HSV-based Color Texture Image Classification using Wavelet Transform and Motif Patterns

Jun-Dong Chang[1], Shyr-Shen Yu[1], Hong-Hao Chen[2], and Chwei-Shyong Tsai[2,*]

[1] Department of Computer Science and Engineering,
National Chung Hsing University,
Taichung 402, Taiwan, ROC
phd9607@cs.nchu.edu.tw, pyu@nchu.edu.tw

[2] Department of Management Information Systems,
National Chung Hsing University,
Taichung 402, Taiwan, ROC
honghowc@gmail.com, tsaics@nchu.edu.tw*

Received 6 December 2009; Revised 6 January 2010; Accepted 9 January 2010

Abstract. In this paper, a novel color texture image classification based on HSV color space, wavelet transform, and motif patterns is introduced. Traditionally, RGB color space is widely used in digital images and hardware. However, RGB color space is not accurate in human visual perception and statistical analysis. Therefore, HSV color space is applied to obtain more accurate color statistics for extracting features. Due to extracting texture features in color texture images, wavelet transform and motif co-occurrence matrix are used in HSV color space for feature extraction. According to characteristic of wavelet transform, the horizontal, vertical and diagonal distributions are presented in sub-bands of a transformed image. Then, texture features of the horizontal, vertical and diagonal sub-bands are extracted by the motif co-occurrence matrix. After feature extraction, support vector machine (SVM) is applied to learn and classify texture classes by the extracted features. From experimental results, the proposed method is better and more correct than recent RGB-based color texture image classification.

Keywords: HSV color space, texture classification, wavelet transform, motif patterns, co-occurrence matrix, support vector machine

1 Introduction

Nowadays, texture analysis plays an important role in many image areas, such as geosciences and remote sensing, medical imaging, defect detection, document processing and image retrieval. Texture is a surface structure formed by uniform or non-uniform repeated patterns. The patterns also can be the perceived surface such as mineral, metal or wood which have tactile properties, or they could be reflectance on a surface such as color. In texture analysis, there are related issues such as texture classification, texture segmentation, and texture synthesis which are concerned by many researchers.

In last few decades, lots of texture classification techniques were proposed in these years. Firstly, the first and second orders statistics [1, 2] and co-occurrence matrix [3] were proposed for obtaining texture features. Further, model-based method such as Markov random field (MRF) [4, 5], Gibbs transform [6] and linear regression [7] are used to obtain distribution probabilities of textures on random fields. In addition, the local binary pattern (LBP) operator was also proposed to discriminate texture patterns by thresholding gray values of the neighboring pixels with binary codes [8]. Recently, multi-resolution methods such as wavelet transform is widely used and applied in texture analysis. Arivazhagan *et al*. proposed to use wavelet co-occurrence features and wavelet statistical features to discriminate texture classes [9]. Selvan *et al*. used singular value decomposition (SVD) on wavelet transform to model a probability density function for texture classification [10]. Moreover, neural networks and machine learning are also applied to learn and classify texture classes by wavelet and information theory features [11-13]. However, most of texture classification methods are proposed to gray-level texture images. Due to the complexity of color space, existing methods cannot perform well classification in color texture images. Therefore, it is an important and imperative work to study on the color texture classification.

This paper proposes a HSV-based color texture image classification using wavelet transform and motif patterns. The rest of this paper is organized as follows. Section 2 introduces HSV color space. Section 3 presents

* Correspondence author

wavelet transform and motif co-occurrence matrix for feature extraction. After that, support vector machine is applied to learn and classify texture images by extracted features in Section 4. Section 5 presents experimental results and performance comparisons. Finally, concluding remarks are given in Section 6.

2 HSV Color Space

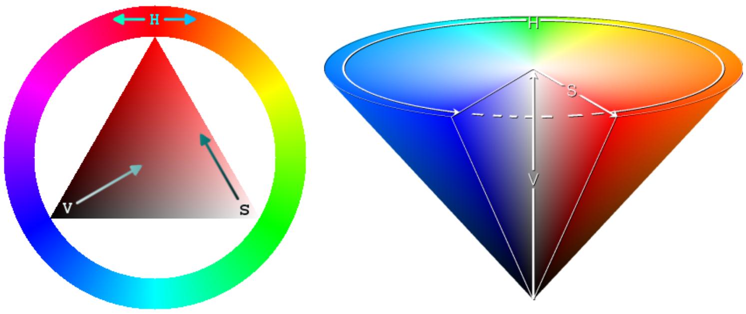


Fig. 1. HSV color model. [18]

In general, RGB color space is a general color space widely used in digital image display and optical instruments. However, RGB color space is not sensitive to human visual perception or statistical analysis. HSV color space is a non-linear transform from RGB color space that can describe perceptual color relationship more accurately than RGB color space. In this paper, HSV color space is adopted in following procedures.

In Fig. 1, HSV color space is formed by hue (H), saturation (S) and value (V). Hue denotes the property of color such as blue, green, red, and so on. Saturation denotes the perceived intensity of a specific color. Value denotes brightness perception of a specific color. Thus it can be seen that HSV color space is different from RGB color space in color variations. When a color pixel-value in RGB color space is adjusted, intensities of red channel, green channel, and blue channel of this color pixel are modified. That means color, intensity, and saturation of a pixel is involved in color variations. It is difficult to observe the color variation in complex color environment or content. However, HSV color space separates the color into hue, saturation, and value which means observation of color variation can be individually discriminated. According to above descriptions about HSV color space, it can obviously observe that HSV color space can describe color detail than RGB color space in color, intensity and brightness. In order to transform RGB color space to HSV color space, the transformation is described as follows:

$$h = \begin{cases} 0, & \text{if } \max = \min \\ (60^\circ \times \frac{g-b}{\max - \min} + 360^\circ) \bmod 360^\circ, & \text{if } \max = r \\ 60^\circ \times \frac{b-r}{\max - \min} + 120^\circ, & \text{if } \max = g \\ 60^\circ \times \frac{r-g}{\max - \min} + 240^\circ, & \text{if } \max = b \end{cases},$$

$$s = \begin{cases} 0, & \text{if } \max = 0 \\ \frac{\max - \min}{\max} = 1 - \frac{\min}{\max}, & \text{otherwise} \end{cases},$$

$$v = \max, \tag{1}$$

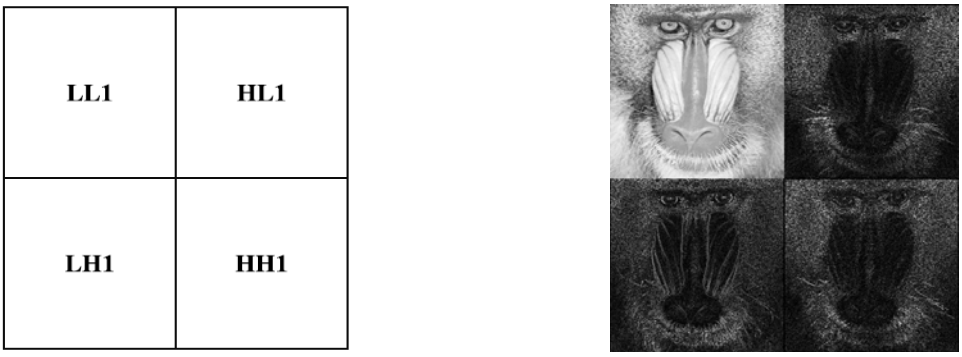
where r, g and b denote red, green and blue normalized in value [0, 1]. In order to quantize the range of the h plane for extracting features specifically, a quantized table is given as follows [14]:

$$\bar{h} = \begin{cases} 0, & \text{if } h \in [0,20) \cup [315,360] \\ 1, & \text{if } h \in [20,50) \\ 2, & \text{if } h \in [50,75) \\ 3, & \text{if } h \in [75,155) \\ 4, & \text{if } h \in [155,195) \\ 5, & \text{if } h \in [195,275) \\ 6, & \text{if } h \in [275,315) \end{cases}, \tag{2}$$

then the quantized hue plane is obtained. After quantizing the original hue plane, the values of s plane and v plane are normalized to $\bar{s}$ plane and $\bar{v}$ plane with value [0, 255] for coming feature extractions.

3 Feature Extraction

3.1 Wavelet Transform



(a) **(b)**

Fig. 2. Wavelet transform: (a) Level-1 decomposition, (b) A decomposed image “Baboon” by Level-1 transform

Wavelet transform (WT) is a multi-resolution decomposition which divides a spatial domain image into four sub-band images in frequency domain. According to recent researches, wavelet transform can observe the spatial information of an image. Due to the filter ability of wavelet transform, the input image (signal) is filtered by low pass and high pass filters. When the input image is filtered by low pass and high pass filters, the filtered signals represent spatial information within sub-band images. Therefore, texture images are transformed to sub-band images by wavelet transform before texture feature extraction. The wavelet transform of 1-dimension signal *f*(*x*) is defined as follow:

$$(W_\psi f)(a,b) = \int f(x)\psi^*_{(a,b)}(x)\mathrm{d}x, \\ \psi_{(a,b)} = \frac{1}{\sqrt{a}}\psi(\frac{x-b}{a}), \tag{3}$$

where a denotes the scaling factor, b is the translation parameter, and $\psi^*(x)$ is the transforming function, also called mother wavelet. Due to applying wavelet transform in a 2-D image, the transforming results are calculated by using a separable product of 1-D filter to the image:

$$\begin{aligned} LL &= \left\lfloor H_x * \left\lfloor H_y * I \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2}(b) \\ HL &= \left\lfloor G_x * \left\lfloor H_y * I \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2}(b) \\ LH &= \left\lfloor H_x * \left\lfloor G_y * I \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2}(b) \\ HH &= \left\lfloor G_x * \left\lfloor G_y * I \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2}(b) \end{aligned}, \tag{4}$$

where I is the input image, H_x and H_y are low pass filters, G_x and G_y are high pass filters, $b \in R^2$, * denotes the convolution operator, ↓2 denotes down-sampling operation. The four sub-band images are denoted as LL, LH, HL, and HH which include wavelet coefficients to present detailed characteristics of an image. According to sub-band images LH, HL, and HH, the image information including vertical, horizontal and diagonal features can be obtained from these three sub-band images respectively. If the image would be processed with further decomposition, the first level LL sub-band image is decomposed by above procedures. For this reason, the normalized $\bar{s}$ plane and $\bar{v}$ plane are decomposed by 1-level wavelet transform that denotes $\bar{s}_{\mathrm{LH1}}$, $\bar{s}_{\mathrm{HL1}}$, $\bar{s}_{\mathrm{HH1}}$ and $\bar{v}_{\mathrm{LH1}}$, $\bar{v}_{\mathrm{HL1}}$, $\bar{v}_{\mathrm{HH1}}$. After that, these six sub-band images and $\bar{h}$ plane are further extracted texture features by the motif co-occurrence matrix.

3.2 Motif Co-occurrence Matrix

Originally, the co-occurrence matrix is proposed by Haralick *et al*. to extract texture features [3]. By scanning each pixel with four orientations (i.e. 0, 45, 90, and 135 degrees) to calculate probabilities of neighboring pixels’ occurrences, a co-occurrence matrix can be obtained. According to the co-occurrence matrix, texture relations can be observed by probabilities of corresponding pixel-value pairs. Then, all pixel-value pairs’ probabilities of the co-occurrence matrix are calculated to obtain 14 texture features (i.e. Angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation, and maximal correlation coefficient) for texture analysis. However, the co-occurrence matrix only observes the pixel distribution for an image globally. Texture discrimination would be weak when the texture content is not regular. In 2004, Jhanwar *et al*. proposed the motif co-occurrence matrix using in content-based image retrieval [15]. By using the motif co-occurrence matrix (MCM), the distribution of local area of an image could be noticeably observed and discriminated. Six motifs (shown in Fig. 3) are different scanning paths to determine which motif does each 2 × 2 grid belongs to. Each grid is scanned from its top-left pixel and then searches its close pixel by minimum pixel-value difference. Before scanning, six motifs are labeled as 0 to 5 respectively. For example, in Fig. 4, the top-left grid contains four pixels {202, 53, 78, 55}, and then the scanning path goes through {202, 78, 55, 53} sequentially. According to this scanning path, the top-left grid is determined as "U motif" and labeled "2" in a motif map. The rest of grids are scanned as above procedures to obtain a motif map of this image (shown in Fig 4(b)). Therefore, the local variance of an image can be observed by the motif map. Further, the motif map is scanned by the co-occurrence matrix technique to generate a motif co-occurrence matrix to obtain texture features.

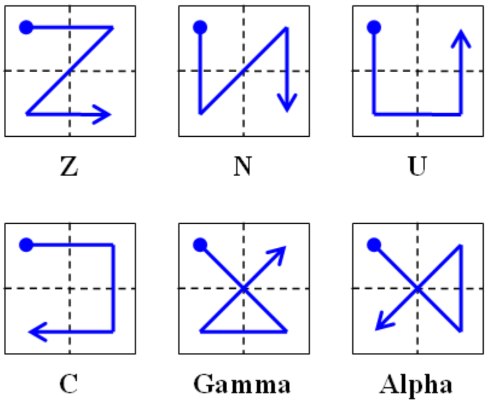
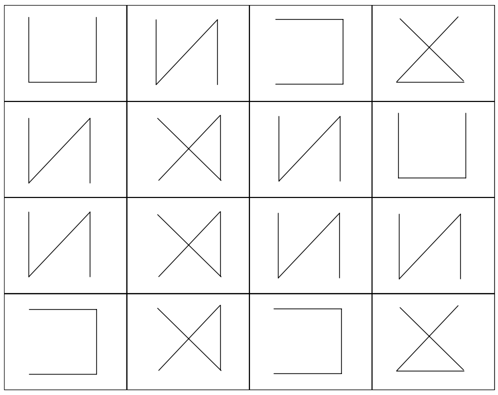


Fig. 3. Six motifs with 2 × 2 grids

202	53	149	54	255	255	255	124
78	55	84	52	57	190	186	250
129	68	35	128	160	38	36	255
183	29	140	68	54	31	144	182
176	52	47	43	47	53	145	156
145	38	61	45	40	62	140	176
150	186	98	488	220	211	87	167
99	196	189	174	155	159	151	106

(a)



(b)

Fig. 4. Example of scan motifs (a) An 8 × 8 image, (b) A motif map of (a)

Then, a motif co-occurrence matrix $P(i, j)$ which i and j = 0, 1 …, 5 is generated by scanning a whole motif map. According to the motif co-occurrence matrix, the features such as contrast, energy and entropycan be calculated and obtained by follow equations:

$$\begin{aligned} &Contrast: C = \sum_{i,j=0}^{N} (i-j)^2 P(i,j), \\ &Energy: E = \sum_{i,j=0}^{N} P(i,j)^2, \\ &Entropy: H = -\sum_{i,j=0}^{N} P(i,j)\log_2 P(i,j), \end{aligned} \tag{5}$$

where N denotes the number of motifs (or color bins).

In this paper, the $\bar{h}$ plane and six sub-band images $\bar{s}_{LH1}$, $\bar{s}_{HL1}$, $\bar{s}_{HH1}$ and $\bar{v}_{LH1}$, $\bar{v}_{HL1}$, $\bar{v}_{HH1}$ for $\bar{s}$ plane and $\bar{v}$ plane are scanned to generate its own motif co-occurrence matrix respectively. Then, 21 features are obtained by calculating these matrices by Eq. (5). These 21 features also can be formed to be a feature set *S* to input to support vector machine to learn and classify texture classes. The details of texture classification are described in Section 4.

4 Texture Classification

In this section, support vector machine (SVM) is applied to train the extracted texture features by following procedures. Support vector machine (SVM), a kind of supervised learning algorithm used for classification and regression, is proposed by Vapnik and collaborators [16, 17]. Due to the significant learning and classification ability, support vector machine is generally applied to bioinformatics, pattern recognition and data mining. Support vector machine maps input vectors to a higher dimensional space where a maximal separating hyperplane is constructed. Two parallel hyperplanes are constructed on each side of the hyperplane that separates the data. The separating hyperplane is the hyper-plane that maximizes the distance between the two parallel hyperplanes. Therefore, the benefit of SVM is that it can perform well without many training samples. Let a sample dataset be described as:

$$S = \{(\mathrm{x}_i, y_i) \mid i = 1,2,...,l\},\ \mathrm{x}_i \in \mathrm{R}^n,\ y_i \in \mathrm{R}, \tag{6}$$

where $\mathrm{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ is the *i*th input pattern with *n* tuples; $y_i \in \{1,-1\}$, is the *i*th label for x_i. x_i belongs to either of two classes. Therefore, SVM finds a hyperplane *H* for *S* also can be represented as:

$$H : \mathrm{w} \cdot \mathrm{x}_i^t + b = 0, \tag{7}$$

where t is the transposing operation, $\mathrm{w} = (w_1, w_2, \ldots, w_n)$ is the weight vector, and *b* is a bias. In order to get an optimal value for (w, *b*) to maximize the margin ρ which is $1/\|\mathrm{w}\|^2$ between $\mathrm{w} \cdot \mathrm{x}_i + \mathrm{b} = 1$ and $\mathrm{w} \cdot \mathrm{x}_i + \mathrm{b} = -1$, the optimal value for (w, *b*) can be obtained by solving the follow condition:

$$\begin{aligned} &\min\ \frac{1}{2}\mathrm{w}\mathrm{w}^t, \\ &\text{s.t.}\ \ y_i(\mathrm{w}\mathrm{x}_i^t + b) \geq 1,\ i = 1,2,...,l. \end{aligned} \tag{8}$$

The condition is also called prime problem, we can use the Karush-Kuhn-Tucker (KKT) theorem and LaGrange multiplier method to solve the dual problem as follow:

$$\begin{aligned} &\max\ \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j K(\mathrm{x}_i^t\mathrm{x}_j), \\ &\text{s.t.}\ \ \sum_{i=1}^{l}\alpha_i y_i = 0,\ \text{where}\ \alpha_i \geq 0,\ i = 1,2,...,l, \end{aligned} \tag{9}$$

where α_i denotes LaGrange multipliers associated with Eq. (9), $K(\mathrm{x}_i^t\mathrm{x}_j)$ is the kernel function. Finally, the optimal decision function $f(\mathrm{x})$ can be defined as:

$$f(\mathrm{x}) = \mathrm{sign}(\sum_{i=1}^{l}\alpha_i y_i \mathrm{x}_i \mathrm{x}^t + b). \tag{10}$$

In this paper, the extracted feature set *S* is inputted to support vector machine to learn and classify texture classes. Then, the following section presents and discusses the performance of the proposed method.

5 Experimental Results

In this section, a color texture image database (University of Oulu texture database, 2005, shown in Fig. 5) is used in the experiments. Sixteen RGB color texture images of 512 × 512 pixels size are used in the experiments for different classes. Then, 100 sub-images are generated by randomly dividing a texture image with 128 × 128 pixels size for each class. That means 1,600 sub-images are generated for 16 classes which include 320 sub-images for training and 1,280 sub-images for testing. First, all of RGB color sub-images are transformed to HSV color sub-images formed by *h*, *s*, and *v* plane. According to the ranges of HSV are $h \in [0,360]$, $s \in [0,1]$, $v \in [0,1]$, these three planes are normalized to $\bar{h} \in [0,6]$, $\bar{s} \in [0,255]$, $\bar{v} \in [0,255]$. Then, the normalized $\bar{s}$ plane and $\bar{v}$ plane are decomposed by wavelet transform to obtain sub-band images denoted as $\bar{s}_{LH1}$, $\bar{s}_{HL1}$, $\bar{s}_{HH1}$ and $\bar{v}_{LH1}$, $\bar{v}_{HL1}$, $\bar{v}_{HH1}$, respectively. After obtaining sub-band images, the $\bar{h}$ plane and six sub-band images are scanned to generate the motif co-occurrence matrices to calculate homogenous features for SVM training and classification.

Fig. 5. Sixteen color texture images; from left to right and top to bottom: Grass, Flowers1, Flowers2, Bark1, Clouds, Fabric7, Leaves, Metal, Misc, Tile, Bark2, Fabric2, Fabric3, Food1, Water, and Food2

Table 1. Comparisons of classification rates (%) for three methods.

Image Type	Methods		
	The Pro-posed Method	WSF-WCF [10]	W-ANFIS [12]
Grass	96%	86%	92%
Flowers1	99%	90%	98%
Flowers2	96%	89%	90%
Bark1	98%	90%	94%
Clouds	100%	97%	100%
Fabric7	97%	92%	91%
Leaves	100%	92%	98%
Metal	93%	90%	90%
Misc	99%	93%	94%
Tile	97%	95%	94%
Bark2	97%	90%	92%
Fabric2	100%	97%	99%
Fabric3	100%	99%	99%
Food1	99%	95%	97%
Water	100%	100%	100%
Food2	95%	90%	89%

In this experiment, two wavelet based texture classification methods are compared to the proposed method, which are the WSF-WCF method and the W-ANFIS method. The WSF-WCF method which is proposed by Arivazhagan *et al*. uses wavelet statistical features and wavelet co-occurrence features for texture classification. On the other hand, the W-ANFIS method also uses wavelet transform and information theory to obtain features, and then adaptive neuro-fuzzy inference system is applied to learn and classify texture classes. In related researches of texture classification, the correct classification rate is calculated by the first order of classified texture images which is belonged to its true class or not. In Table 1, it can obviously observe that the correct classification rate of the proposed method is better than the other methods. Due to properties of the HSV color system, color, intensity and brightness of the texture images can be observed particularly. Comparing to the WSF-WCF method, the proposed method uses less features to achieve high correct classification rate. On the other hand, although the W-ANFIS uses fewer features but RGB color space is not sufficient for better perceptual vision in feature extraction. Also, the training process of W-ANFIS takes more time in selecting optimal neurons and parameters. In brief, the proposed method achieves better correct classification rate with fewer features and low time-consuming than the WSF-WCF method and the W-ANFIS method.

6 Conclusions

In this paper, a HSV-based color texture image classification technique using wavelet transform and motif patterns is proposed. According to perceptual ability of HSV color space, RGB color texture images are transformed to HSV color space. Then, normalized HSV planes are scanned by motifs to generate three motif co-occurrence matrices to obtain texture features. By using support vector machine, texture images are learned and classified by extracted features. From experimental results, the proposed method is better than the WSF-WCF method and the W-ANFIS method in correct classification rates. According to the performance of the proposed method, it can be applied to medical image recognition such as breast ultrasound images or bone age assessment for texture analysis in future works.

7 Acknowledgement

This work is supported in part by the Ministry of Education, Taiwan, R.O.C. under the ATU plan.

References

[1] J. S. Weszka, C. R. Dyer, A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," *IEEE Transaction on Systems, Man, and Cybernetics*, Vol. 6, No. 4, pp. 269-286, 1976.

[2] O. D. Faugeras and W. K. Pratt, "Decorrelation Methods of Texture Feature Extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, No. 4, pp. 323-332, 1980.

[3] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on System, Man and Cybernetics*, Vol. 3, No. 6, pp. 610-621, 1973.

[4] G. R. Cross and A. K. Jain, "Markov Random Field Texture Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 5, No. 1, pp. 25-39, 1983.

[5] R. L. Kashyap and A. Khotanzed, "A Model based Method for Rotation Invariant Texture Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 4, pp. 472-481, 1986.

[6] H. Derin and H. Elliot, "Modeling and Segmentation of Noisy and Textured Images using Gibbs Random Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 1, pp. 39-59, 1987.

[7] Z. Z. Wang and J. H. Yong, "Texture Analysis and Classification with Linear Regression Model based on Wavelet Transform," *IEEE Transactions on Image Processing*, Vol. 17, No. 8, pp. 1421-1430, 2008.

[8] H. Zhou, R. Wang, C. Wang, "A Novel Extended Local-binary-pattern Operator for Texture Analysis," *Information Sciences*, Vol. 178, No. 22, pp. 4314-4325, 2008.

[9] S. Arivazhagan and L. Ganesan, "Texture Classification using Wavelet Transform," *Pattern Recognition Letters*, Vol. 24, No. 9-10, pp. 1513-1521, 2003.

[10] S. Selvan and S. Ramakrishnan, "SVD-based Modeling for Image Texture Classification using Wavelet Transformation," *IEEE Transactions on Image Processing*, Vol. 16, No. 11, pp. 2688-2696, 2007.

[11] A. Sengur, "Wavelet Transform and Adaptive Neuro-fuzzy Inference System for Color Texture Classification," *Expert Systems with Applications*, Vol. 34, No. 3, pp. 2120-2128, 2008.

[12] A. Sengur, I. Turkoglu, M. C. Ince, "Wavelet Packet Neural Networks for Texture Classification," *Expert Systems with Applications*, Vol. 32, No. 2, pp. 527-533, 2007.

[13] I. Turkoglu and E. Avci, "Comparison of Wavelet-SVM and Wavelet-adaptive network based Fuzzy Inference System for Texture Classification," *Digital Signal Processing*, Vol. 18, No. 1, pp. 15-24, 2008.

[14] X. Zhu, J. Zhao, J. Yuan, H. Xu, "A Fuzzy Quantization Approach to Image Retrieval based on Color and Texture," *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication (ICUIMC2009)*, Suwon, Korea, pp. 141-149, 2009.

[15] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, "Content-based Image Retrieval using Motif Cooccurrence Matrix," *Image and Vision Computing*, Vol. 22, No. 14, pp. 1211-1220, 2004.

[16] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.

[17] V. Vapnik, *Statistical Learning Theory*, John Wiley, New York, 1998.

[18] http://en.wikipedia.org/wiki/HSL_and_HSV