DOI: 10.21917/ijivp.2016.0171

MULTILEVEL APPROACH OF CBIR TECHNIQUES FOR VEGETABLE CLASSIFICATION USING HYBRID IMAGE FEATURES

D. Latha[1], M. Mohamed Sathik[2] and Y. Jacob Vetha Raj[3]
[1,3]*Department of Computer Science, Nesamony Memorial Christian College, India*
E-mail: [1]dlathasatheesh@gmail.com, [3]jacobvetharaj@gmail.com
[2]*Sadakathullah Appa College, India*
E-mail: mmdsadiq@yahoo.com

Abstract

CBIR is a technique to retrieve images semantically relevant to query image from an image database. The challenge in CBIR is to develop a method that should increase the retrieval accuracy and reduce the retrieval time. In order to improve the retrieval accuracy and runtime, a multilevel CBIR approach is proposed in this paper. In the first level, the color attributes like mean and standard deviations are proposed to calculate on HSV color space to retrieve the images with minimum disparity distance from the database. In order to minimize search area, in the second level Local Ternary Pattern is proposed on images which were selected from the first level. Experimental results and comparisons demonstrate the superiority of the proposed approach.

Keywords

Content Based Image Retrieval (CBIR), Gray Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Local Ternary Pattern (LTP)

1. INTRODUCTION

In internet applications, every minute image is added to the databases of the server and also images are retrieved from the database. Hence, an image retrieval system becomes a need.

Early image retrieval techniques were generally text based. In this text-based image retrieval system [1], each image is manually annotated by the text descriptor and then this descriptor is used by the database management system to perform image retrieval. Generally, a lot of manpower is required for annotating text descriptors to every image of the large database. Hence it is a difficult and expensive work [3]. There is also a problem of annotation inaccuracy because of the subjectivity of human perception.

To overcome the problems in text-based system approach, content based image retrieval (CBIR) system was introduced in early 1980's. In this, the images are retrieved according to the image content. The primary goal of the CBIR system is to construct meaningful descriptor based on the image features like color, texture and shape of the images, which are automatically calculated by the computer to facilitate efficient and effective retrieval. In CBIR, color is the widely used feature to construct meaningful descriptor for the image [2].

Usually colors are defined in three dimensional color spaces. They could either be RGB (Red, Blue, Green), HSV (Hue, saturation, Value), or HSB Hue, Saturation, Brightness). The CBIR techniques using color feature usually compares the color histograms in the RGB or HSV color space. Generally, there are two types of color histograms, namely Global Color Histograms (GCHs) and Local Color Histograms (LCHs) [8]. A GCH represents one whole image with a single color histogram. An LCH divides an image into fixed blocks and takes the color histogram of each of these blocks.

Although retrieval methods using color feature information are widely used, it has the disadvantage of being very sensitive to changes in the histogram itself, including changes in brightness and color.

Texture [11] [8] is a feature that describes about the structural arrangement of the surface such as smoothness, coarseness, and regularity. The texture information is captured by several texture extraction methods. Gray Level Co-Occurrence Matrix (GLCM) is one of the widely used methods for texture feature extraction. The GLCM is used to extract certain properties about the spatial distribution of the gray level in the texture image. In order to estimate different gray level co-occurrence matrices, many statistical features like energy, entropy, contrast and homogeneity are extracted

Local Binary Pattern (LBP) [11] is one of the popular texture classification features. LBP encodes the pixel differences between the center pixel (Pc) and the neighboring pixels on a circle of radius. LBP is robust to illumination and contrast variations as it considers the pixel difference. But, it is sensitive to noise and small fluctuations of pixel values. To handle this problem, Local Ternary pattern (LTP) [11] has been introduced.

LTP uses two thresholds which creates three states as compared to two in LBP. So, LTP is more resistant to noise and small pixel variations. In our proposed system Local Ternary Pattern (LTP) is used to evaluate the texture feature.

Shape describes the configuration of an object, as outline or contour. It permits an object to be distinguished from its surroundings by its outline. Retrieval methods using shape feature [1] information distinguishes contours of objects contained in an image. It is not affected by object size and position. However, since contour of an object is sensitive to changes in shape or direction, contour extraction is difficult in itself. The most successful representations for shape categories are Fourier Descriptor and Moment Invariance: The main idea of Fourier Descriptor is to use the transformed boundary. Also, the idea of moment invariants is to use region based moments that are invariant to transformation.

With these image features, region based image retrieval methods using image segmentation were introduced into the CBIR to improve the retrieval accuracy of CBIR system. But the performance of these methods relies in the result of segmentation.

Generally, CBIR involves two steps. The first step is extracting image features like color, texture and shape (edge) to a distinguishable extent. The second step is matching these features to yield a result that is visually similar.

The problem involves entering an image as a query into an application that is designed to employ CBIR techniques in extracting visual properties to retrieve images in the database that are visually similar to the query image.

Our proposed system works in multilevel to reduce run time but also to improve the success classification ratio. Color attributes like mean and standard deviations on HSV color space is proposed in the first level. In the second level one of the texture classifier LTP is proposed.

The remainder of this paper is organized as follows: Section 2 discusses some of the earlier proposed research work on content based image retrieval. Section 3 discusses image features. Section 4 describes proposed work. Section 5 describes the experimental results and provides comparative performance. Finally, section 6 presents the conclusion.

2. RELATED WORKS

HSV color space and texture characteristics are used as feature for the image in the work proposed by Fan-Hui Kong [7]. In this he represents the one dimensional vector G by constructing a commutative histogram of the color characteristics of image after using non-interval HSV quantization for G. Texture features are extracted using Gray Level Co-occurrence Matrix (GCM) or Color Co-occurrence Matrix (CCM) respectively. Through the image retrieval experiment he indicates that the use of color features and texture based on CGM provides efficient retrieval.

P.S. Hiremath [1] presented a novel framework by combining color, texture and shape information for image retrieval and achieves higher image retrieval efficiency. In this paper, the image has been partitioned into equal sized non-overlapping tiles and the image features computed on these tiles serves as descriptors of color and texture. Gabor filter responses are used as texture features and color moments serve as color features. GVF (gradient vector flow) is used to obtain the edge image, which captures the object's shape information; GVF fields give excellent results in determining the object boundaries irrespective of the concavities involved. Invariant moments are used as shape features. A combination of these features provides a robust feature set for image retrieval.

Images are represented by three types of popular global features such as color, texture and shape in the work proposed by Wei Bian et al. [2]. HSV histogram is selected as color feature. Hue and saturation were both quantized into 8 bits but value into 4bits. For texture, a Pyramidal Wavelet Transfer (PWT) was extracted. For shape the edge directional histogram was calculated. When a query is given, the visual features are extracted, and then all images in the database are sorted on the base of Euclidean measure. If the user is satisfied with the results, the retrieval process comes to an end. However, most of the time, Relevance Feedback (RF) is required because of a poor retrieval performance. Biased Discriminate Embedding (BDEE) algorithm which is a reduction algorithm for relevant feedback procedure is also proposed. Also BDEE is compared with popular RF algorithms.

Jing Zhang et al. [3] explored a novel approach for color and texture based retrieval using three Region of Interest (ROI). To extract color features, the hue histograms of HSV color space have been used. The Gray Level Co-occurrence (GLCM) matrix is used to extract certain properties about the spatial distribution of the gray level in the texture image. In order to estimate different gray level co-occurrence matrices, many statistical features like energy, entropy, contrast and homogeneity are extracted. The processing of selecting ROIs is simple and the computation is fast because it makes partial match. It first uses K-means algorithm to segment image to three regions and then respectively select one ROI from every region. Finally extract color features and texture features from three ROIs. The similarity of two images will be determined by the similarities between pairs of ROIs. Retrieval performance is evaluated using nature images.

A user-oriented mechanism for CBIR method based on interactive genetic algorithm (IGA) is proposed by Chih-Chih [4]. In this work, color attributes like the mean value and the standard derivation and also the image bitmap of a color image are used as features for retrieval. Also, the entropy based on the gray level co-occurrence matrix and edge histogram of an image is considered as the texture features. Again to reduce the gap between the retrieval results and the use's expectation, the IGA (Interactive Genetic Algorithm) is employed to help the users to identify the images which are most satisfied to the user's need.

Amit Satpathy et al. [11], introduced two sets of edge-texture features, Discriminative Robust Local Binary Pattern (DRLBP) and Ternary Pattern (DRLTP) for object recognition.

Gwenole Quelle et al. [6], concentrated on image signature which is derived from one of the texture feature extraction methods, an adopted non separable wavelet transform and performance is compared with an adapted separable wavelet transform.

Although, many methods were proposed as in related works, each paper had its own shortcomings, especially in retrieval accuracy and execution time. So, in order to improve the retrieval accuracy and execution time of CBIR, a multi-level CBIR approach is proposed in this paper.

3. IMAGE FEATURES

The primary goal of the CBIR system is selecting appropriate image descriptor to facilitate efficient and effective retrieval of the image. Some of the important image features which are used in the proposed system are given below.

3.1 COLOR DESCRIPTOR

One of the important features that make the image recognition easily by human is color. The color image has been defined in three dimensional color space like RGB, HSV, and HSB and so on. But the HSV color system is considerably closer than the RGB system to the way in which humans experience and describes color sensation. So HSV color space is used in the proposed system.

The three color components of HSV are Hue, Saturation (lightness) and Value (brightness). Each component of the HSV color space contributes directly to the human conceptual understanding of colors. In HSV, Hue (H) distinguishes colors, Saturation (S) gives percentage of white added to pure color space and Value (V) represents the intensity of perceived light [2], [1]. The advantage of HSV color space is that it is closer to human conceptual understanding of colors and has the ability to separate chromatic and achromatic components. Mean (μ_t) and standard deviation (σ_t) for a color component of an image is defined as

$$\mu_t = \frac{1}{N}\sum_{i=1}^{N} P_i^t \tag{1}$$

$$\sigma_t = \left[\frac{1}{N-1}\sum_{i=1}^{N}\left(P_i^t - \mu\right)^2\right]^{1/2} \tag{2}$$

where, $t \in \{H, S, V\}$. Here P_i^t indicate the i^{th} pixel of t^{th} color component of HSV color space of an image, N is the total number of pixels in the image and μ_t and σ_t represent the mean and standard deviation of t color component in HSV color space of an image.

3.2 TEXTURE DESCRIPTOR

Texture is an image feature that describes about the structural arrangement of the surface. It refers to innate surface properties of the object and their relationship to the surrounding environment. In our proposed system Local Ternary Pattern (LTP) on the gray scale image is used to evaluate the texture feature. Local Binary Pattern (LBP) encodes the pixel differences between the neighboring pixels on a circle of radius and the center pixel (P_c).

For (3×3) neighbors, the pixel differences between the neighbors and the center pixel is calculated by (P_b-P_c), where P_c is the center pixel and P_b is the seven neighbors around P_c on a circle. In Fig.1, P_0, P_1, P_2, P_3, P_4, P_5, P_6 and P_7 are the neighbors around P_c.

P_0	P_1	P_2
P_7	P_c	P_3
P_6	P_5	P_4

Fig.1. Pixel representation of 3×3 image block

The LBP code at (x, y) is calculated as follows:

$$\text{LBP } x, y = \sum_{i=0}^{7} s(P_i - P_c)2^i \tag{3}$$

where, $s(z)$ is the threshold function.

$$s(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}$$

LBP is robust to illumination and contrast variations as it considers the pixel difference. But, it is sensitive to noise and small fluctuations of pixel values. To handle this problem, Local Ternary pattern (LTP) has been proposed.

LTP uses two thresholds which creates three states as compared to two in LBP. So, LTP is more resistant to noise and small pixel variations.

The LTP code at (x, y) is calculated as follows:

$$\text{LTP } x, y = \sum_{i=0}^{7} s'(P_i - P_c)3^i \tag{4}$$

where,

$$s'(z) = \begin{cases} 1, & z \geq T \\ 0, & -T < z < T \\ -1, & z \leq -T \end{cases}$$

Generally T is a user-defined value. In our proposed system T is selected empirically. The value of T is used as 7 in the proposed system as it gives better performance.

In the proposed system gray level co-occurrence matrix (GLCM) is calculated for the LTP values of each pixel. The GLCM calculates how often a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j.

The GLCM [4], [7] can be written as the probabilities of two pixels in an image, which are located with distance d and angle θ have gray levels i and j.

GLCM mathematically defined as, $p(i, j; d, \theta) = n\{(x_1, y_1)(x_2, y_2)\}$ such that,

$$g(x_1, y_1) = i,\ g(x_2, y_2) = j,\ |(x_1, y_1) - (x_2, y_2)| = d,$$

and

$$\angle((x_1, y_1), (x_2, y_2)) = \theta \tag{5}$$

where, 'n' denotes the number of occurrences of pixel in the gray image g, with i and j being the intensity levels of the pixels at position (x_1, y_1), and (x_2, y_2) respectively. In the proposed system GLCM is computed with $\theta = 0$ and $d = 1$.

The following figures Fig.2 and Fig.3 show the calculation of GLCM for the sample input of size 4 by 5. The Fig.2 shows the sample input. Element (1, 1) in the GLCM contains the value 1 because there is only one instance of horizontally adjacent inputs have the values 1 and 1 in the sample input. Element (1, 2) in the GLCM contains the value 2 because there is two instance of horizontally adjacent inputs have the values 1 and 2 in the sample input. Similarly all the values in the GLCM are calculated.

1	1	5	6	8
2	4	5	7	1
4	5	7	1	2
8	5	1	2	5

Fig.2. Sample input

	1	2	3	4	5	6	7	8
1	1	2	0	0	1	0	0	0
2	0	0	0	1	1	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	2	0	0	0
5	1	0	0	0	0	1	2	0
6	0	0	0	0	0	0	0	1
7	2	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0

Fig.3. GLCM values for the sample input

4. PROPOSED SYSTEM

In the proposed system, in order to improve the retrieval accuracy and runtime, a multilevel approach is introduced.

The proposed system collects the query image from the user in the first level. Then the color attributes like mean and standard deviations is calculated for the HSV color space of the query image. Similarly the same color attributes are calculated for the HSV color space of all database images. Then the proposed system evaluates the similarity between the query image and every database images using Euclidean distance based on the calculated color attributes. Then the system retrieves four images with smaller distance.

The system operates in five steps in level 1.

Querying: The user has to provide a sample image as the query for the system.

Color space: The system represents the query and all the database images in HSV color space.

Attribute evaluation: The system evaluates the mean and standard deviation for the query image and for all the database images.

Disparity computation: The system computes the disparity between the query and all database images by calculating the Disparity (*q*, *c*) as in Eq.(6).

Retrieval: The system retrieves some images with minimum disparity.

The disparity measure between the images is defined as,

$$\text{Disparity}(q,c)=\sqrt{\sum_{t\in\{H.S.V\}}\left(\mu_t^q-\mu_t^c\right)^2+\sum_{t\in\{H.S.V\}}\left(\sigma_t^q-\sigma_t^c\right)^2} \quad (6)$$

where, μ_t^q and σ_t^q represent the mean and standard deviation of the query image *q*, μ_t^c and σ_t^c represent the mean and standard deviation of the image *c* in the database.

In the second level, the proposed system calculates the LTP values for each pixel of query image by using Eq.(4). Then GLCM is calculated for the LTP values of query image. Similarly, LTP values are calculated for the images which were selected in the level 1 then GLCM is calculated for the LTP values of each pixel.

Then the system computes the disparity between the query image and every image selected in the first level using Euclidean distance based on the calculated GLCM values. Finally, the system retrieves an image with minimum disparity distance.

The disparity measure between the images is defined as,

$$\text{Disparity}(q,c)=\sum\left|\text{glcm}_\text{q}-\text{glcm}_\text{c}\right| \quad (7)$$

where, *q* denotes the query image and *c* denotes the database images. glcm_q denotes the GLCM value calculated for the query image on the LTP values. Similarly, glcm_c denotes the GLCM value calculated for the LTP values of database images.

5. EXPERIMENTAL RESULTS

The performance of the proposed system is evaluated with test image database under the system with i5 processor, 4GB RAM and 32-bit Windows 7 ultimate operating system. This test image database contains totally 99 images of 11 categories of vegetables. The database image is given in Fig.5.

5.1 PERFORMANCE OF THE PROPOSED SYSTEM

The performance of the proposed method is given in Table.1, 86% of success rate for 99 images when only color attributes like mean and standard deviation on HSV color space is used. In the same way, 74% of success rate when only LTP is used on the same set of images. But, the multilevel approach with color attributes like mean and standard deviations calculated on HSV color space of all database images in the first level and the Local Ternary Pattern (LTP) on the images selected from the first level gives 90% of success rate. Also the Table.1 shows that while using LBP method, 78% of successful completion occurred for the same set of input images.

The Table.2 also shows the run time of different methods. CBIR using only color attributes took 16.98 seconds, whereas only LTP took 24.67 seconds. But when both methods used in levels reduce the runtime as the search area is reduced in the second level. When only LBP is used it could take 24.01 seconds.

When color attribute and LBP values of query image are added and compared with the added values of database images using Euclidean distance, it gave 78% of successful retrieval and took 25.97 seconds of execution time. But color attributes added with LTP gave 75% of successful retrieval and it took 27.25 seconds.

But, the multilevel approach with color attributes like mean and standard deviations calculated on HSV color space of all database images in the first level and the Local Binary Pattern (LBP) on the images selected from the first level gives 88% of success rate and it took 25.17 seconds of runtime.

The Fig.4 shows the number of success for 100 vegetables using different methods. The Fig.6 shows the success rate for each vegetable category using different methods. Thus the superiority of the proposed system has been proved. The Table.4 shows the retrieval results at the end of level 1. The Table.5 shows the retrieval result at the end of level 2.

Table.1. Percentage of successful classification

Method	% of successful classification
Color attributes	86
LBP	78
LTP	74
Color Attributes +LTP	75
Color Attributes + LBP	78
Color attributes and LBP in multilevel	88
Color attributes and LTP in multilevel	90

5.2 COMPARISON WITH OTHER METHOD

In order to show the superiority of the proposed approach, the proposed method is compared with those in [4]. The author in [4] has used color and texture image feature. In that paper, mean and standard deviation have been used as color attributes. In addition, entropy on the GLCM and edge histogram is considered as texture descriptor values to classify the images. The experimental results are shown in Table.3. The comparison reveals that the proposed approach performs well than other methods. The vegetable categories Beans, Brinjal, Carrot, Dark Chillies and Garlic are classified well with 100%. But, for the Chillies and Ladies Finger the performance of the proposed method is little inferior to that obtained in [4].

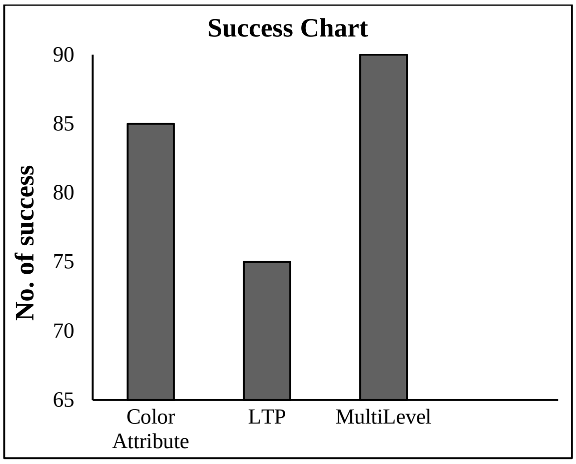

Fig.4. Number of success for 99 vegetables using different methods

Table.2. Average Run Time Comparison

Method	Runtime in seconds
Color attributes	16.96
LBP	24.01
LTP	24.7
Color Attributes + LTP	27.25
Color Attributes + LBP	25.97
Color attributes and LBP in multilevel	25.17
Color attributes and LTP in multilevel	23.92

Table.3. Success Rate using two methods

Categories	Chih-Chih [4]	Proposed Method
Banana	78	78
Beans	89	100
Brinjal	100	100
Carrot	89	100
Chillies	89	78
Chow-Chow	56	89
Dark Chillies	89	100
Garlic	100	100
Ladies finger	78	67
Onion	100	100
Tomato	67	89

Fig.5. Database Images

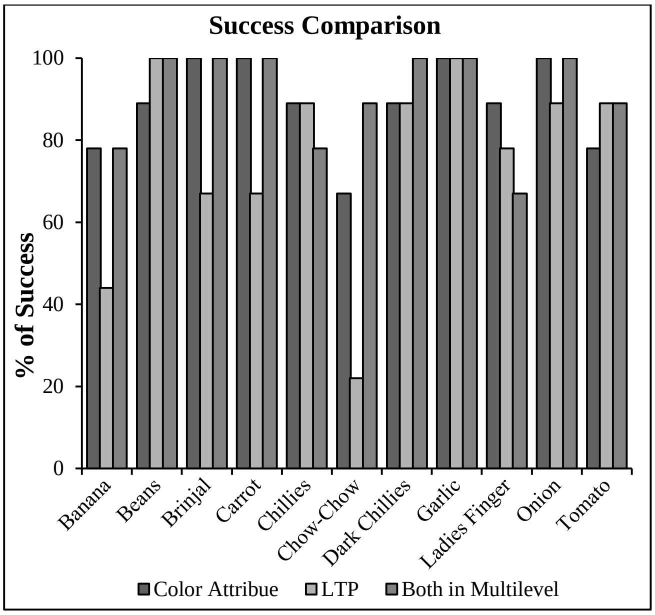

Fig.6. Success rate for different categories of vegetables using different methods

6. CONCLUSION

A multi-level approach for vegetable image classification using CBIR techniques is presented in this work. In the first level, color attributes like mean and standard deviation are calculated on HSV color space and that value is used as descriptor for image retrieval. In this level 86% of successful classifications have been obtained. While LTP is used in the second level, the performance is increased to 90%. Experimental results of proposed system have proved the improvement in retrieval performance and in run time. In future enhancement, efficient features might be identified and adopted in the proposed work as layers to improve the performance with 100% of successful classification.

Table.4. Retrieval results at the end of level 1

Query Images	**Images Retrieved from the Database at the end of level 1**			
				
				
				
				

Table.5. Retrieval result at the end of level 2

Query Images	**Retrieved Image from the Database at the end of level 2**
	
	
	
	

REFERENCES

[1] P.S. Hiremath and Jagadeesh Pujari, "Content based Image Retrieval using Color, Texture and Shape Features", *Proceedings of 15th International Conference on Advanced Computing and Communications*, pp. 780-784, 2007.

[2] Wei Bian and Dacheng Tao, "Biased Discriminant Euclidean Embedding for CBIR", *IEEE Transactions on Image Processing*, Vol. 19, No. 2, pp. 545-554, 2009.

[3] Jing Zhang, Choong-Woong Yoo and Seok-Wun Ha "ROI Based Natural Image Retrieval using Color and Texture Feature", *Proceedings of 4th International Conference on Fuzzy system and Knowledge Discovery*, pp. 740-744, 2007.

[4] Chih-Chih Lai and Ying- Chuan Chen, "A User-Oriented Image Retrieval System Based on Interactive Genetic Algorithm", *IEEE Transactions on Instrumentation and Measurement*, Vol. 60, No. 10, pp. 3318-3325, 2011.

[5] Gwenole Quellec, Mathieu Lamard, Guy Cazuguel, Beatrice Cochener and Christian Roux, "Adaptive Nonseparable Wavelet Transform via Lifting and its Application to Content Based Image Retrieval", *IEEE Transactions on Image Processing*, Vol. 19, No. 1, pp. 25-35, 2010.

[6] Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, pp. 62-66, 1979.

[7] Fan-Hui Kong, "Image Retrieval using both color and Texture feature", *Proceedings on International conference on Machine Learning and Cybernetics*, pp. 2228-2232, 2009.

[8] Subrahmanyam Murala, Anil Balaji Gonde and R.P. Maheshwari, "Color and Texture Features for Image Indexing and Retrieval", *Proceedings on IEEE International Advance Computing Conference*, pp. 1411-1416, 2009.

[9] Jeong-Yo Ha, Gye-Young Kim and Hyung-II Choi, "The Content-based Image Retrieval Method Using Multiple Features", *Proceedings of 4th International Conference on Network Computing and Advanced Information Management*, pp. 652-657, 2008.

[10] Youngeun An, Sungbum Pan and Jongan Park, "Image Retrieval Based on Color Tone Variance Difference Feature", *Proceedings on International Conference on Machine Learning and Cybernetics*, Vol. 7, pp. 3777-3780, 2008.

[11] Amit Satpathy, Xudong Jiang and How-Lung Eng, "LBP-based Edge-Texture Features for Object Recognition", *IEEE Transactions on Image Processing*, Vol. 23, No. 5, pp. 1953-64, 2014.

[12] Hatice Cinar Akakin and Metin N. Gurcan "Content based Microscopic Image Retrieval System for Multi-Image Queries", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 16, No. 4, pp. 758-769, 2012.