.97
Elephant	0.76	0.7	0.67
Flower	1	0.99	0.98
Horse	0.8	0.72	0.69
Mountain	0.82	0.75	0.74
Food	0.74	0.68	0.63

TABLE II. COMPARISON OF PROPOSED TEHCNIQUE WITH EXISTING METHODS

Average Precision (No. of Retrieved Images = 30)				
Image Class	**Method Used**			
	Pujari et. al's [23]	***Duc et. al's [24]***	***Das et. al's [25]***	***Proposed Method***
African	0.48	0.44	0.50	0.59
Sea	0.34	0.43	0.42	0.75
Building	0.33	0.49	0.39	0.73
Bus	0.52	0.64	0.64	0.96
Dinosaur	0.95	0.98	0.99	0.97
Elephant	0.40	0.48	0.55	0.67
Flower	0.60	0.77	0.84	0.98
Horse	0.70	0.75	0.86	0.69
Mountain	0.36	0.31	0.44	0.74
Food	0.46	0.32	0.56	0.63
Average	**0.51**	**0.56**	**0.62**	**0.77**

Performance of the distance measures is also evaluated by taking 9 different images as query from each class and retrieving top 30 images. Each time the precision is calculated. Table III shows the average precision for both the distance measures. It can be observed that Manhattan distance gives better results (shown by bold numerals) for two image classes i.e. (Africans, and Building).

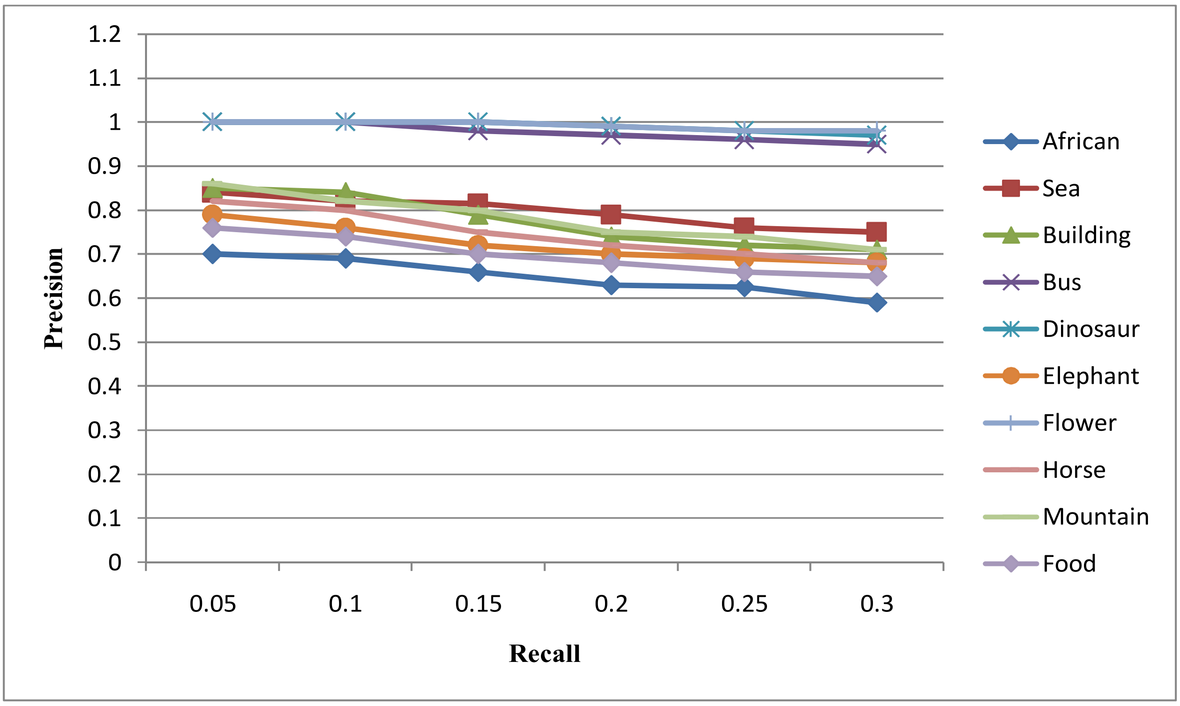


Fig. 10. Precision vs. Recall Plot

But Euclidean distance outperforms it in rest of classes. Hence we have considered Euclidean distance metric for our method.

TABLE III. PERFORMANCE COMPARISON FOR DISTANCE MEASURES

Average Precision		
Image Class	**Type of Distance Measure**	
	Manhattan Distance	**Euclidean Distance**
African	**0.69**	0.65
Sea	0.62	**0.69**
Building	**0.77**	0.65
Bus	0.86	**1**
Dinosaur	0.89	**1**
Elephant	0.43	**0.61**
Flower	0.76	**1**
Horse	0.49	**0.52**
Mountain	0.66	**0.69**
Food	0.48	**0.51**

V. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a content based image retrieval technique using shape and texture feature. The strengths of SWT and Edge Histogram Descriptor are utilized to achieve a higher accuracy in retrieving images. Use of Edge Histogram Descriptor makes the method robust by extracting the global as well as local composition of edges in the image. The integration of shape measures help in improving the retrieval accuracy.The proposed technique is compared with three different existing methods evaluated on the same database. Experimental results clearly depict the effectiveness and superiority of the proposed technique over these methods as well as other techniques utilizing colour information for retrieval. The proposed method gives an average accuracy of 77% on Wang's Image Database. As a part of future work we look forward to analyse the performance of our algorithm with different wavelet families.

REFERENCES

[1] P. Aigrain, H. Zhang, and D. Petkovic, " Content-Based Representation and Retrieval of Visual Media: A State of the Art Review", Multimedia Tools and Applications, Vol. 3, pp.179-202, 1996.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain,"Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 1349–1380, Dec-2000.

[3] D. Feng, W. C Siu, and H.J. Zhang, "Fundamentals of Content Based Image Retrieval, in Multimedia Information Retrieval and Management-Technological Fundamentals and Applications", New York: Springer, 2003.

[4] Rishav Chakravarti, Xiannong Meng, "A Study of Color Histogram Based Image Retrieval", Sixth International Conference on Information Technology: New Generations, pp.1323-1328, April-2009.

[5] Arjunan R.V, Kumar V.V, "Image Classification in CBIR Systems with Colour Histogram Features", International Conferences on Advances in Recent Technologies in Communication and Computing, pp.593-595, October-2009.

[6] Farsi H., Mohamadzadeh S., "Colour and Texture Feature Based Image Retrieval by using Hadamard Matrix in Discrete Wavelet Transform", Image Processing, IET, Vol.7, Issue. 3, pp.212-218, April-2013.

[7] Nivya Sasheedran, C. Bhuvaneswari, "An Effective Content Based Image Retrieval Approach Using Ripplet Transform", International conference on Circuits, Power and Computing Technologies, pp. 917-922, 2013.

[8] Wan Siti, H. Munirah, W. Ahmad, M. Faizal and A. Fauzi, "Comparison of Different Feature Extraction Techniques in Content- Based Image Retrieval for CT Brain Images," 10th IEEE workshop on Multimedia Signal Processing, pp. 503-508, 2008.

[9] Awais Adnan, Saleem Gul, Muhammad Ali, Amir Hanif Dar, "Content based Image Retrieval using Geometrical Shapes of Objects in Image", International Conference on Emerging Technologies, pp. 222-225, November-2007.

[10] Tian Yumin, Mei Lixia, "Image Retrieval Based on Multiple Features Using Wavelet," 5th IEEE International Conference on Computational Intelligence and Multimedia Applications, pp. 137-142, 2003.

[11] P. S. Hiremath, S. Shivashankar, J. Pujari, "Wavelet Based Features for Color Texture classification with Application to CBIR," International Journal of Computer Science and Network Security, Vol. 6, No.9A, September-2006.

[12] Yudong Zhang, Zhengchao Dong, Lenan Wu, Shuihua Wang, Zhenyu Zhou, "Feature extraction of brain MRI by Stationary Wavelet Transform", International conference on Biomedical Engineering and Computer Science, pp. 1-4, April 2010.

[13] G. P. Nason and B. W. Silverman-The stationary wavelet transform and some statistical applications in wavelet and statistics. In: Antoniadis A ed. Lecture Notes in Statistics. Berlin: Spinger Verlag, 281-299, 1995.

[14] S. J. Park, D. K. Park, C. S. Won, "Core experiments on MPEG-7 edge histogram descriptor," MPEG document M5984, Geneva, May, 2000.

[15] ISO/IEC/JTC1/SC29/WG11: "Core Experiment Results for Edge Histogram Descriptor (CT4)," MPEG document M6174, Beijing, July 2000.

[16] Chee Sun Won, Dong Kwon Park, and Soo-Jun Park, "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, Vol.24, no.1, February-2002.

[17] Dengsheng Zhang, Guojun Lu, "Review of Shape Representation and Description Techniques", Pattern Recognition, pp. 1-19, 2004.

[18] R.C Gonzalez, R.E Woods, Digital Image Processing, 3rd Edition, Prentice Hall, 2007.

[19] Sung-Hyuk Cha, Sargur N. Srihari, "On Measuring the Distance between Histograms", Pattern Recognition, pp.1355-1370, 2002.

[20] J. Z. Wang, "Wang Database," [Online], Available at: http://wang.ist.psu.edu/. Last accessed December 30, 2013.

[21] D. Tao, "The COREL Database for CBIR", [Online],Available at: https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval , Last accessed December 30, 2013.

[22] Hennig Muller, Wolfgang Muller, David McG. Squire, Stephane Marchand-Maillet, Thierry Pun, "Performance Evaluation in Content Based Image Retrieval: Overview and Proposals", Image/Video Indexing and Retrieval (Elsevier), Vol. 22, No. 5, pp. 593-601, April 2001.

[23] P. S. Hiremath and J. Pujari, "Content Based Image Retrieval using Color Boosted Salient Points and Shape features of an image", International Journal of Image Processing, 2(1):10-17, 2008.

[24] H. N. Duc, T. L. Tien, T. D. Hong, C. B. Thu and T. N. Xuan, "Image retrieval using contourlet based interest points", Proc. of the 10th Int. Conf. on Information Science, Signal Processing and their Applications (ISSPA 2010), 93-96, 2010, Kuala Lumpur.

[25] Manish Chowdhury, Sudeb Das and Malay Kumar Kundu, "Novel CBIR System Based on Ripplet Transform Using Interactive Neuro-Fuzzy Technique", Electronic Letters on Computer Vision and Image Analysis 11(1), pp.1-13, 2012.