Wavelet Transform Analysis and Classification of the Liver from Computed Tomography Datasets

Tryphon Lambrou, *MIEEE*, Alfred D. Linney, and Andrew Todd-Pokropek, *MIEEE*

***Abstract*— In this paper a feasibility study of liver CT dataset classification, using features from different scales of the wavelet transform analysis in conjunction with statistical pattern recognition methods is presented. In our study 850 extracted sub-images from 19 liver CT scans were used, in order to establish which features distinguish better between the normal/cancer classes. Statistical measurements were collected; from the sub-images as well as from their different scale wavelet transform coefficients. We found by using the Leave-One-Out method that the combination of the features from the 1st and 2nd Order statistics, achieved overall classification accuracy > 90.0%, both specificity and sensitivity > 90.0%. Features selected by the spatial domain performed better than the wavelet based techniques, under the classification rule of Quadratic Classifier (QC). In addition, features selected by the 3rd scale wavelet transform coefficients performed better than those collected from the other wavelet scales, under the classification rule of Bayesian Classifier (BC).**

I. INTRODUCTION

WAVELET theory is a mathematical tool for hierarchically decomposing functions. Wavelet transform analysis has been applied to medical images mainly for compression, and mammographic image analysis [1].

Although Computed Tomography (CT) is only slightly more accurate than ultrasound in showing focal hepatic lesions, it has several advantages. All the upper abdominal anatomy is displayed on the CT images, providing information about extrahepatic processes that can influence clinical interpretation. Also, intravenous injection of water-soluble contrast medium increases the detection rate of small masses. In addition it should be noted that two major improvements in CT technology have established the modality as the preferred single imaging technique for screening the abdomen. Firstly, the improvement of scan speed using spiral/helical CT technology allowed the creation of multiphasic contrast enhanced CT examination of the whole liver. Secondly, the introduction of multidetector row CT (MDCT) scanners which are replacing the spiral/helical CT systems, and offer high spatial resolution isotropic imaging with high temporal resolution, and enable scan duration throughout whole liver of around 10s. However, as our bibliographic review shows, until very recently there has been little published research focused on liver CT.

Yoshino et.al. [2][3], developed an image diagnosis system that had a three-layer neural network back-propagation utilizing the back-propagation algorithm. Yoshido and co-workers classified parenchymal patterns of cirrhotic liver into three types according to the size of nodules, using magnetic resonance images and ultrasound datasets.

Chen et.al. [4], presented a CT liver image diagnostic classification system which automatically finds/extracts the CT liver boundary and further classifies liver disease. Their reported system comprises a detect-before-extract Brownian motion model to delineate the liver boundary, and a modified probabilistic neural network to distinguish between normal liver and hepatoma and hemageoma. The reported classification accuracy was about 83%.

Lee et.al. [5], proposed a method for diffuse liver disease classification of ultrasound liver datasets, using multiscale wavelet based analysis and a probabilistic neural networks. Their dataset included, normal liver, hepatitis and cirrhosis, and achieved classification accuracy rate of around 88%.

Lee et.al. [6] used features based on M-band wavelet transform to classify ultrasonic liver images – normal liver, cirrhosis, and hepatoma. Their proposed hierarchical classifier achieved 96.7% accuracy in the distinction between normal – abnormal liver images, and was at least 93.6% accurate in the distinction between cirrhosis and hepatoma liver images.

Manuscript received June 30, 2006.

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) and Medical Research Council (MRC) under the Interdisciplinary Research Consortium scheme – "From Medical Images and Signals to Clinical Information".

T. Lambrou is a Research Fellow with the Department of Medical Physics & Bioengineering, University College London, Gower Street, London WC1E 6BT, UK (phone: +44-020-7679-0286; fax: +44-020-7679 -0255; e-mail: tlambrou@medphys.ucl.ac.uk).

A. D. Linney is a Professor with the Centre of Auditory Research, UCL Ear Institute, 332-336 Gray's Inn Road, London WC1X 8EE, UK (e-mail: a.linney@ucl.ac.uk).

A. Todd-Pokropek is a Professor and Head of the Department of Medical Physics & Bioengineering, University College London, Gower Street, London WC1E 6BT, UK (e-mail: atoddpok@medphys.ucl.ac.uk).

Yoshida et.al. [7] addressed the problem of distinguishing benign (hemangiomas) from malignant (hepatocellular carcinomas (HCCs) and metastases) focal liver lesions in B-mode ultrasound images. Multiscale texture features from the wavelet packet analysis were combined by an artificial neural network; the performance was measured by the area under the curve (Az). Their reported results yielded a Az value of 0.92 in distinguishing benign from malignant lesions, 0.93 in distinguishing hemangiomas from HCCs, and 0.94 in distinguishing hemangiomas from metastases.

Gletsos et.al. [8], presented a computer-aided diagnostic system for classifying hepatic lesions from computed tomography images. CT images of normal liver, hepatic cysts, hemangiomas, and hepatocellular carcinomas were used as input. Texture characteristics from the co-occurrence matrices were collected, and their classification scheme consisted of three sequentially placed feed-forward neural networks.

This paper attempts an investigation on the usage of statistical features collected from the spatial and wavelet transform domains, using several different classifiers, for applications on Liver CT image classification and retrieval.

II. MATERIALS & METHODS

In this study we used 850, 32x32x8 bit, image extracts from 19 Liver CT scans (425 normal and 425 cancer), for the training stage of the classification procedure. The images were analyzed in Spatial domain, and using the first three levels of decomposition of the overcomplete wavelet transform [9][10] architecture. The application of the overcomplete wavelet decomposition, in which the output of the filter banks are not subsampled, should result in a texture description scheme invariant with respect to translations of the input signal. This is not the case with other wavelet based approaches or decomposition schemes. The Daubechies 4-TAP wavelet filter was used.

Our statistical pattern recognition approach uses the classical steps of feature extraction, classification and feature selection, which are further described below.

The first step of our pattern recognition approach is the feature extraction step, which is the transformation of patterns into features that are regarded as a compacted representation. The usage of statistical features for the analysis and classification of textured images has been extensively demonstrated in the literature. Overall twenty-two statistical image features were collected from each image, given by category as:

First Order Statistics [11] which express the distribution of grey levels within the image. These features are based on the histogram of the image, since they represent the frequency distribution of the grey level in the image, i.e. Mean, Variance, Skewness, and Kurtosis.

Second Order Statistics [11] which are measurements collected by the Grey-Tone Spatial-Dependent Matrices. These matrices of grey-tone spatial-dependence frequencies are a function of the angular relationship between neighboring image elements, and additionally a function of the distance between them. The features which can be extracted by the grey-tone spatial-dependence matrices are: Angular Second Moment, Correlation, Entropy, Sum of Squares: Variance, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance and Difference Entropy.

Grey Level Run Lengths [11] which are measurements on matrices which represent the set of the consecutive image elements which have the same grey level (grey level run), and the number of image elements which belong to such a grey level run. The following features can be extracted from these matrices: Short Runs Emphasis, Long Runs Emphasis, Gray Level Non-Uniformity, Run Lengths Non-Uniformity, and Run Percentage.

In addition, from the wavelet decomposed images the features collected were, their First Order Statistics: i.e. Mean, Variance, Skewness, and Kurtosis. The measures of Root Mean Square (RMS) Variation, the Non-Normalized Energy, the Normalized Energy, the Normalized Shannon Entropy, and the Non-Normalized Shannon Entropy.

Three statistical classifiers were constructed and employed in this study. The classifiers used are:

Minimum Distance Classifier (MDC) [12], which employs as classification criterion the minimum Euclidean distance between the unknown entry and the mean values of each of the other classes.

Quadratic Minimum Distance Classifier (QC) [12], where the classification rule is again the minimum Euclidean distance between the unknown entry and the mean values of each of the other classes, using a quadratic equation within the least squares technique in order to minimize the errors.

Bayes Classifier (BC)[12], which minimizes the expected cost of misclassified data.

The performance of the classifiers was evaluated by using the Leave-One-Out method. This involves the re-classification of all the images (one at the time) to their a priori known categories (or classes). In addition, for each set of features all possible combinations were tested up to three-dimensional decision space. Those features, which achieve the best classification rate, were used in the pattern recognition process. This phase is called feature selection, and aims to reduce the features set to a subset, which consists only of meaningful information (i.e. features which characterize best) about the images we want to classify.

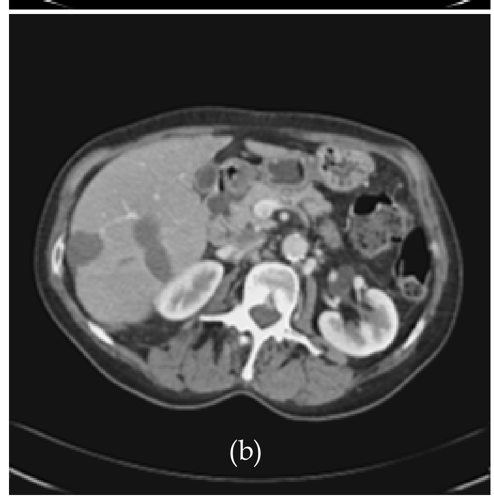
The classification accuracy results presented in this paper are those, which fulfill all of the three requirements: a) the classification accuracy of the normal class (specificity) is more than 80%, b) the classification accuracy of the abnormal class (sensitivity) is more than 80%, and c) the overall accuracy is more than 80%.

III. RESULTS

The wavelet transform analysis was performed using the overcomplete logarithmic splitting algorithm, and all the images were decomposed up to three levels of decomposition. The effect of such processing is demonstrated in Figure 1.

In the Spatial domain the best overall classification accuracy result achieved was 96.35% (specificity 93.88%, sensitivity 98.82%), using the feature combination Mean – Skewness –Entropy and Skewness – Sum Variance - Entropy from the 1st and 2nd Order statistics, and the Quadratic Classifier. In the 1st scale wavelet transform domain, the best overall classification accuracy was 90.47% (specificity 92.94%, sensitivity 88.00%), using the feature combination of Mean – Normalised Energy – Non Normalised Energy, and by the Quadratic Classifier. In the 2nd scale wavelet transform domain, the overall best classification accuracy was 92.35% (specificity 93.65%, sensitivity 91.06%), using features from the 2nd Order statistics Normalised Energy – Normalised Entropy, and by the Quadratic classifier. And finally from the 3rd scale wavelet transform domain, the best overall classification accuracy was 94.24% (specificity 95.06%, sensitivity 93.41%), using the 1st and 2nd Order statistics feature combination of Skewness – Normalised Shannon Entropy, and by the Quadratic classifier.

In terms of the performance of the Classifiers used in this study, we concluded that: the Quadratic Classifier performed better for features selected from the Spatial and the 1st, 2nd, and 3rd Scale Wavelet Transform Domains. The Minimum Distance Classifier performed well for features collected from the Spatial and the 1st Scale Wavelet Transform Domain. The Bayesian Classifier provided the comparable classification accuracy results to those obtained by the Minimum Distance Classifier, for features collected from all the domains. Except in the case of the 3rd Scale Wavelet Transform Domain, were it performed better. Tables 1 to 4, provide the best classification accuracy results of each of the classifiers for features collected from the Spatial and 1st, 2nd, and 3rd Scales of the Wavelet Transform Domains, respectively.



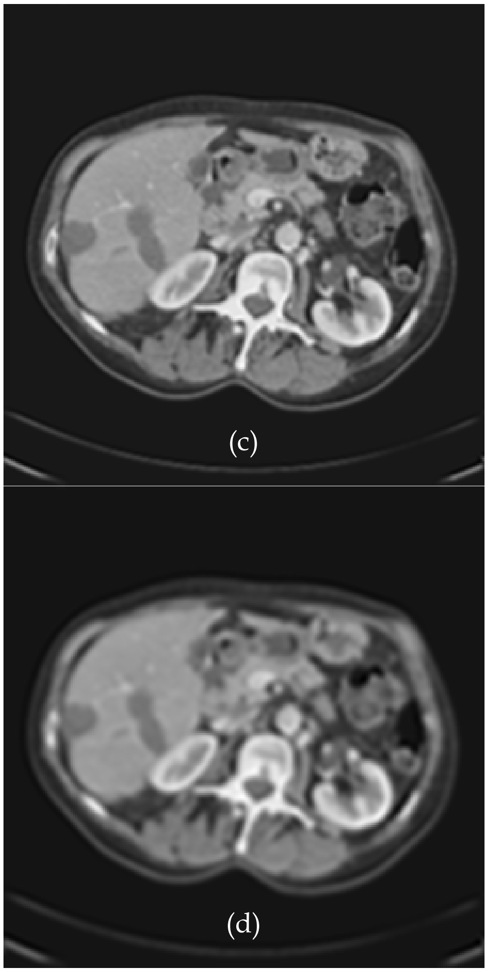


Figure 1. The effect of the overcomplete wavelet transform analysis, (a) Liver CT slice on the spatial domain, and (b)(c)(d) its 1st to 3rd Scale Wavelet Transform coefficients, respectively.

In terms of the performance of the Statistical Features extracted from all the liver CT images, we concluded that: Features from the 1st Order Statistics obtained by all Domains, produced classification accuracy results above the thresholds set. Features from the 2nd Order Statistics obtained by all the Domains produced the best classification accuracy results. Finally, features from the Grey Level Run Lengths obtained by the Spatial Domain produced classification accuracy results above the thresholds set. Figures 2 to 5, present the distributions of the number of feature combinations versus their overall classification accuracy, from the Spatial, and 1st, 2nd, and 3rd overcomplete wavelet transform domains, respectively.

TABLE I
CLASSIFICATION ACCURACY RESULTS FROM SPATIAL DOMAIN

Classifier	Sensitivity	Specificity	Overall
BC	90.82%	91.76%	91.29%
MD	94.35%	90.82%	92.59%
QC	93.88%	98.82%	96.35%

TABLE II
CLASSIFICATION ACCURACY RESULTS FROM THE 1ST SCALE OF THE WAVELET DOMAIN

Classifier	Sensitivity	Specificity	Overall
BC	84.00 %	92.24%	88.47%
MD	86.12%	91.06%	88.59%
QC	92.94%	88.00%	90.47%

TABLE III
CLASSIFICATION ACCURACY RESULTS FROM THE 2ND SCALE OF THE WAVELET DOMAIN

Classifier	Sensitivity	Specificity	Overall
BC	88.94%	90.12%	89.53%
MD	86.12%	87.76%	86.94%
QC	93.65%	91.06%	92.35%

TABLE IV
CLASSIFICATION ACCURACY RESULTS FROM THE 3RD SCALE OF THE WAVELET DOMAIN

Classifier	Sensitivity	Specificity	Overall
BC	91.53%	95.76%	93.65%
MD	86.12%	88.24%	87.18%
QC	95.06%	93.41%	94.24%

TABLE V
BEST CLASSIFICATION ACCURACY RESULTS FROM ALL DOMAINS

	Sensitivity	Specificity	Overall
Spatial	93.88%	98.82%	96.35%
1st Level	92.94%	88.00%	90.47%
2nd Level	93.65%	91.06%	92.35%
3rd Level	95.06%	93.41%	94.24%

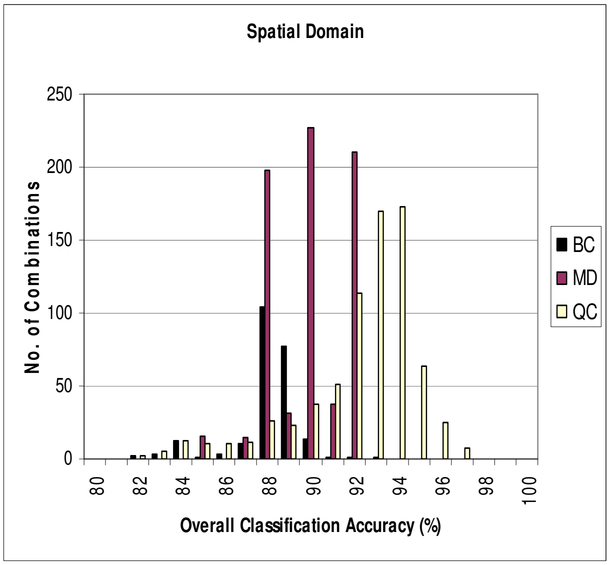


Figure 2. Distribution of the number of feature combinations vs. the overall classification accuracy for features collected from the Spatial Domain.

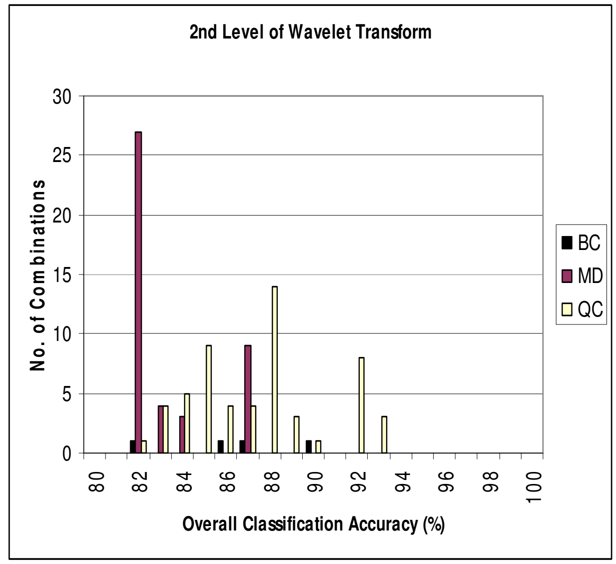


Figure 4. Distribution of the number of feature combinations vs. the overall classification accuracy for features collected from the 2nd level of the overcomplete wavelet transform domain.

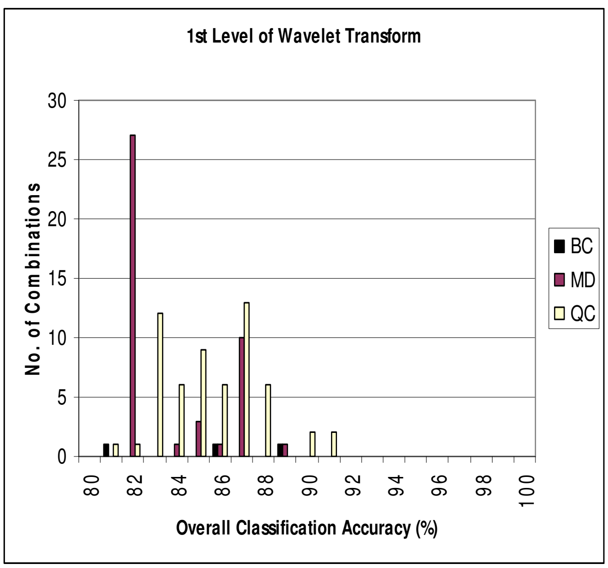


Figure 3. Distribution of the number of feature combinations vs. the overall classification accuracy for features collected from the 1st level of the overcomplete wavelet transform domain.

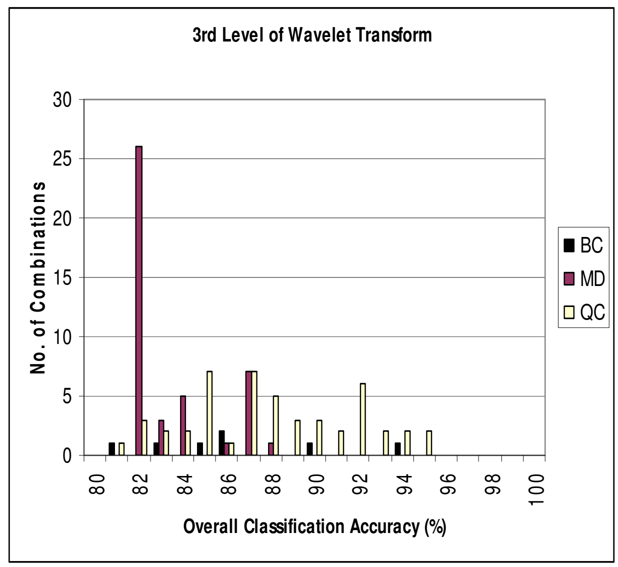


Figure 5. Distribution of the number of feature combinations vs. the overall classification accuracy for features collected from the 3rd level of the overcomplete wavelet transform domain.

IV. DISCUSSION

The aim of this study is to examine the performance of the Wavelet Transform based analysis and classification on Liver CT datasets, and in particular to determine whether we can distinguish between the general classes of normal and cancer liver tissue.

The usage of statistical features for the analysis and classification of textured images has been extensively demonstrated in the literature. Our results suggest that features from the 2nd Order Statistics achieved the best classification accuracy results, since such measurements focus on the overall nature of the texture such as homogeneity, contrast, the presence of organized structure, complexity, and the grey tone transitions within the image.

Although numerous publications have presented and evaluated different Computer Aided Diagnosis schemes, one has to keep in mind that the detection accuracy of any CAD system depends upon the set of images used. This includes the number of images used throughout the training stage of the classification scheme, as well as properties of the images, such as resolution and depth, type of abnormalities included etc.

V. CONCLUSIONS

In this paper a feasibility study of liver CT dataset classification, using different scales of the wavelet transform analysis in conjunction with statistical pattern recognition methods is presented. In our study 850 extracted sub-images from 19 Liver CT were used, in order to establish which features distinguish better between the normal/cancer classes. Twenty statistical measurements were collected; from the images as well as from their different scale wavelet transform coefficients. We found by using the Leave-One-Out method that the combination of the features from the 1st and 2nd Order statistics, achieved overall classification accuracy more than 90.0%, both specificity and sensitivity more than 90.0%. Features selected by the spatial domain performed better than the wavelet-based techniques, under the classification rule of Quadratic Classifier (QC). In addition, features selected by the 3rd scale wavelet transform coefficients performed better than the other wavelet-based techniques, under the classification rule of Bayesian Classifier (BC).

Another advantage of using the wavelet transform coefficients, instead of the spatial domain signal, is that the processing delay/cost needed in the feature extraction stage is a lot less due to the compacted representation of the wavelet transform. In addition we demonstrated that high classification accuracy could be achieved using only compacted data. Possible applications of systems like the one presented in this paper are in content-based classification, search and retrieval of images, and for image processing and classification.

REFERENCES

[1] H.D. Cheng, X. Cai, X. Chen, L. Hu, X. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: a survey", Pattern Recognition, 2003, 36, pp. 2967-2991.

[2] S. Yoshino, A. Kobayashi, T. Yahagi, H. Fukuda, M. Ebara, M. Ohto, "Neural Network Approach to Characterization of Cirrhotic Parenchymal Echo Patterns", IEICE Transactions on Fundamentals, 1993, E76-A, pp. 1316-1322.

[3] S. Yoshino, A. Kobayashi, T. Yahagi, H. Fukuda, M. Ebara, M. Ohto, "Quantitative Diagnosis on Magnetic Resonance Images of Chronic Liver Disease Using Neural Networks", IEICE Transactions on Fundamentals, 1994, E77-A, pp. 1846-1850.

[4] E.L. Chen, P.C. Chung, C.L. Chen, H.M. Tsai, C.I. Chang, "An Automatic Diagnostic System for CT Liver Image Classification", IEEE Transactions on Biomedical Engineering, 2000, 45, pp. 783-794.

[5] J.S. Lee, Y.N. Sun, X.Z. Lin, "A New Approach to Ultrasonic Liver Image Classification", IEICE Transactions on Information & Systems, 2000, E83-D, pp. 1301-1308.

[6] W.L. Lee, Y.C. Chen, K.S. Hsieh, "Ultrasonic Liver Tissue Classification by Fractal Feature Vector Based on M-Band Wavelet Transform", IEEE Transactions on Medical Imaging, 2003, 22, pp.382-392.

[7] H. Yoshida, D.D. Casalina, B. Keserci, A. Coskun, O. Ozturk, A. Savranlar, "Wavelet-Packet-Based Texture Analysis for Differentiation Between Benign and Malignant Liver Tumours in Ultrasound Images", Physics in Medicine and Biology, 2003, 48, pp. 3735-3753.

[8] M. Gletsos, S.G. Mougiakakou, G.K. Matsopoulos, K.S. Nikita, A.S. Nikita, D. Kelekis, "A Computer-Aided Diagnostic System to Characterize CT Focal Liver Lesions: Design and Optimization of a Neural Network Classifier", IEEE Transactions on Information Technology in Biomedicine, 2003, 7, pp. 153-162.

[9] M. Vetterli, J. Kovacevic, "Wavelets and Subband Coding", Prentice Hall, Englewood Cliffs N.J., USA, 1995.

[10] M.V. Wickerhauser, (1994): "Adapted Wavelet Analysis from Theory to Software", A.K. Peters, Wellesley Massachusetts, USA, 1994.

[11] A.K. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall International Editions, N.J., USA, 1989.

[12] K. Fukunaga, "Introduction to Statistical Pattern Recognition", Second Edition, Academic Press, San Diego CA., USA, 1990.