International Journal of Innovative
Computing, Information and Control
Volume 8, Number 1(B), January 2012
 ISSN 1349-4198
pp. 659–676

NOVEL SHAPE-TEXTURE FEATURE EXTRACTION FOR MEDICAL X-RAY IMAGE CLASSIFICATION

Seyyed Mohammad Mohammadi, Mohammad Sadegh Helfroush and Kamran Kazemi

Faculty of Electrical and Electronic Engineering
Shiraz University of Technology
Modarres Bolvd. P.O. Box, Shiraz 71555-313, Iran
mohammad.mohammadi@gmail.com; { ms_helfroush; kazemi }@sutech.ac.ir

Received September 2010; revised June 2011

Abstract. *Since medical x-ray images are gray scale images with almost the same texture characteristics, conventional color or texture features cannot be used for appropriate categorization in medical x-ray image archives. Therefore, for this type of images with these special conditions, more complicated methods are needed to improve the classification results. In this paper, with regard to medical x-ray images characteristics and also to avoid the complexity, a novel feature is proposed which is the combination of shape and texture features. The feature extraction process is started by edge and shape information extraction from original medical x-ray images. Finally, Gabor filter is used to extract spectral texture features from shape images. Furthermore, in order to study the effect of feature fusion on the classification performance, different effective features like local binary pattern and gray level co-occurrence matrix are utilized to combine with novel features. In order to evaluate the proposed shape-texture feature and feature combination in medical x-ray images categorization, a set of well-known classifiers such as Euclidean distance matching criterion, probabilistic neural network (PNN) and support vector machine (SVM) are applied. The proposed feature is extracted from 4402 medical x-ray images in 21 classes of IRMA database and the best classification accuracy rate becomes 88.77% by SVM. The classification accuracy rate of 94.2% for combination of novel shape-texture feature, local binary pattern and gray level co-occurrence matrix has been obtained by SVM.*
Keywords: Classification accuracy rate, Feature extraction, Feature combination, Medical x-ray images, Novel shape-texture feature (NSTF)

1. **Introduction.** Traditional classifiers and retrieval systems are based on text or keywords. This means that keywords which explain contents of images are attached to the images and guide the traditional search engines such as Google, Yahoo and AltaVista search sites. With increment of multimedia information, annotation process would be considerably time consuming. Furthermore, this process is an ambiguous task because of different human perceptions. To solve this problem, researchers focused on the actual contents of images. To access images based on their contents, several features such as color, shape and texture are extracted from them. Content-based image retrieval (CBIR) systems [1-4] are playing an important role in the clinical processes applications. Classification and indexing stages are main primary stages for CBIR systems which should be done prior to any processing to offer a reliable and appropriate retrieval. On the other hand, confident classification plays a great role in obtaining an effective retrieval. The applications of x-ray images for medical diagnostics are explosively increasing; therefore, the need for powerful and reliable archiving, classification and retrieval tools for this type of images has become a great and undeniable concern. In medical x-ray image classification

application, several schemes and algorithms have been proposed in the literature [5-9]. In [5], a hierarchical medical image classification method including two levels using a set of various shape and texture features was proposed. These features were moment invariants, Fourier descriptor, major axis orientation, eccentricity, major and minor axis length, gray level co-occurrence matrix (GLCM), the tessellation-based spectral feature in multi-scale space and the directional histogram in multi-scale space. These nine features were analyzed separately and the best feature for each class was selected through assigning a weight to it. In each level of the hierarchical classifier, using a new merging scheme and multi-layer perceptron (MLP) classifiers (merging-based classification), homogenous (semantic) classes were created from overlapping classes in the database. The proposed algorithm was evaluated on a database consisting of 9100 medical x-ray images of 40 classes. It provides on accuracy rate of 90.83% on 25 merged classes in the first level. In [6], an automatic detection of body parts in x-ray images was proposed. In the first step of this scheme, average gray descriptor (AGD), color layout descriptor (CLD), edge histogram descriptor (EHD) and local binary patterns (LBP) features were extracted and then the performance of these five different feature types was evaluated by support vector machine (SVM). The proposed feature extraction scheme was implemented on a set of radiographs of 116 classes from IRMA database. 84.7% was the best classification accuracy rate obtained by evaluating local binary pattern. In [7], a classification method of x-ray images using grid approach was proposed. This scheme consisted of 6 stages: object segmentation, identification of the object boundary, transformation using grid approach, representation using Freeman code, calculation and matching process. The best classification accuracy rate obtained by Jeffrey divergence was 83%. In [8], x-ray chest image retrieval method based on feature fusion was presented. In this scheme, color auto-correlogram, dominant color of partition, GLCM, gray-gradient co-occurrence matrix and shape invariant moments were extracted as retrieval features. After comparing retrieval results of these features, feature fusion and relevance feedback were applied. This framework was evaluated on 122 x-ray chest images and the best obtained classification accuracy rate was 74%. In [9], a multilevel feature extraction and medical x-ray images classification scheme was introduced. In this scheme, first, features were extracted in three levels: global, local and pixel. Then they were combined to generate one big feature vector. This feature was extracted from 9000 training medical x-ray images and 1000 test ones in 57 classes and evaluated by both SVM and K-nearest neighborhood (KNN) classifiers. The best achieved classification accuracy rate was 89%. The significant and important shortage of [5-9] is that the number of training images utilized for classification is much larger than the number of test ones. In addition, in [5], assigning special weight to features, using few medical x-ray image classes in [7,8], shows that utilized features are not general and powerful enough to cover all varieties of images. Furthermore, in all mentioned methods, the scale of images is an important point, and to extract practical features all the image scales have to be normalized. The aim of this paper is to introduce an applicable and effective novel feature extraction framework for medical x-ray images based on their special characteristics and independent to their scales and gray levels. Medical x-ray images are gray-scale images with quite similar texture characteristics; thus, color and texture would not be appropriate features to classify them. Furthermore, as shown in Figure 1, many of x-ray images categories, have similar shapes, e.g., cervical spine, coronal cranium and coronal facial cranium, or even have unidentifiable shapes, e.g., pelvis angiography. Therefore, it is difficult to find appropriate and applicable features based on only shape features. With regard to these conditions and characteristics, in this paper, a novel feature extraction framework based on both shape and texture features for medical x-ray images categorization is proposed. In this framework, first, the shape and directional information are extracted from original images carefully. Then, texture feature is extracted from shape and directional information and makes the shape-texture feature. This feature, called novel shape-texture feature (NSTF), is independent of image scales and gray level intensities. Experimental results show that the proposed NSTF is an applicable and appropriate classification criterion in comparison with other traditional features in medical x-ray images classification. In order to demonstrate the capability and accuracy of NSTF, three different classifiers consisting of Euclidean distance matching criterion, probabilistic neural network (PNN) and SVM are utilized.

(a) (b) (c) (d)

Figure 1. (a) Coronal cranium; (b) cervical spine; (c) coronal facial cranium; (d) coronal pelvis angiography by cardiovascular system. (a)-(c) are three sample classes with similar shape and (d) is a class with uncertain shape.

This paper is organized as follows. In Section 2, the most commonly used conventional features in medical application are introduced. In Section 3, the NSTF is entirely explained. In this section, all fundamental steps of the proposed feature are brought. In Section 4, the utilized classification algorithms, i.e., Euclidean distance matching criterion, PNN and SVM are described. Experimental results are brought in Section 5, and Section 6 is designated to discussion and conclusion.

2. **Conventional Features in Medical CBIR Applications.** There are several feature extraction methods for different applications. However, in medical x-ray image processing and classification application we have to use special features correspond to their characteristics. As the medical x-ray images are gray scale images, the color feature would not be appropriate for classification and we are ineluctable to use shape or texture feature extraction methods. In this section, some of known and widely used shape and texture features in medical CBIR application are introduced.

2.1. **Gray level co-occurrence matrix.** Gray level co-occurrence matrix (GLCM) [10], as one of the most known texture analysis methods, estimates image properties related to second-order statistics. Each entry (x, y) in GLCM corresponds to the number of occurrences of the pair of gray levels i and j which are a distance d apart in original image. Assume $I(i, j)$ is an image with size $P \times Q$ and a set of N_g gray levels, co-occurrence matrix $F(x, y, d, \theta)$ is defined as follows:

$$\begin{aligned} F(x,y,d,\theta) = \sharp\{&((i_1,j_1),(i_2,j_2)) \\ &\in (P\times Q)\times(P\times Q) \mid (i_2,j_2) \\ = &(i_1,j_1)+(d\cos\theta, d\sin\theta), I(i_1,j_1)=i, I(i_2,j_2)=j, \\ &1 \le x,y < N_g\} \end{aligned} \tag{1}$$

where d and θ denote the distance and the orientation aligning between pixels (i_1, j_1) and (i_2, j_2) in the image, respectively. "$\sharp$" denotes the number of elements in the set. In order to estimate the similarity between different gray level co-occurrence matrices, Haralick [10] proposed 14 statistical features. To reduce the computational complexity, only some of these features are selected. The description of 4 most relevant features that are widely used in literature [5,11] is given in (2). Here, the distance d is set to 50 and the θ varies in 4 directions 0, 45, 90, 135 degree.

$$
\begin{aligned}
\text{Energy}(d,\theta) &= \sum_{x=1}^{N_g}\sum_{y=1}^{N_g} F(x,y,d,\theta)^2 \\
\text{Contrast}(d,\theta) &= \sum_{x=1}^{N_g}\sum_{y=1}^{N_g} (i-j)^2 F(x,y,d,\theta) \\
\text{Correlation}(d,\theta) &= \frac{\sum_{x=1}^{N_g}\sum_{y=1}^{N_g}(x-m_i)(y-m_j)F(x,y,d,\theta)}{\sigma_i\sigma_j} \\
\text{Homogenity}(d,\theta) &= \sum_{x=1}^{N_g}\sum_{y=1}^{N_g} \frac{1}{1+\mid x-y\mid} F(x,y,d,\theta)
\end{aligned}
\tag{2}
$$

where $m_i(m_j)$ and $\sigma_i(\sigma_j)$ are mean and standard deviation of pixels value in row (column) direction of the GLCM, respectively.

2.2. **Moment invariants.** Moment invariant [3,5] was first introduced by Hu [1]. It was derived from the theory of algebraic invariant. This technique is chosen to extract image features which the generated features are Rotation Scale Translation (RST)-invariant. Two-dimensional moments of a digitally sampled $P \times Q$ image $I(i,j)$ is given as:

$$
m_{pq} = \sum_{i=0}^{P-1}\sum_{j=0}^{Q-1}(i)^p(j)^q I(i,j) \qquad p,q = 0,1,2,3,\ldots \tag{3}
$$

The moments $I(i,j)$ shifted by an amount (a,b) are defined as:

$$
\mu_{pq} = \sum_{i=0}^{P-1}\sum_{j=0}^{Q-1}(i+a)^p(j+b)^q I(i,j) \qquad p,q = 0,1,2,3,\ldots \tag{4}
$$

Thus, the central moments μ_{pq} can be computed from (4) on substituting $a = -\bar{i}$ and $b = -\bar{j}$ as $\bar{i} = \frac{m_{10}}{m_{00}}$ and $\bar{j} = \frac{m_{01}}{m_{00}}$,

$$
\mu_{pq} = \sum_{i=0}^{P-1}\sum_{j=0}^{Q-1}(i-\bar{i})^p(j-\bar{j})^q I(i,j) \qquad p,q = 0,1,2,3,\ldots \tag{5}
$$

When a scaling normalization is applied the central moments change as:

$$
\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \qquad \gamma = \left[\frac{p+q}{2}\right] + 1 \tag{6}
$$

In particular, Hu defines seven values, computed by normalizing central moments through order three, that are invariant to object scale, position and orientation. In terms of the central moments, the seven moments are given as follows:

$$\begin{aligned}
M_1 &= \eta_{20} + \eta_{02} \\
M_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
M_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
M_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
M_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
M_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
M_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad +(3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]
\end{aligned} \tag{7}$$

2.3. **Region properties.** The common region properties, [5,12], are perimeter, eccentricity, Euler number, area, major axis length, minor axis length and orientation of the image. The major and minor axis lengths are lengths of the major and minor axes of the ellipse that has the same normalized second central moments as the region or an object, respectively. The eccentricity can be defined as the ratio of the smallest eigenvalue to the largest one. The Euler number is the subtraction of number of connected components and number of holes in the object or region. The orientation can be defined as the direction of the largest eigenvector of the second order covariance matrix of a region or an object. The perimeter is the number of pixels located on the object boundary and the area is the number of pixels located in the region within the boundary.

2.4. **Tamura feature.** There are 6 different Tamura features: coarseness, contrast, directionality, linelikeness, regularity and roughness [13]. In the literature [14], the first three features are used since they are strongly correlated with human perception.

2.5. **Entropy.** Entropy is a general statistical parameter. Log-Energy entropy, Sure entropy and Shanon entropy [15,16], shown in (8), are generally used to extract features [11].

$$\begin{aligned}
&\text{Shanon entropy} = -\sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} I(i,j)^2 \log\left(I(i,j)^2\right) \\
&\text{Sure entropy} = N - \sharp\{(i,j) \mid I(i,j) \le \tau\} + \sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} \min\left(I(i,j)^2, \tau^2\right) \\
&\text{Log-energy entropy} = -\sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} \log(I(i,j)^2)
\end{aligned} \tag{8}$$

where $I(i,j)$, $i = 0,\dots,P-1$ and $j = 0,\dots,Q-1$, is the image with size $P \times Q$, τ is a constant where $\tau \ge 1$, N is the total number of pixels in image I and "$\sharp$" denotes the number of elements in the set. As the information theoretic features like Entropy, depend on gray level intensities, the images gray level should be normalized before feature extraction until exact and accurate features are generated. Hence, to normalize the gray level intensities, (9) is utilized. If $I(i,j)$ is the original image pixel, the normalization equation is:

$$J(i,j) = \begin{cases} m + \sqrt{\dfrac{v \times (I(i,j) - m_0)^2}{v_0}} & \text{if } I(i,j) > m, \\ m - \sqrt{\dfrac{v \times (I(i,j) - m_0)^2}{v_0}} & \text{if } I(i,j) < m, \end{cases} \tag{9}$$

where $m_0(m)$ and $v_0(v)$ are the initial (desired) mean and variance of original image I, respectively and $J(i,j)$ is the normalized image. In this paper, the desired mean and variance are selected 100 and 2000 respectively. After normalization, three different types of Entropy features are extracted which provide a three element feature vector.

2.6. **Local binary pattern.** Local Binary Pattern (LBP) [18] is a simple and efficient texture descriptor with low computational complexity. It labels the pixels of input image by thresholding the neighborhood of each pixel with the value of the center pixel and considering the results as a binary number. The neighborhood is formed by a symmetric neighbor set of P pixels on a circle of radius R. Formally, given a pixel at (i,j), the resulting LBP code can be expressed in the decimal form as follows:

$$LBP_{P,R}(i,j) = \sum_{n=0}^{P-1} s(I_n - I_c)2^n \tag{10}$$

where n runs over the P neighbors of the central pixel, I_c and I_n are the gray level values of the central pixel and the neighbor pixel, and $s(x)$ is 1 if $x \geq 0$ and 0 otherwise. After labeling an image with LBP operator, histogram of the labeled image $I_l(i,j)$ can be defined as:

$$H_c = \sum_{i,j} F(I_l(i,j) = c) \tag{11}$$

where $c = 0, 1, \ldots, L$ and L is the number of different labels produced by the LBP operator and

$$F(A) = \begin{cases} 1, & \text{if A is true,} \\ 0, & \text{if A is false.} \end{cases} \tag{12}$$

The derived LBP histogram contains information about the distribution of local features, such as edges, spots and flat areas over the image. It can be used to statistically describe image characteristics. In this paper, P and R are set to 24 and 3, respectively.

2.7. **Wavelet feature.** Wavelet transform has been one of the most widely used and powerful texture feature extraction methods in medical images classification application [11,17]. In this paper, in order to extract wavelet features, energy of LH and HL sub bands of one level Symlet 4 wavelet coefficients are computed.

3. **Proposed Novel Shape-texture Feature.** In this section, the NSTF is comprehensively introduced. In order to overcome the weakness of existing traditional feature extraction methods, we proposed a novel feature which is the combination of shape and texture features. As it is shown in Figure 2, in the first step of proposed feature extraction method, the histogram of all images are adjusted. Then, after removing noise from adjusted images, the edge and shape information are extracted. Afterwards, the shape images are splitted to 25 sub images and the Gabor features are extracted from these sub images. In the following subsections, the fundamental steps of proposed shape-texture feature are given.

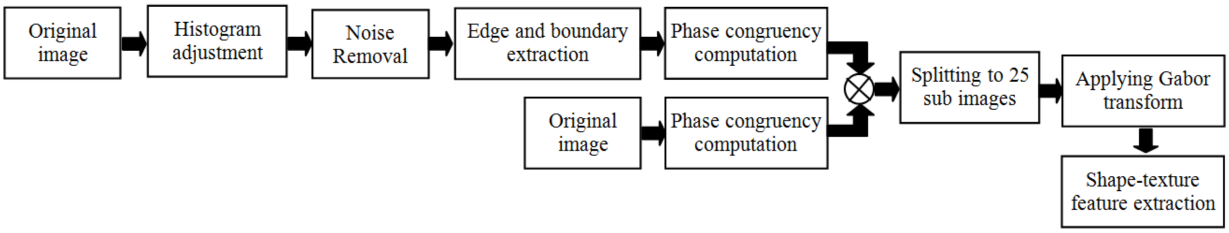

FIGURE 2. Fundamental steps of the proposed NSTF extraction framework

3.1. **Histogram adjustment.** In order to emphasize and highlight the details of input images and to justify their gray level histogram, in the first step of proposed feature extraction method, the histogram of input images is adjusted. In this case, the contrast of all images is increased by mapping the values of the input intensity image to new values such that, 1% of the data is saturated at low and high intensities of the input data.

3.2. **Noise removal.** To get rid of noise and unnecessary information, the anisotropic diffusion filter as proposed by P. Perona and J. Malik [19] is utilized. The filter removes noise from input image while preserves important parts of image contents such as edges and major boundaries. The filtered image is modeled as the solution to the anisotropic diffusion equation as follows:

$$\frac{\partial u(x,y,t)}{\partial t} = \mathrm{div}(g(\mid \nabla u(x,y,t) \mid)\nabla u(x,y,t)) \tag{13}$$

where $u(x,y,t) : \Omega \times [0,+\infty) \to R$ is a scale image and $g(\mid \nabla u \mid)$ is a decreasing function depending on the gradient of u.

3.3. **Edge and boundary extraction.** In order to extract the edges and boundaries from images, Canny edge detection algorithm [20] is utilized. Canny edge detection operator finds optimum edge detection algorithm which includes the following steps [20]: smoothing, finding gradients, non-maximum suppression, double thresholding and edge tracking by hysteresis. In this task, Canny edge detection is chosen between 6 different edge and boundary detection algorithms, i.e., Sobel, Prewitt, Roberts, Laplacian of Gaussian (LOG) and Zero-crossing, [21-23] methods, because of its considerable capability and acceptable results.

3.4. **Phase congruency computation.** As shown in Figure 2, to remove unnecessary and redundant edges and boundaries and also to increase the edge features accuracy, the phase congruency of shape image and original image are computed. Then, these two resultant images are multiplied. Phase congruency is a measure of feature importance in images; it is a method of boundary detection that is especially robust against changes in illumination and contrast. The measure of phase congruency developed by Morrone et al. [24] is:

$$PC(x) = \frac{\mid E(x) \mid}{\sum_n A_n(x)} \tag{14}$$

where $\mid E(x) \mid$ (local energy), is the magnitude of the vector from the origin to the end point and $A_n(x)$ is the amplitude of local, complex valued, Fourier components at a location x in the signal. Figure 3 shows the effect of applying phase congruency to the original image and the shape image of Step 3 and their noticeable multiplication result. As it is illustrated in this figure, most of the unnecessary and redundant edges and boundaries are entirely omitted or taken low intensity.

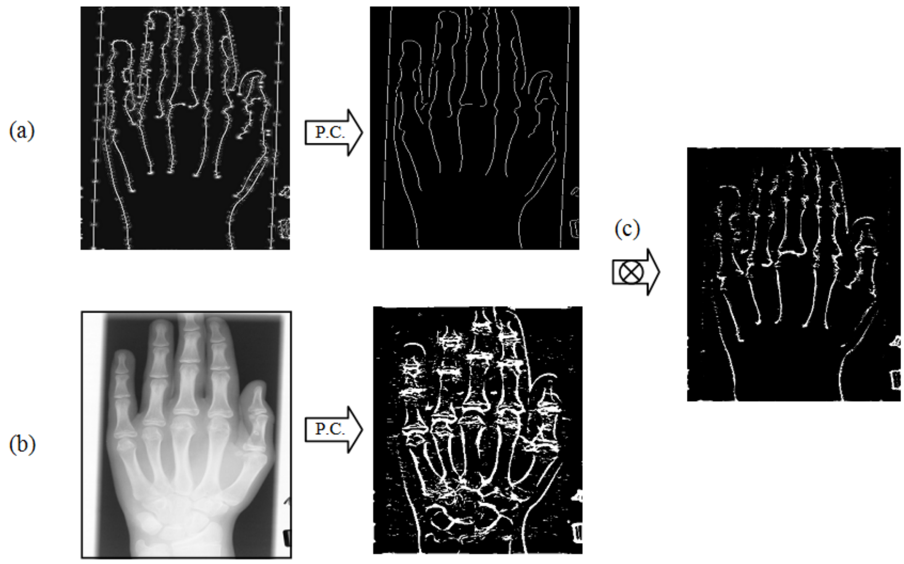

FIGURE 3. Phase congruency (P.C.) computation results: (a) computing phase congruency of edge images; (b) computing phase congruency of original images; (c) multiplication results of shape images phase congruency and original images phase congruency

3.5. **Splitting to 25 sub images.** The medical x-ray images have complex edge and directional information variations which are entirely dependent to spatial location in the image. Therefore, considering total shape of an object in each image as shape feature may not be a good strategy. In order to have an exact and effective feature, the processed image of previous subsection is splitted to several sub images. In Figure 4, image splitting circumstance to 25 sub images is illustrated. Also, in order to avoid omitting the available information on the sub images boundaries, selected sub images have overlap.

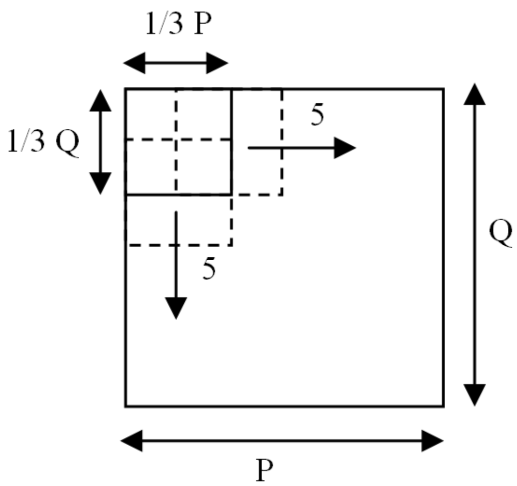

FIGURE 4. Splitting image to 25 sub images

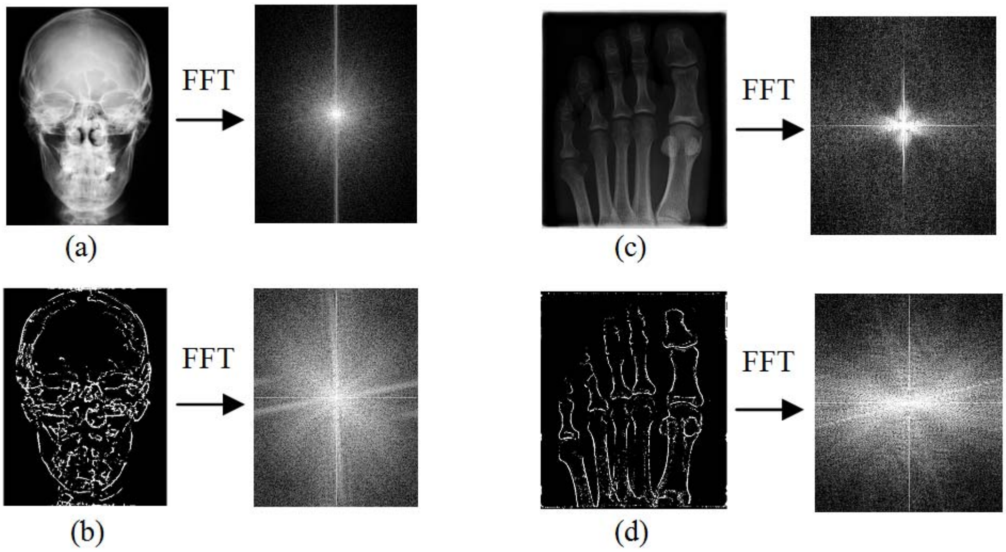

FIGURE 5. Fourier spectrum differences: (a), (c) original images Fourier spectrum; (b), (d) shape images Fourier spectrum

3.6. **Applying Gabor transform.** The spectral characteristics of x-ray images are considerably located in low frequencies. However, Gabor filtering extracts features which belong to mid frequency bands or higher. Hence, feature extraction from Gabor filtering cannot give distinguishable features. As illustrated in Figure 5, edge extraction of any image causes the spectrum to become more spread. In this case, Gabor filtering will be more efficient and produce more effective features. For a given image $I(i, j)$ with size $P \times Q$, the discrete Gabor transform [25] is given by a convolution:

$$G_{mn}(i,j) = \sum_{s}\sum_{t} I(i-s, j-t)\phi^*_{mn}(s,t) \tag{15}$$

where s and t are the filter mask size variables, and ϕ^*_{mn} is the complex conjugate of ϕ_{mn} which is a class of self-similar functions generated from expansion and rotation of the following mother wavelet:

$$\phi_{mn} = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi\omega x) \tag{16}$$

where ω is the modulation frequency. The self-similar Gabor wavelets are obtained through the generating function:

$$\phi_{mn}(x,y) = a^{-m}\phi(\tilde{x},\tilde{y}) \tag{17}$$

where m and n indicate the scale (frequency) and direction of the wavelet respectively, with $m = 0, 1, \ldots M-1$, $n = 0, 1, \ldots, N-1$, and

$$\begin{aligned} \tilde{x} &= a^{-m}(x\cos\theta + y\sin\theta) \\ \tilde{y} &= a^{-m}(-x\sin\theta + y\cos\theta) \end{aligned} \tag{18}$$

where $a > 1$ and $\theta = n\pi/N$. The variables in the previous equations are defined as follows:

$$\begin{aligned}
a &= (U_h/U_l)^{\frac{1}{M-1}} \\
W_{m,n} &= a^m U_l \\
\sigma_{x,m,n} &= \frac{(a+1)\sqrt{2\ln 2}}{2\pi a^m (a-1) U_l} \\
\sigma_{y,m,n} &= \frac{1}{2\pi \tan\left(\frac{\pi}{2n}\right)\sqrt{\frac{U_h^2}{2\ln 2} - \left(\frac{1}{2\pi\sigma_{x,m,n}}\right)^2}}
\end{aligned} \tag{19}$$

Here, $U_l = 0.05$, $U_h = 0.4$, $M = 3$, $N = 6$, s and t range from 0 to 33. Please note that, the Gabor filter bank is applied to each of 25 sub images of shape images.

3.7. **Shape-texture feature extraction.** In the final step, features are extracted from filtered sub images. In this paper, Diagonal moment, Entropy, Sure Entropy (8), Log energy Entropy (8), Norm Entropy and Standard Deviation [15,16], (20), are chosen as shape-texture features in order to extract the final information from Gabor transform.

$$\begin{aligned}
\text{Diagonal moment} &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sqrt{\frac{\mid i-j \mid I'(i,j)}{2}} \\
\text{Entropy} &= -\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} I'(i,j)\log(I'(i,j)) \\
\text{Norm Entropy} &= -\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} \mid I'(i,j) \mid^{\tau} \\
\text{Standard deviation} &= \sqrt{\frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}(I'(i,j)-m)^2}{M\times N-1}}
\end{aligned} \tag{20}$$

where $I'(i,j)$, ($i = 0,\dots,M-1$ and $j = 0,\dots N-1$), is a pixel value in a sub image and τ is a constant where $\tau \geq 1$ and m is the mean values of I'.

4. **Classification Methods.** In order to evaluate capability of NSTF and to compare it with existing features for content based medical images retrieval (CBMIR), three different classifiers including Euclidean distance, probabilistic neural network (PNN) and support vector machine (SVM) are utilized.

4.1. **Euclidean distance.** An image feature vector with fixed size L can be shown by $x = \{x^1, x^2, \dots, x^L\}$. The Euclidean distance $d_E(x_1, x_2)$ [26] between feature vectors x_1 and x_2 is defined as:

$$d_E^2(x_1, x_2) = \sum_{k=1}^{L}\left(x_1^k - x_2^k\right)^2 = (x_1 - x_2)^T(x_1 - x_2) \tag{21}$$

Here, x_1 is the feature vector of a test image and x_2 is the average of feature vectors of training samples for each class. T in this case denotes the transpose of the term $(x_1 - x_2)$.

4.2. **Probabilistic neural network.** The probabilistic approach to neural networks has been developed in the framework of statistical pattern recognition [27]. The design of probabilistic neural network (PNN) is based on approximating the class-conditional probability distributions by finite mixtures which can be used to make Bayesian decision. The mixture parameters can be computed by means of expectation maximization algorithm. The PNN network, as described in Figure 6, includes an input layer, two hidden layers (one each for example/pattern and a class/summation layers) and an output layer.

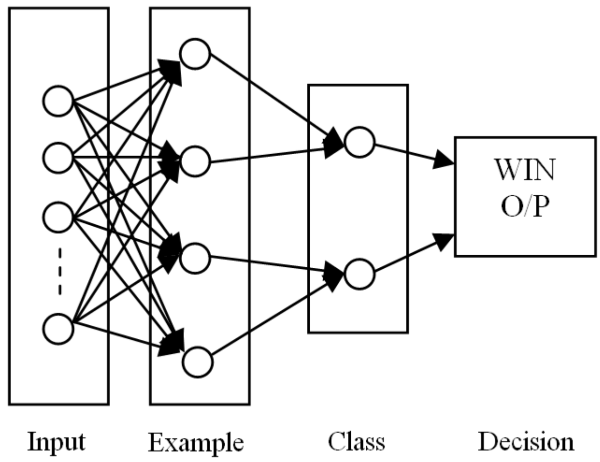

Figure 6. Structure of PNN

4.3. **Support vector machine.** In this paper, the one-against-one Multiclass Support Vector Machine (SVM) training framework for image classification is utilized. The SVM have been developed by Vapnik [4,28] which gained popularity because of many promising features such as better empirical performance. SVM is a set of classification and regression related supervised learning methods and belongs to a family of generalized linear classifiers. On the other hand, it is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy. SVM can be defined as linear classifier in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory.

4.4. **Support vector machine.** In this paper, the one-against-one Multiclass Support Vector Machine (SVM) training framework for image classification is utilized. The SVM have been developed by Vapnik [4,28] which gained popularity because of many promising features such as better empirical performance. SVM is a set of classification and regression related supervised learning methods and belongs to a family of generalized linear classifiers. On the other hand, it is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy. SVM can be defined as linear classifier in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory.

SVM is originally designed for binary pattern classification. For multi-class pattern recognition, a combination of binary SVMs and a decision strategy to decide the class of the input test image are commonly used. Each SVM is independently trained. The training data set (x_w, c_w) consists of L examples belonging to K classes. $c_w \in 1, 2, \ldots, K$

are the classes labels. The one-against-one method was first introduced on SVM as pairwise SVM. In this approach, a SVM is constructed for each pair of classes. Thus, the number of SVMs used in this approach is $K(K-1)/2$. A SVM for a pair of classes (k,l) is constructed using training examples belonging to these two classes. The desired output, y_w, for a training example, x_w, is defined as follows:

$$y_i = \begin{cases} +1, & \text{if } c_w = k, \\ -1, & \text{if } c_w = l. \end{cases} \tag{22}$$

5. **Experimental Results.** In this section, the former features and the NSTF will be evaluated by three known and widely used classifiers. First, the utilized database will be introduced and then the feature extraction details and the features evaluation will be thoroughly explained.

5.1. **Utilized database.** The utilized images in this paper are chosen from an IRMA database [29], containing 9100 medical x-ray images in 57 classes. In our task, to evaluate the proposed novel shape-texture feature and to compare it with other features, 21 classes out of these 57 classes are selected. These classes are introduced in Table 1. Note that they may have different scales with different contrasts even in one class. In addition, they may be corrupted, rotated or contain signs such as L or R. Some examples of images with these conditions are shown in Figure 7. With regard to these conditions, proposed features are extracted from them to make a comprehensive result.

TABLE 1. The 21 utilized classes of IRMA database medical x-ray images

Class Number	Anatomic Part	Direction
1	Pelvis (cardiovascular system)	Coronal
2	Left Breast	Axial
3	Right Breast	Axial
4	Knee	Axial
5	Chest	Coronal
6	Elbow	Sagittal
7	Elbow	Coronal
8	Ankle joint	Coronal
9	Knee	Coronal
10	Pelvis (musculoskeletal system)	Coronal
11	Cervical Spine	Coronal
12	Lumber Spine	Coronal
13	Chest	Sagittal
14	Neuro Cranium	Sagittal
15	Facial Cranium	Coronal
16	Cranium	Coronal
17	Cervical Spine	Sagittal
18	Lumber Spine	Sagittal
19	Thoracic Spine	Sagittal
20	Lower Leg (cardiovascular system)	Coronal
21	Ankle joint	Sagittal

FIGURE 7. Examples of some images in utilized database

5.2. **Feature evaluation.** To evaluate and compare the features, they are all applied to 4402 medical x-ray images of 21 classes, shown in Table 1, containing 315 training images and 4087 test ones (15 training images and 20 to 2000 test images for each class). Note that the training and test images of each class are not the same. The other important point is that the number of test images is much more than that of training ones. The GLCM, moment invariants, region properties, LBP and Tamura features are extracted from whole original images and the Entropy feature is extracted from whole normalized images as they are introduced in Section 2. In this paper, in order to improve the capability of wavelet feature, instead of extracting it from original images, it is extracted from shape images. The classification result of extracting wavelet feature from shape images was improved by near 26% in comparison with classification results of extracting it from original images. As discussed in Subsection 3.6, extracting spectral texture features from shape images will improve the capability of those features; therefore, improvement of the reformed wavelet feature classification results has been predictable. Hence, in this paper, the reformed wavelet feature is selected to be compared with NSTF. Finally, the NSTF is extracted from medical x-ray images as it was introduced in Section 3. In this case, the scales and gray levels of medical x-ray images do not play great role in classification result. On the other hand, the NSTF is independent to scales and gray levels of the medical x-ray images. After feature extraction, in order to reduce the feature space dimensionality and computational complexity and also to increase the features capability, the principal component analysis (PCA) algorithm is utilized [30]. PCA is theoretically the optimal linear scheme, in terms of least mean square error, for compressing a set of high dimensional vectors into a set of lower dimensional vectors and then reconstructing the original set. The lenght of moment invariants, region properties, Entropy and Tamura feature vectors become 7, 7, 3 and 4 after feature extraction, respectively; hence, reducing these short feature vectors by PCA algorithm is not necessary. The remaining feature vector lengths before and after reducing by PCA are brought in Table 2. The shortened features are evaluated by means of three common classifiers, consequently. In Table 3, the classification results of these features including the NSTF are brought. As it is clearly apparent in Table 3, the NSTF has the most accurate classification result amongst others. In the next subsection, the combination of features is studied.

5.3. **Feature combination.** Considering the classification results shown in Table 3, the idea of the feature combination is studied. In this case, different combinations of features are considered. As the classification accuracy rate of NSTF is much more than others, the NSTF is decided to be the permanent part of the combination. First, dual combination of NSTF and each remaining features is evaluated. Note that these features are exactly the same as features evaluated in the previous subsection. In order to evaluate this combination, after feature extraction from 4402 images, first the feature vector lengths are reduced by PCA, separately. The reduced feature vector lengths are shown in Table

Table 2. Feature vector length before and after reducing by PCA for classification

Features	Length before reduction	Reduced length by PCA
Novel shape texture feature (NSTF)	2700	240
GLCM	52	30
Reformed wavelet feature	50	25
LBP	78	20

Table 3. The classification accuracy rates of 8 features for 4402 medical x-ray images in 21 classes obtained by SVM, PNN and Euclidean distance

Features	SVM	PNN	Euclidean Distance
GLCM	65.4%	35.21%	53.02%
Moment Invariants	4.76%	5%	16.93%
Region properties	41.04%	22.74%	17.56%
Tamura features	27.02%	33.65%	28.96%
Entropy	25.23%	24.62%	21.44%
LBP	56.9%	28.66%	46.09%
Reformed wavelet feature	41.9%	15.76%	34.86%
The novel shape-texture feature	88.77%	64.01%	83.55%

2. After feature vector length reduction, the two feature vectors are joined to each other and make a unique feature vector. Now, to evaluate the feature combinations, three different classifiers are utilized. The classification accuracy rates of dual combination are illustrated in Table 4. The results shown in Table 4 indicate that the combination of the NSTF and GLCM improves the total classification result. Therefore, in order to study the other combination conditions, triple combination of the NSTF plus GLCM and other 6 remaining features are evaluated. The triple feature combination classification results are shown in Table 5. The results indicate that the combination of three features, i.e., NSTF, GLCM and LBP improves the dual combination classification result by almost 4%. The obtained classification result of combining only three features for 4402 medical x-ray images is really marvelous. In order to further study, quad combination of features is evaluated. The evaluation results of quad combination of the NSTF, GLCM, LBP and 5 other remaining features by three classifiers are shown in Table 6. As the results show, further combination of features can not improve the classification accuracy rates, i.e.,

Table 4. The classification accuracy rates of dual combination of the novel shape-texture feature and other 7 features for 4402 medical x-ray images in 21 classes obtained by SVM, PNN and Euclidean distance

Features	SVM	PNN	Euclidean Distance
NSTF+ GLCM	90.89%	81.1%	89.52%
NSTF+ Moment Invariants	88.77%	64.01%	83.55%
NSTF+ Region properties	55.35%	22.74%	17.56%
NSTF+ Tamura features	59.69%	34.42%	28.96%
NSTF+ Entropy	86.68%	75.86%	65.51%
NSTF+ Reformed wavelet feature	88.53%	63.62%	84.01%
NSTF+ LBP	89.9%	64.43%	84.55%

Table 5. The classification accuracy rates of triple combination of the novel shape-texture feature plus GLCM and other 6 features for 4402 medical x-ray images in 21 classes obtained by SVM, PNN and Euclidean distance

Features	**SVM**	**PNN**	**Euclidean Distance**
NSTF+ GLCM + Moment Invariants	90.89%	81.1%	89.52%
NSTF+ GLCM + Region properties	55.54%	22.74%	17.56%
NSTF+ GLCM + Tamura features	60.1%	34.42%	28.96%
NSTF+ GLCM + Entropy	89.73%	79.58%	68.93%
NSTF+ GLCM + Reformed wavelet feature	90.66%	77.37%	89.32%
NSTF+ GLCM + LBP	94.2%	78.58%	90.36%

Table 6. The classification accuracy rates of quad combination of the novel shape-texture feature, GLCM, LBP and other 5 features for 4402 medical x-ray images in 21 classes obtained by SVM, PNN and Euclidean distance

Features	**SVM**	**PNN**	**Euclidean Distance**
NSTF+ GLCM + LBP + Moment Invariants	94.2%	78.58%	90.36%
NSTF+ GLCM + LBP + Region properties	55.85%	22.74%	17.56%
NSTF+ GLCM + LBP + Tamura features	59.5%	34.43%	28.96%
NSTF+ GLCM + LBP + Entropy	89.94%	79.2%	71.28%
NSTF+ GLCM + LBP + Reformed wavelet feature	91.53%	76.95%	90.61%

adding the fourth feature does not provide additional discriminatory information suitable for classification improvement. Therefore, utilizing the combination of three features, i.e., NSTF, GLCM and LBP as classification criterions is the best choice to classify the medical x-ray images. Classification accuracy of 94.2% obtained by SVM, for combination of three features confirms the capability of them in medical x-ray images classification. In addition, the selected features are general for this variety of medical x-ray images, because they are not needed to be assigned special weights for each class. Classification accuracy rates of 21 classes for combination of these three features are separately shown in Table 7.

5.4. **Computational complexity order of shape-texture feature.** As the NSTF extraction includes several steps, in order to compute its computational complexity, we have to analyze each step separately and then compute the total complexity.

- The anisotropic diffusion process computational complexity order is $(P \times Q)N_{it}$, where N_{it} is the number of iteration in the anisotropic diffusion process and $P \times Q$ is the size of the image.
- The Canny edge detector is a projection operator G_n that when convolved with a 2D image I, the zero-crossings of the resulting function $G_n * I$ are the edges of the image. Hence, the computational complexity order of applying Canny edge detector is $(P \times Q)(k \times l)$, where $P \times Q$ is the image size and $k \times l$ is size of the projection operator.
- The phase congruency computational complexity order for an image with size $P \times Q$ is $(P \times Q)$.
- The 2D Gabor kernel can be represented as a convolution of two orthogonal 1D components. These components are: a Gaussian $g(x)$ and a Wavelet $w(x)$ (a complex

TABLE 7. Evaluating the combination of novel shape texture feature, GLCM and LBP for 4402 medical x-ray images in 21 classes by SVM, Euclidean distance and PNN classifiers

Class Number	SVM	Euclidean Distance	PNN
1	100%	100%	90%
2	95%	95%	95%
3	95%	100%	90%
4	95%	95%	85%
5	88.22%	83.38%	75.26%
6	100%	95%	65%
7	90%	80%	40%
8	95%	85%	80%
9	95%	75%	70%
10	90%	75%	65%
11	95%	85%	90%
12	100%	100%	100%
13	91.59%	93.64%	40.23%
14	88.38%	80.63%	64.79%
15	100%	100%	70%
16	95%	100%	100%
17	85%	75%	75%
18	90 %	90%	90%
19	100%	95%	75%
20	95%	100%	100%
21	95%	95%	90%
Total Accuracy Rates	**94.2%**	**90.36%**	**78.58%**

wave enveloped by a Gaussian), defined respectively by:

$$g(x) = \frac{1}{2\pi\sigma} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2}\right)\right] \tag{23}$$

$$w(x) = g(x)\exp(j2\pi\omega x) \tag{24}$$

where ω is the frequency of the wavelet. These functions describe the separable components of a Gabor filter kernel. It follows that convolution of a Gabor kernel with an image can be calculated separably. For example, a horizontally aligned $n \times n$ Gabor kernel K can be written as:

$$\mathbf{K} = \mathbf{g} * \mathbf{w} \tag{25}$$

where $\mathbf{g}$ and $\mathbf{w}$ are $n \times 1$ vectors whose elements are define by regularly sampling $g(x)$ and $w(x)$ across intervals centered at $x = 0$. The convolution of $\mathbf{K}$ with an image $\mathbf{I}$ is then:

$$\mathbf{I} * \mathbf{K} = \mathbf{I} * (\mathbf{g} * \mathbf{w}) = (\mathbf{I} * \mathbf{g}) * \mathbf{w} \tag{26}$$

The 2D convolution computation is $O((P \times Q)n^2)$ for an $n \times n$ kernel and an image with $P \times Q$ pixels.

To implement this research the utilized system was Pentium 4 with 3 GHz CPU and 1 GB RAM and all programs have been implemented in MATLAB software.

6. **Conclusion.** In this paper, a new feature extraction framework for medical x-ray images is proposed. Medical x-ray images have unique characteristics so the common features may not be proper classification criterions for this type of images. With regard to medical x-ray images characteristics, an effective and powerful combination of shape and texture features, called novel shape-texture feature (NSTF) is introduced. The results brought in experimental results section, express the efficiency of this novel feature. We can enumerate the advantages of proposed framework as following points:

a) The shape texture feature in comparison with other widely used features in medical images classification application showed a marvelous capability;
b) Combination of only three features for a large group and variety of medical x-ray images without specifying a weight to each feature, expresses the generality and capability of them;
c) Utilizing few training samples and examining large number of test ones proves the great capability and efficiency of proposed features;
d) The low computational complexity and straight forward implementation.

With regard to these advantages we can strongly claim that the proposed feature in this paper is the most powerful and reliable feature for medical x-ray images classification. Finally, to study the feature combination effect on classification results, two other effective features are examined and added to the proposed feature which resulted magnificent classification accuracy rates. The utilized data base has many deficiencies like severe shift or rotation. The proposed shape-texture feature is dependent to shapes and directional information; hence, if the objects in the images are extremely shifted or rotated, using novel shape texture feature may result in misclassifications. Of course, this deficiency is not common for x-ray imaging systems.

Acknowledgment. The authors would like to thank TM Lehmann, Dept. of Medical Informatics, RWTH Aachen, Germany, for making the database available for the experiments.

REFERENCES

[1] A. Ohno and H. Murao, A similarity measuring method for images based on the feature extraction algorithm using reference vectors, *International Journal of Innovative Computing, Information and Control*, vol.5, no.3, pp.763-772, 2009.
[2] S. Miyata, A. Yanou, H. Nakamura and S. Takehara, Road sign feature extraction and recognition using dynamic image processing, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(B), pp.4105-4114, 2009.
[3] M. Pan, J. Tang and X. Yang, An algorithm for medical image tilt correction using B-spline and moment invariants, *ICIC Express Letters*, vol.4, no.1, pp.57-64, 2010.
[4] A. Jaffar, B. Ahmed, N. Naveed, A. Hussain, F. Jabeen and A. M. Mirza, Multi domain features based classification of mammogram images using SVM and MLP, *ICIC Express Letters*, vol.4, no.3(B), pp.937-942, 2010.
[5] H. Pourghassem and H. Ghassemian, Content-based medical image classification using a new hierarchical merging scheme, *Computerized Medical Imaging & Graphics*, vol.32, no.8, pp.651-661, 2008.
[6] V. Jacquet, V. Jeanne and D. Unay, Automatic detection of body parts in x-ray images, *IEEE Computer Society Conf. on Computer Vision & Pattern Recognition Workshop*, Miami, FL, USA, pp.25-30, 2009.
[7] Bertalya, Prihandoko, D. Kerami and T. M. Kusuma, Classification of x-ray images using grid approach, *IEEE International Conf. on Signal Image Technology & Internet Based Systems*, Washington, DC, USA, pp.314-319, 2008.
[8] Q. Zhang and X. Tai, The x-ray chest image retrieval based on feature fusion, *International Conf. on Audio, Language & Image Processing*, Shanghai, China, pp.899-903, 2008.

[9] A. Mueen, M. S. Baba and R. Zainuddin, Multilevel feature extraction and x-ray image classification, *Journal of Applied Science*, vol.7, no.8, pp.1224-1229, 2007.

[10] R. M. Haralick, K. Shanmugam and I. Dinstein, Textural features for image classification, *IEEE Trans. Systems, Man, & Cybernetics*, vol.3, no.6, pp.610-621, 1973.

[11] M. Vasantha, D. V. S. Bharathi and R. Dhamodharan, Medical image feature, extraction, selection and classification, *International Journal of Engineering Science & Technology*, vol.2, no.6, pp.2071-2076, 2010.

[12] K. Masood and N. Rajpoot, Spatial analysis for colon biopsy classification from hyperspectral imagery, *Annals of the BMVA*, vol.2008, no.4, pp.1-16, 2008.

[13] H. Muller, J. Kalpathy-Cramer, I. Eggel, S. Bedrick, J. Reisetter, C. E. Kahn Jr. and W. Hersh, Overview of the CLEF 2010 medical image retrieval track, *Proc. of the CLEF 2010 Workshop*, 2010.

[14] B. G. Prasad and A. N. Krishna, Performance evaluation of statistical texture features for medical image classification, *Proc. of the National Conference on Emerging Trends in Computing Science*, 2011.

[15] R. R. Coifman and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. on Information Theory*, vol.38, no.2, pp.713-718, 1992.

[16] D. L. Donoho and I. M. Johnstone, Ideal de-noising in an orthonormal basis chosen from a library of bases, *C.R.A.S. Paris, t. 319, Ser. I*, pp.1317-1322, 1994.

[17] C. Ray and K. Sasmal, ECG image classification and detection of trend of heart diseases, *Proc. of International Conf. on Advances in Computer Science*, 2010.

[18] M. Eisa, A. ElGamal, R. Ghoneim and A. Bahey, Local binary patterns as texture descriptors for user attitude recognition, *International Journal of Computer Science & Network Security*, vol.10, no.6, pp.222-229, 2010.

[19] S. Fu, Q. Ruan, W. Wang and Y. Li, Adaptive anisotropic diffusion for ultrasonic image denoising and edge enhancement, *International Journal of Information Technology*, vol.2, no.4, pp.284-292, 2006.

[20] J. Canny, A computational approach to edge detection, *IEEE Trans. on Pattern Analysis & Machine Intelligence*, vol.8, no.6, pp.679-698, 1986.

[21] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, Prentice Hall, NJ, USA, 1990.

[22] J. R. Parker, *Algorithms for Image Processing and Computer Vision*, John Wiley & Sons, New York, USA, 1997.

[23] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd Edition, Prentice-Hall, Upper Saddle River, New Jersey, USA, 2008.

[24] M. C. Morrone, J. R. Ross, D. C. Burr and R. A. Owens, Mach bands are phase de-pendent, *Nature*, vol.324, pp.250-253, 1986.

[25] S. A. Veni and K. A. Narayanankutty, Image enhancement of medical images using Gabor filter bank on hexagonal sampled grids, *World Academy of Science, Engineering and Technology*, vol.65, 2010.

[26] E. Deza and M. M. Deza, *Encyclopedia of Distances*, Springer, 2009.

[27] J. Grim, P. Somol and P. Pudil, Probabilistic neural network playing and learning Tic-Tac-Toe, *Pattern Recognition Letters, Special Issue: Artificial Neural Networks in Pattern Recognition*, vol.26, no.12, 2005.

[28] C. Burges, A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, vol.2, no.2, pp.121-167, 1998.

[29] *http://irma-project.org.*

[30] A. K. Jain and B. Bhandrasekaran, Dimensionality and sample size considerations in pattern recognition practice, *Handbook of Statistics*, pp.835-855, 1987.