

## Computer-Aided Diagnosis of Liver Cirrhosis by Simultaneous Comparisons of the Ultrasound Images of Liver and Spleen

Henry Horng-Shing Lu<sup>1</sup>, Chung-Ming Chen<sup>2</sup>, Yi-Ming Huang<sup>1</sup>  
and Jen-Shian Wu<sup>1</sup>

<sup>1</sup>*National Chiao-Tung University* and <sup>2</sup>*National Taiwan University*

*Abstract:* Ultrasound imaging is an important tool for early detection and regular check-ups of liver cirrhosis. The diagnosis can be performed by analysis of echo textures of the liver and of the accompanying spleen. The simultaneous comparison of liver and spleen images for the same person at the same system setup can be used to reduce subject, machine, and system variations. This study aims to investigate the computer-aided diagnosis of features derived from the ultrasound images of livers and the accompanying spleens. We will incorporate the techniques of an early vision model, dimension reduction, fractal dimension, nonparametric discriminant rules by kernel density estimation and classification trees to improve the statistical analysis methods. These methods are tested by the clinical images collected at National Taiwan University Hospital with 64 normal livers and 30 cirrhosis ones. The smallest overall bootstrap prediction error is found to be 5.29% by these new methods.

*Key words:* Classification trees, dimension reduction, early vision model, fractal dimension, kernel density estimation, liver cirrhosis, ultrasound.

### 1. Introduction

For many years, hepatic cancers, chronic liver diseases, and liver cirrhosis have remained one of the most popular causes of death in Taiwan according to statistics by the Department of Health<sup>2</sup>. Therefore, it is important to have a reliable diagnosis for the diffuse liver diseases in early detection and regular check ups.

Ultrasound imaging systems are used to diagnose diffuse liver diseases because of their non-invasiveness, ability to do real-time scanning, low cost, and versatility. However, due to the heterogeneous characteristics of ultrasound imaging systems, previous studies on computer-assisted diagnosis usually only consider one

---

<sup>2</sup>See <http://www.doh.gov.tw/>

system (Parker *et al.*, 1988; Momenann *et al.*, 1988; Garra *et al.*, 1989; Hartman *et al.*, 1993; Sun *et al.*, 1996; Lu *et al.*, 1999; Pavlopoulos *et al.*, 2000). Moreover, many parameters related to the physical instrument of ultrasound imaging are subject to a fixed setup. This limitation has become a serious problem when one attempts to diagnose a variety of patients with various tissue structures by different system setups.

This study is thus motivated to develop a more robust computer-aided diagnosis system for diagnosing liver cirrhosis on ultrasound images based on the diagnosis principle employed by the medical doctors at the National Taiwan University Hospital (NTUH). One example of normal and cirrhosis cases is demonstrated in Figure 1. We can use the echo texture of the same person's spleen as a reference in making a diagnosis, which alleviates the limitation imposed on the algorithms by the discrepancy of human bodies, ultrasound systems, and parameter setups. We also take into account the periodic pattern or the human early vision model in an attempt to mimic the natural way that a medical doctor makes the diagnosis (Chen *et al.*, 2000, 2001; Chen and Lu, 2001; Chen *et al.*, 2003). Advanced data mining techniques, including dimension reduction (Li, 1991; Li *et al.*, 2000), fractal dimension (Cherkassky and Mulier, 1998; Akiyama *et al.*, 1990), the nonparametric discriminant rule by kernel density estimation (Fortin *et al.*, 1992), and the classification tree (Silverman, 1986; Breiman *et al.*, 1984), are integrated to explore features and improve prediction of computer-aided diagnosis in this study.

The methods and materials in this study are described in Section 2. Empirical results and comparisons are provided in Section 3. Conclusions and discussions are given in Section 4.

## 2. Methods and Materials

Instead of applying features related to the physical setups of ultrasound images for livers, like attenuation coefficients, statistics of diffuse (random), and specular (structural) backscatter intensities in literature (Parker *et al.*, 1988; Momenann *et al.*, 1988; Garra *et al.*, 1989; Hartman *et al.*, 1993; Lu *et al.*, 1999), we consider features that are robust to the system setups in comparing the ultrasound images of liver and of the accompanied spleen. Features in the space and frequency domains are considered to represent the spatial and periodic pattern, including moving blocks of images, the Fourier coefficients, and an early vision model by Gabor filter banks (Chen *et al.*, 2000, 2001; Chen and Lu, 2001; Chen *et al.*, 2003). Selected statistics, dimension reduction (Li, 1991; Li *et al.*, 2000; Cherkassky and Mulier, 1998) and fractal dimension (Sun *et al.*, 1996; Pavlopoulos *et al.*, 2000; Akiyama *et al.*, 1990; Fortin *et al.*, 1992) are integrated to extract the major variations of echo textures. Classification rules of kernel density

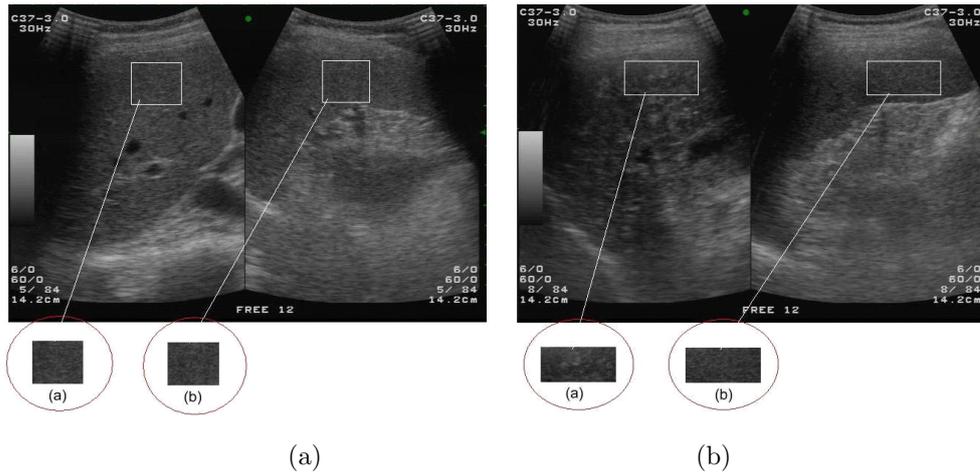


Figure 1: There are similar and fine echo-textures for (a) the liver and (b) the accompanied spleen for that person with a normal liver in part I. Another person with liver cirrhosis has coarser echo-texture in (a) the liver than that in (b) the accompanied spleen in part II.

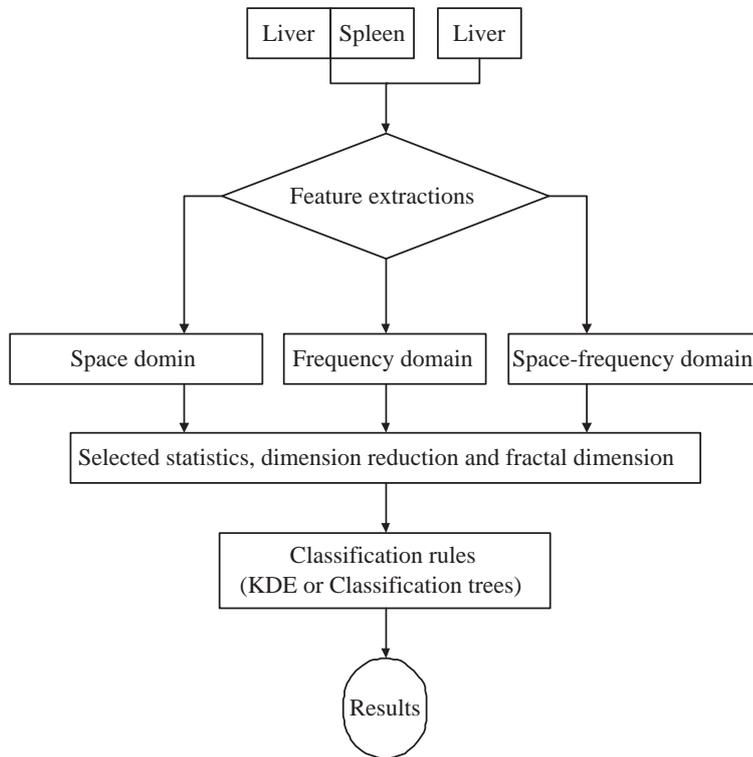


Figure 2: The flowchart of our CAD for liver cirrhosis by ultrasound images is displayed.

estimation and the classification tree (Silverman, 1986; Breiman *et al.*, 1984; Hand, 1997) are then applied to discriminate normal and cirrhosis cases. The

flowchart of analysis is displayed in Figure 2. The block 'Liver/Spleen' represents the input of liver and spleen images simultaneously. In order to compare the results of simultaneous comparisons of liver/spleen with the analysis results of liver only, similar procedures of analysis are applied solely to the liver images. Hence, there is another block for 'Liver' in Figure 2, which denotes unique input of liver images. Finally, these two strategies can be combined in Figure 2.

I. A normal case (Figure 1a),

II. A cirrhosis case (Figure 1b).

The test images were clinical images selected by medical doctors at NTUH and captured from a Toshiba SSA-380A clinical ultrasound imaging system through a frame grabber card. Images were from the RGB output of the Toshiba SSA-380A and were captured by the frame grabber card, Meteor-II card, made by the Matrox Electronic System Ltd. The captured image was stored in the BMP format with 8-bit resolution for each color channel. There were 94 samples, with 64 normal livers and 30 cirrhosis ones diagnosed by experienced medical doctors, collected from clinics at NTUH from August 1998 to January 1999. Two typical examples are given in Figure 1. Experienced physicians select one region of interest (ROI) of the liver image and the corresponding ROI of the spleen image to compare of echo-textures. The system setups and depths of both ROI's are the same to control the variations of machine and system setups. The sizes of both ROI's are the same for a liver and the accompanied spleen of the same person, which are denoted by  $M$  and  $N$ . For different persons or different scans, the sizes of the ROI's may be different due to the varying sizes of echo images. The purpose of this study is to develop computer-aided diagnosis that can automatically distinguish the echo textures of livers and the accompanied spleens. To avoid bias caused by physicians, the selection of ROI's should occur before the diagnosis is known. Because textures are local properties, the intensities of neighboring pixels are used as feature vectors in the space domain. For instance, the picture elements (pixels) in a block with size equal to  $m$  by  $n$ , like 16 by 16 or 8 by 8, can be formed as a feature of the central pixel, as illustrated in Figure 3 and plotted in Figure 4. But this kind of feature in the space domain is very sensitive to the shifting of the center even if the texture structure remains the same. Hence, it is necessary to consider other kinds of transformation to preserve the textures in the presence of translated centers. In order to have more distinguishable feature vectors, the absolute value of Fourier transforms and an early vision model by Gabor filter banks (Chen *et al.*, 2000, 2001; Chen and Lu, 2001; Chen *et al.*, 2003) have been employed to construct the feature vectors in the frequency and the space-frequency domains, as illustrated in Figures 5 and 6.

I. For the normal case in Figure 1 (Figures 3a and 3b):

II For the cirrhosis case in Figure 1 (Figures 3c and 3d):

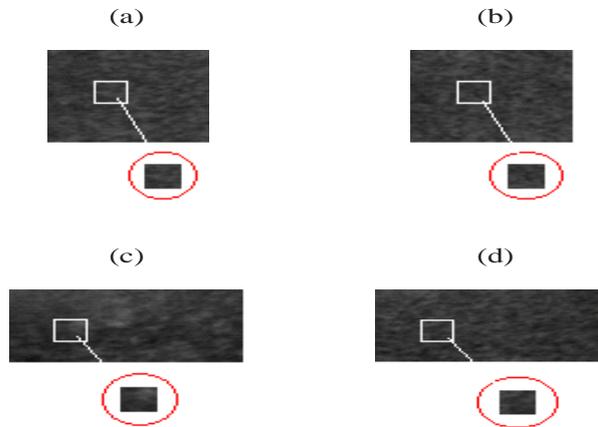


Figure 3: Part I: (a) A local block of a normal liver and (b) a local block of the accompanied spleen are displayed in the part I of Figure 1. Part II: (a) A local block of a cirrhosis liver and (b) a local block of the accompanied spleen are illustrated in part II of Figure 1.

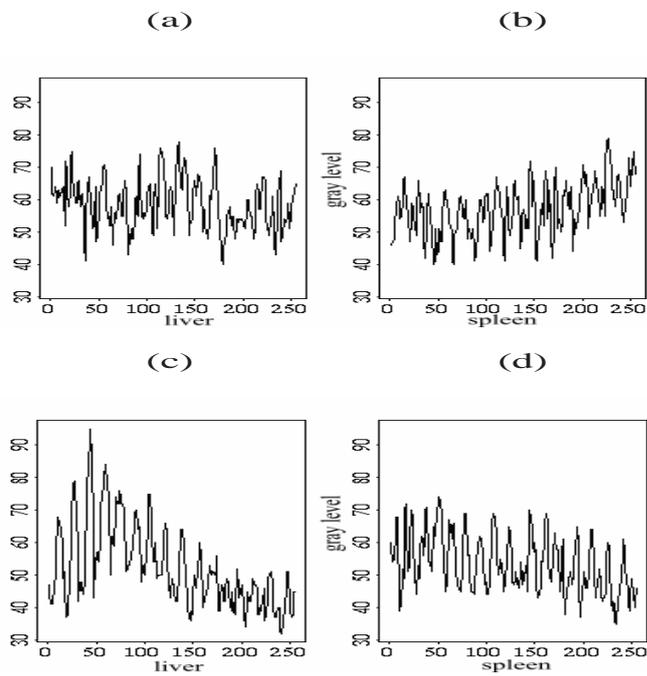


Figure 4: Plots of the feature vectors in those four local blocks of Figure 3 in the space domain.

I. For the normal case in Figure 1 (see Figures 4a and 4b):

II. For the cirrhosis case in Figure 1 (see Figures 4c and 4c):

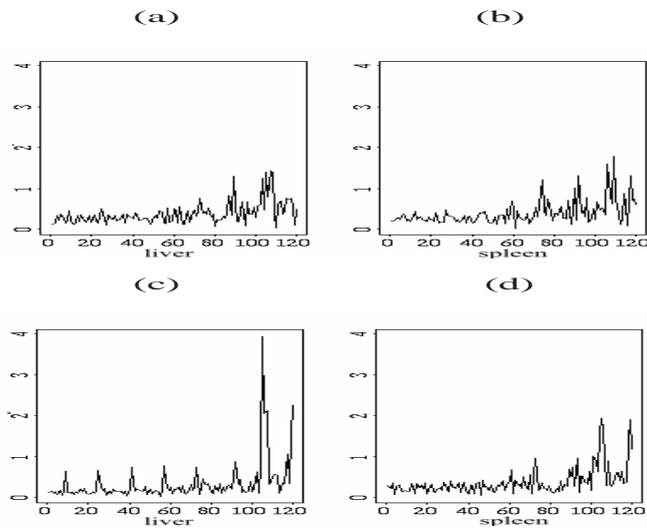


Figure 5: Plots of the feature vectors in those four local blocks of Figure 3 in the frequency domain

I. For the normal case in Figure 1 (see Figures 5a and 5b):

II. For the cirrhosis case in Figure 1 (see Figures 5c and 5d):

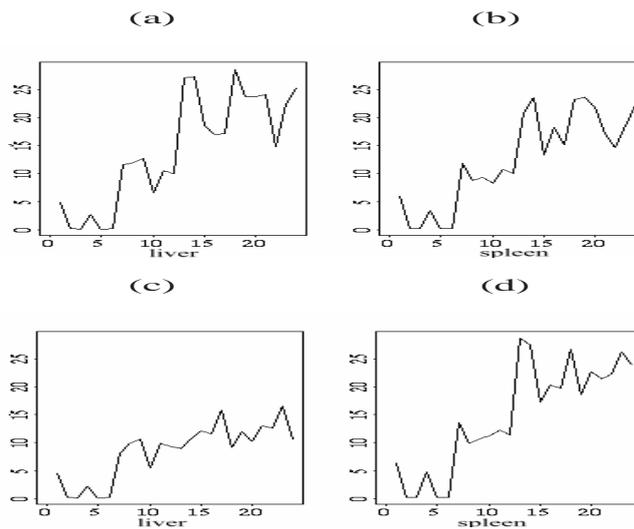


Figure 6: Plots of the feature vectors in those four local blocks of Figure 3 in the space-frequency domain.

I. For the normal case in Figure 1 (see Figures 6a and 6b):

II. For the cirrhosis case in Figure 1 (see Figures 6c and 6d):

Suppose the ROI is of size  $M$  by  $N$ , and a local block in the ROI is  $m$  by  $n$ . There are  $(M - m + 1) \times (N - n + 1)$  moving blocks in total. Because the dimension of each feature vector computed for each block is high, dimension reduction would be required to alleviate the curse of dimensionality. Also, because the data size, i.e.,  $(M - m + 1) \times (N - n + 1)$ , is large and varying, data reduction to the same size for all images would be necessary for the purpose of comparison. Principal component analysis (PCA) and sliced inverse regression (SIR) are applied in this study (Li, 1991; Chen *et al.*, 2000; Cherkassky and Mulier, 1998). SIR is a weighted PCA that employs the information of classification labels, which leads a more effective dimension reduction for classification (Li, 1991; Chen *et al.*, 2000). Several selections of statistics are also applied in order to reduce the data size while preserving statistical information such as mean, median, standard deviation (STD), interquartile range (IQR), coefficient of variation (CV, which is defined to be standard deviation divided by mean), skewness coefficient (SC), and kurtosis coefficient (KC). In particular, it is found that the sum of the first and second largest eigenvalues in our data set is greater than 80% of the total sum of all eigenvalues. Hence, the leading two eigenvectors are used as the major projection directions. Twenty-one dimension reduction techniques investigated in this study are summarized in Table 1. Take method 4 in Table 1 as an example; the feature vectors are projected onto the first PCA directions of liver and spleen in the frequency domain. Suppose the size of a local block is 8 by 8. Because the image intensity is real-valued, the absolute value of a two-dimensional Fourier transform is symmetric about the origin. Moreover, the DC components are affected by the setup of ultrasound imaging and can be removed. Since the origin is located at the fifth row and the fifth column after the two-dimensional fast Fourier transform of a 8 by 8 block of a real-valued image, we will only keep the first four rows except those four DC components at the fifth column. The size of feature vector becomes 28 by 1. For an ROI of  $M \times N$  pixels in a liver (or a spleen), there are  $(M - 8 + 1) \times (N - 8 + 1)$  blocks of feature vectors. Applying PCA on these feature vectors of a liver (or a spleen), the leading eigenvector with size of 28 by 1 is obtained. The inner products of the feature vectors and the largest eigenvector generate  $(M - 7) \times (N - 7)$  values. The dimension is reduced to 1 now.

Table 1: Twenty-one possible dimension reduction techniques are investigated in this study

---

Projection onto the PCA directions of livers (and spleens):
Space domain: 1: onto the first PCA direction 2: onto the second PCA direction 3: onto the leading two PCA direction Frequency domain: 4: onto the first PCA direction 5: onto the second PCA direction 6: onto the leading two PCA direction Space-frequency domain: 7: onto the first PCA direction 8: onto the second PCA direction 9: onto the leading two PCA direction
Projection onto the PCA directions of spleens: Space domain: 10: onto the first PCA direction 11: onto the second PCA direction 12: onto the leading two PCA direction Frequency domain: 13: onto the first PCA direction 14: onto the second PCA direction 15: onto the leading two PCA direction Space-frequency domain: 16: onto the first PCA direction 17: onto the second PCA direction 18: onto the leading two PCA direction
Projection onto the SIR directions of livers and spleens: 19: space domain 20: frequency domain 21: space-frequency domain

---

The histograms of the projection values by method 4 for a normal liver and a cirrhosis liver in Figure 1 are displayed in Figure 7. It is evident that a normal case has similar histograms of projected feature vectors for echo textures in liver and the accompanied spleen. On the contrary, the histograms for a cirrhosis liver and the accompanied spleen are different. The statistics like the standard deviation can then be used to distinguish these two cases from each other. If the second largest PCA direction is used for projection in the method 5, the histograms for normal and cirrhosis cases are displayed in Figure 8. The histograms for a cirrhosis liver are quite distinguishable when they are compared with those for a normal one. Statistics like the mean can be used for distinction in this example. The scatter plots for the projections onto the leading two PCA directions in the frequency domain of method 6 are displayed in Figure 9.

I. For the normal case in Figure 1 (see Figure 7, left panel):

II. For the cirrhosis case in Figure 1 (see Figure 7, right panel):

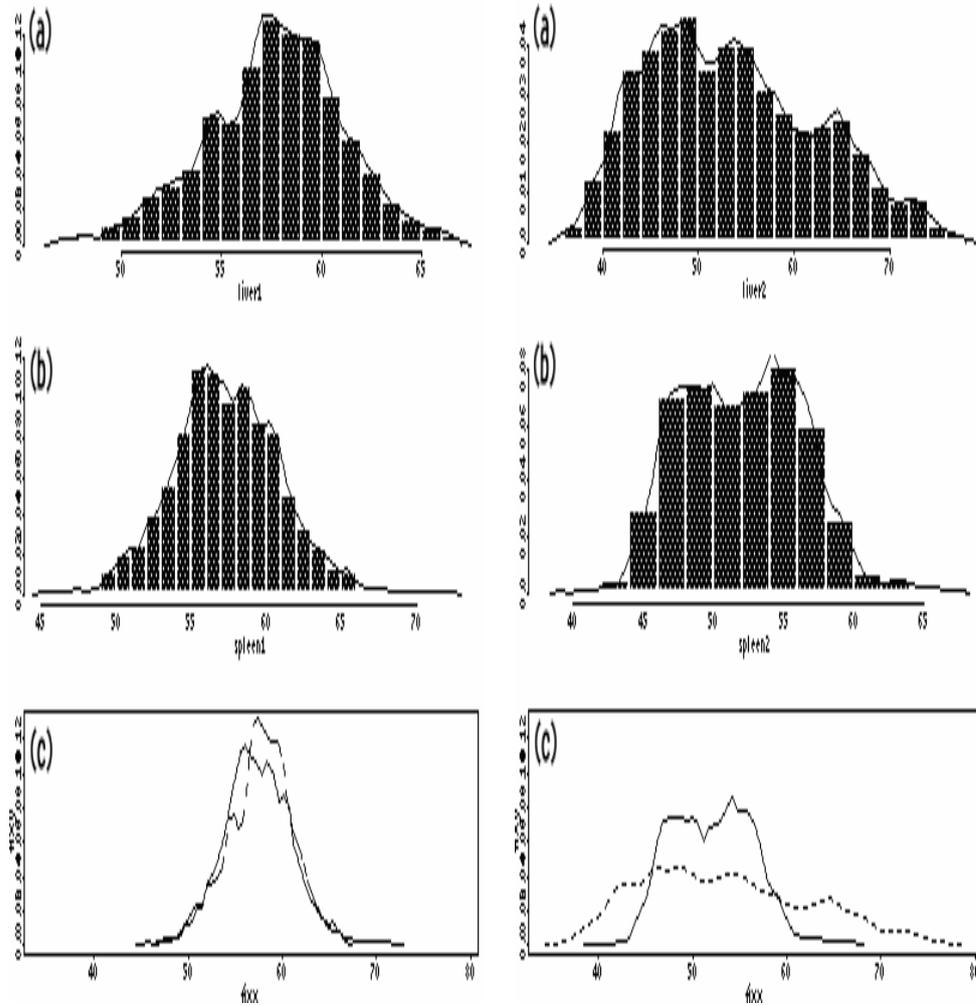


Figure 7: The histograms of projected feature vectors in Figure 1 onto the first PCA directions of liver and spleen in the frequency domain of method 4. Part I: (a) The histogram for a normal liver, (b) the histogram for the accompanied spleen, and (c) the combination of (a) and (b) are displayed. Part II: (a) The histogram for a cirrhosis liver, (b) the histogram for the accompanied spleen, and (c) the combination of (a) and (b) are demonstrated.

I. For the normal case in Figure 1 (see Figures 8, left panel):

I.I For the cirrhosis case in Figure 1 (see Figures 8, right panel):

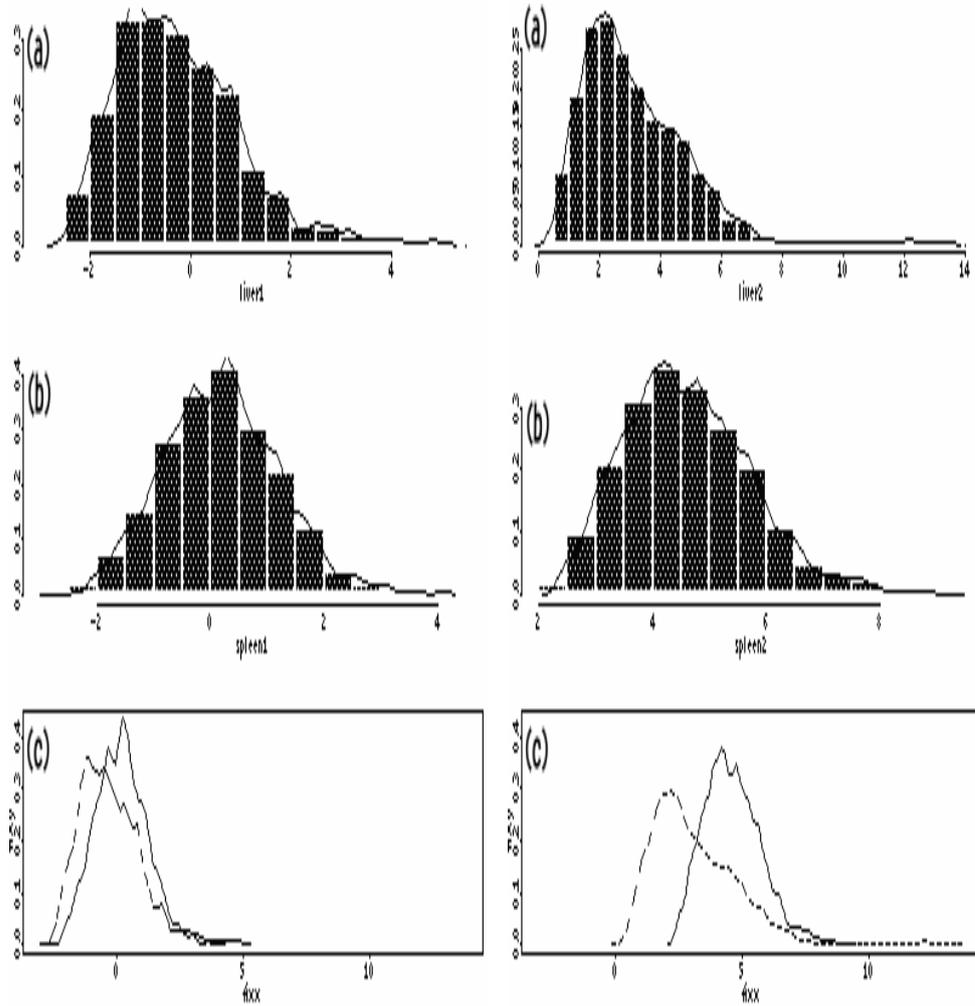


Figure 8: The histograms of projected feature vectors in Figure 1 onto the second PCA directions of liver and spleen in the frequency domain of method 5. Part I: (a) The histogram for a normal liver, (b) the histogram for the accompanied spleen, and (c) the combination of (a) and (b) are displayed. Part II: (a) The histogram for a cirrhosis liver, (b) the histogram for the accompanied spleen, and (c) the combination of (a) and (b) are illustrated.

I. For the normal case in Figure 1 (see Figure 9, left panel):

I.I For the cirrhosis case in Figure 1 (see Figure 9, right panel):

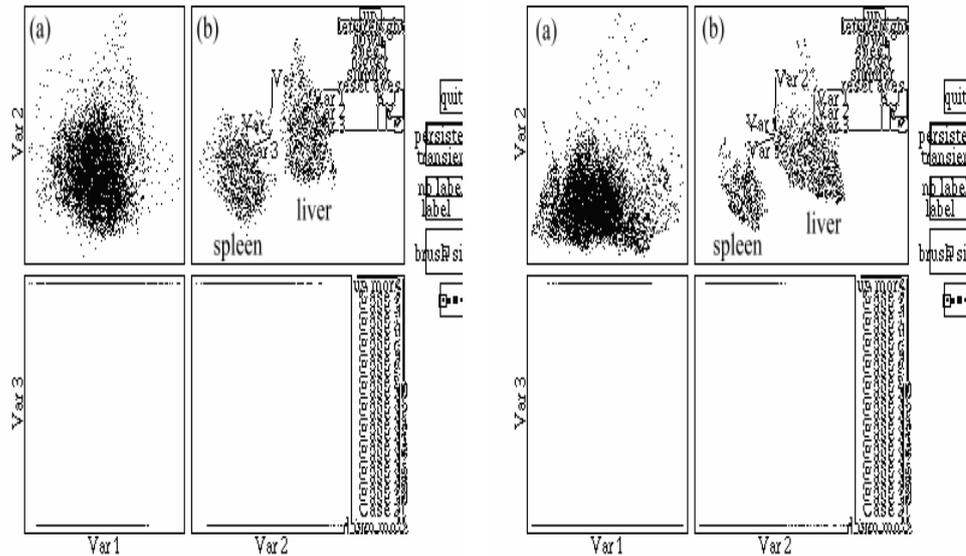


Figure 9: The scatter plots of projected feature vectors in Figure 1 onto the leading two PCA directions of liver and spleen in the frequency domain of method 6. Part I: (a) The 2D scatter plot of a normal liver and the accompanied spleen as well as (b) the 3D scatter plots of a normal liver and the accompanied spleen are displayed. Part II: (a) The 2D scatter plot of a cirrhosis liver and the accompanied spleen as well as (b) the 3D scatter plots of a cirrhosis liver and the accompanied spleen are demonstrated.

Because the echo-textures of spleens in normal and cirrhosis cases are similar, we can also consider the PCA directions of spleens for projection and simultaneous comparisons in methods 10-18. They will be compared with those in methods 1-9. Furthermore, linear combinations of statistics of projected feature vectors are also sought using SIR for classification purposes by incorporating information from classification labels. The linear combination of linear and nonlinear features can even improve the classification and prediction due to the adjustments of weights by class information such that the linear combination of linear and nonlinear features can distinguish two classes effectively (Li, 1991; Li *et al.*, 2000).

If the class information of the whole data set is used in the selection of features or the design of classifiers, then the prediction errors by the resubstitution and the hold-out methods are often biased (Sahiner *et al.*, 2000). To reduce this kind of bias when we use the class information in feature selection (with SIR) and discriminant rules (with the kernel density estimation or the classification tree), we apply the leave-one-out (Jackknife) and bootstrap methods (Efron, 1983; Efron

and Tibshirani, 1993) during the selection of features and the design of classifiers. The leave-one-out method deletes one sample each time. The remaining samples are used to select features and design classifiers for a CAD system. Then, the deleted sample is used to find out the prediction error of this CAD system. The process repeats, and the prediction error of this CAD system is found. The bootstrap method generates re-samples, like 1000 re-samples in this study, and they are used to select features and design classifiers. The prediction errors from bootstrap data set, which do not contain the sample being predicted, are then used to obtain the '.632 bootstrap estimator' of prediction error of this CAD system (Efron, 1983; Efron and Tibshirani, 1993). That is, the SIR method and classifiers are a part of the leave-one-out strategy and/or bootstrapping in our studies. Hence, the bias of the prediction errors found by the linear combination of the SIR method and the classifiers are minimized.

Because the distribution of feature vectors may not be Gaussian distributed, the nonparametric discriminant rule by kernel density estimation (KDE) with a Gaussian kernel (Silverman, 1986) or the classification tree (Breiman *et al.*, 1984; Hand, 1997) is used to perform the last classification step. Since the resulting classifiers are nonlinear, there are no standard tests of stepwise feature selection, like F-to-enter and F-to-remove tests, can be applied here. The prediction errors estimated by the leave-one-out and bootstrap methods are used as a guideline to select the features forwardly.

Fractal dimension has been studied to be useful for classification of texture images and ultrasound images in literature (Sun *et al.*, 1996; Pavlopoulos *et al.*, 2000; Akiyama *et al.*, 1990; Hand, 1997). Hence, this will be considered as a possible feature to be included in our CAD system. The fractal dimension is related to the Hurst coefficient for a fractional Brownian motion (or surface) (Mandelbrot, 1985). The maximum of absolute deviation of image intensities within a neighboring block of varying size can be used to estimate the variance within that block. The relationship between the estimated variance and size gives the estimate of Hurst coefficient (Russ, 1990). This fast computation method of fractal dimension is used in this study. Other methods are discussed in (Peitgen *et al.*, 1992; Wornell, 1996; Bauer and Kohavi, 1999). The empirical results and comparisons of these methods in clinical images will be reported in the next section.

### 3. Empirical Results and Comparisons

(1) L/S (liver/spleen): We start with simple features and statistics to distinguish the differences of echo textures for normal and cirrhosis livers with the reference of the accompanied spleens. Empirical studies for the clinic images are studied

for methods 1-12 in Table 1 with two block sizes, 8 by 8 and 16 by 16. The non-parametric discriminant rule by kernel density estimation (KDE) is used now. The smallest overall bootstrap prediction errors are marked with stars and reported in Table 2 for the feature vectors in the space domain. Thus, the smallest overall bootstrap prediction error can be reduced to 16.04% by the simultaneous comparison with method 12 in Table 1. The false positive and negative errors are also reported in Table 2 that answer another perspective when one kind of error, like the misdiagnosis of cirrhosis livers, is more serious than the other. It is also found that the prediction errors are smaller for the features derived from the 8 by 8 block size both in the space and the frequency domains. On the other hand, the prediction errors are smaller for the features computed from the 16 by 16 block size in the space-frequency domain used in this study.

Table 2: The performance of the method that has the smallest overall bootstrap prediction error in percentage found in the space domain, in which the best one is marked with a star.

Statistics	Error Rates	Jackknife	Overall	Bootstrap	Overall
Liver Mean - Spleen Mean	Normal	31.15	34.04	31.10	32.18
	Cirrhosis	39.39		34.18	
Liver Median - Spleen Median	Normal	31.15	32.98	28.88	31.21
	Cirrhosis	36.36		35.51	
Liver STD / Spleen STD	Normal	21.31	23.40	20.40	23.51
	Cirrhosis	27.27		29.28	
Liver IQR / Spleen IQR	Normal	24.59	23.40	22.29	21.70
	Cirrhosis	21.21		20.61	
Liver CV / Spleen CV	Normal	1.64	31.91	1.81	27.50
	Cirrhosis	87.88		75.01	
Liver SC - Spleen SC	Normal	37.70	46.81	31.41	39.45
	Cirrhosis	63.63		54.31	
Liver KC - Spleen KC	Normal	55.74	42.55	54.00	42.17
	Cirrhosis	18.18		20.32	
Linear Combination	Normal	9.84	11.70	12.73	16.04*
	Cirrhosis	15.15		22.16	

Method 12:

Feature extraction: Space domain.

Block size: 8 by 8.

Dimension reduction: Projection onto the leading two PCA directions of spleens.

It is noted that the methods of projection onto individual PCA directions or spleen PCA directions in the space or the frequency domains with 8 by 8 blocks have small overall prediction errors, which are 16.04% and 16.84%. For simplicity, we will focus on methods 1-6 and 10-15 in the further investigation. The smallest

overall bootstrap prediction errors are reported in Table 3. The smallest overall prediction errors in the space and the frequency domains are marked with stars. Assuming binomial distribution, 95% confidence intervals of prediction error rates for those smallest rates found can be computed. From these empirical results, we observe the following comparison results.

Table 3: The smallest overall prediction errors in percentage found among projection methods 1-6 and 7-12 with or without fractal dimension ( $f$ ) are reported. The classification rules are based on KDE. The smallest one in each category in the space and the frequency domains are marked with stars, respectively. The results for those methods that do not produced smallest overall prediction errors are omitted for clarity of tabulation.

(1)						
		Feature: $L/S$			Feature: $L/S + f$	
		Prediction Error:			Prediction Error:	
Proj.	Bootstrap	Jackknife		Bootstrap	Jackknife	
2	21.83	22.34		14.67*	13.83	*space
3	17.05	14.98		17.54	19.15	
5	24.36	24.47		20.62	21.28	
6	22.42	23.40		19.22	22.34	
12	16.04*	11.70	*space	15.68	15.96	
14	16.84*	17.02	*freq	14.67*	13.83	*freq.
(2)						
		Feature: $L(S)$			Feature: $L(S) + f$	
		Prediction Error:			Prediction Error:	
Proj.	Bootstrap	Jackknife		Bootstrap	Jackknife	
2	18.66	19.15		18.70	19.15	
3	10.59*	11.70	*space	10.59*	11.70	*space
5	21.88	22.34		22.28	23.40	
6	13.30*	13.83	*freq.	13.54*	15.96	*freq.
12	11.67	17.02		11.67	12.77	
14	27.60	30.85		25.31	26.60	
(3)						
		Feature: $L(S) + L/S$			Feature: $L(S) + L/S + f$	
		Prediction Error:			Prediction Error:	
Proj.	Bootstrap	Jackknife		Bootstrap	Jackknife	
2	19.87	21.28		19.87	21.28	
3	12.58	14.89		12.58	14.89	
5	13.50*	12.77	*freq.	14.16	13.83	
6	16.30	18.09		16.11	18.09	
12	10.78*	12.77	*space	10.78*	12.77	*space
14	14.16	13.83		9.98*	9.57	*freq.

(2)  $L/S$  vs.  $L/S + f$  (fractal dimension):

From part (1) of Table 3, fractal dimension can reduce the smallest overall bootstrap prediction errors for the simultaneous comparisons of levers and spleens in both space and frequency domains, from 16.04% and 16.84% to 14.67%.

(3)  $L(S)(+f)$  vs.  $L/S(+f)$ :

The smallest overall bootstrap prediction errors using mainly the statistics of echo textures for liver images are smaller than those by simultaneously comparing livers and spleens based on part (1) and (2) of Table 3. Projection methods 3 and 6 using the projection onto the leading two PCA directions of livers produce the overall bootstrap prediction errors of 10.59% and 13.30%, which are the smallest among the test methods in space and frequency domains. These prediction errors remain the same no matter whether fractal dimensions are used in this study. These results suggest that the statistics of echo textures for liver images are very effective by our methods and these features shall be further combined with the simultaneous comparisons of livers and spleens to reduce the prediction errors. These are studies and investigated next.

(4)  $L(S) + L/S(+f)$ :

The results for using the echo textures of livers as well as the simultaneous comparisons of livers and spleens are reported in part (3) of Table 3. The smallest overall bootstrap prediction errors occur at 9.98% when these features and fractal dimensions are used in the frequency domain. Next, we would like to investigate the performance if we replace the nonparametric classification rule of KDE by classification trees.

Table 4: The smallest prediction errors in percentage found by KDE with different methods using all features in both of the space and the frequency domains.

(1) $L(S) + L/S$		
Feature extraction: Space + Frequency domain		
Prediction errors of linear combination:		
Proj.	Bootstrap	Jackknife
1&4	7.82	7.45
2&5	10.13	9.57
3&6	29.34	24.47
10&13	9.78	9.57
11&14	7.34	7.45
12&15	29.23	18.09
(2) $L(S) + L/S + f$		
Feature extraction: Space + Frequency domain		
Prediction errors of linear combination:		
Proj.	Bootstrap	Jackknife
1&4	7.81	7.45
2&5	9.35	8.51
3&6	29.87	18.85
10&13	9.80	9.57
11&14	5.29*	4.25*
12&15	29.64	18.09

Table 5: The prediction errors in percentage of false positives and negatives for those methods that use all features, statistics, and fractal dimension in the part (2) of Table 4.

(a) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the first PCA directions of livers and spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	6.78	7.81	6.56	7.45
Cirrhosis	9.71		9.09	
(b) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the second PCA directions of livers and spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	8.81	9.35	8.20	8.51
Cirrhosis	10.35		9.09	
(c) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the leading two PCA directions of livers and spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	31.68	29.87	21.31	18.85
Cirrhosis	26.52		16.39	
(d) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the first PCA directions of spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	6.50	9.80	6.56	9.57
Cirrhosis	15.89		15.15	
(e) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the second PCA directions of spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	2.81	5.29*	1.64	4.26*
Cirrhosis	9.86		9.09	
(f) Feature extraction: Space + frequency domains. Dimension reduction: Projection onto the leading two PCA directions of spleens. Statistics: linear combination.				
Error Rates	Bootstrap	Overall	Jackknife	Overall
Normal	2.81	5.29	16.39	18.09
Cirrhosis	9.86	21.21		

#### (5) Classification trees:

The Jackknife prediction errors of classification trees in various methods are evaluated. Based on the comparisons of prediction errors with the corresponding errors in Table 3, classification trees do not reduce prediction errors more than KDE does in our studies. More advanced techniques in classification trees, like bagging, boosting, and other variants, would be necessary to reduce the prediction errors of classification trees with more computation efforts (Dietterich, 2000;

Haralick *et al.*, 1973) in future research.

(6) Space + Frequency domains:

The results of combining all feature vectors and statistics in the space and frequency domains with KDE are reported in Table 4. The smallest overall bootstrap prediction error is reduced to 5.29% when methods 11 and 14 are used with all feature vectors, statistics, and fractal dimensions in the space and frequency domains. Again, the classification trees result in higher Jackknife prediction errors in our studies. The confidence interval of the lowest prediction error of 5.29% is [0.76%, 9.81%] by assuming binomial distribution. This is compatible to the performance of previous approaches in terms of correct classification rates of cirrhosis, e.g. 88.00%-97.30% (Hartman *et al.*, 1993), 66.75% (Sun *et al.*, 1996), 68.00%-80.00% (Pavlopoulos *et al.*, 2000). However, it should be noted that previous approaches required all images be acquired with the same system setup, but this constraint has been relaxed in our method. That is, our method is more robust to varying subject, machine and system setups than previous approaches. The prediction errors of false positives and negatives for those methods that have smallest overall prediction errors in the part (2) of Table 4 are reported in Table 5. This provides a basis for the selection of methods when asymmetric loss is considered.

#### 4. Discussions and Conclusions

We have constructed a CAD system that uses advanced data mining techniques to compare echo textures of livers and their accompanied spleens as medical doctors practice at National Taiwan University Hospital. We start with simple methods to construct this CAD system and gradually increase the complexity by introducing new features and analysis tools when they are shown to be useful in reducing the prediction errors for clinical images. The smallest bootstrap prediction error is found to be 5.29% by combining dimension reduction, KDE, derived features of fractal dimension, liver textures, and simultaneous comparisons of echo textures for livers vs. the accompanied spleens. Intermediate improvements of different components are also evaluated and reported in this study. Other features (Parker *et al.*, 1988; Momenann *et al.*, 1988; Garra *et al.*, 1989; Hartman *et al.*, 1993; Lu *et al.*, 1999; Wu and Chen, 1992; He and Wang, 1990) and analysis methods (Sun *et al.*, 1996; Pavlopoulos *et al.*, 2000; He *et al.*, 1989; Specht, 1990) can be explored and integrated into our current CAD system in the future. For instance, more advanced techniques in classification trees, like bagging, boosting, and other variants, may be used to reduce the prediction errors of classification trees with extra computational costs (Bauer and Kohavi, 1999; Dietterich, 2000). Other methods for computing the fractal dimension (Peitgen *et al.*, 1992; Korvin,

1992; Wornell, 1996) are of interest in comparisons. More sophisticated methods to analyze the information in the space-frequency domains can be investigated. In this study, experienced medical doctors at National Taiwan University Hospital perform the diagnosis of cirrhosis. It is the aim of our future studies to collect biopsies or CT/MR scans. More inputs features of related clinical tests and clinical information for this CAD system will certainly help to improve these methods in practice.

### Acknowledgements

This study was supported in part by grants from the National Science Council in Taiwan, R. O. C.

### References

- Akiyama, I., Saito, T., Nakamura, M., Taniguchi, N., and Itoh, K. (1990). Tissue characterization by using fractal dimension of b-scan image. *IEEE Ultrasound Symposium* **3**, 1353-1355.
- Bauer, E. and Kohavi, R. (1999). An empirical comparison of voting classification algorithms: bagging, boosting, and variants. *Machine Learning* **36**, 105-139.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, J. C. (1984). *Classification and Regression Trees*. Wadsworth.
- Chen, C. M. and Lu, H. H. S. (2001). An adaptive snake model for ultrasound image segmentation: modified trimmed mean filter, ramp integration and adaptive weighting parameters. *Ultrasonic Imaging* **22**, 214-236.
- Chen, C. M., Lu, H. H. S., and Chen, Y. L. (2003). A discrete region competition approach incorporating weak edge enhancement for ultrasound image segmentation. *Pattern Recognition Letters* **24**, 693-704.
- Chen, C. M., Lu, H. H. S., and Han, K. C. (2001). A textural approach based on gabor functions for texture edge detection in ultrasound images. *Ultrasound in Medicine and Biology* **27**, 515-534.
- Chen, C. M., Lu, H. H. S., and Lin, Y. C. (2000). An early vision based snake model for ultrasound image segmentation. *Ultrasound in Medicine and Biology* **26**, 273-285.
- Cherkassky, V. and Mulier, F. (1998). *Learning from Data: Concepts, Theory, and Methods*. Wiley.
- Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. *Machine Learning* **40**, 139-158.
- Efron, B. (1983). Estimating the error rate of a prediction rule: improvements of cross-validation. *J. Amer. Statist. Assoc.* **78**, 316-331.

- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.
- Fortin, C. S., Kumaresan, R., Ohley, W. J., and Hofer, S. (1992). Fractal dimension in the analysis of medical images. *IEEE Engineering in Medicine and Biology* **11**, 65-71.
- Garra, B. S., Insana, M. F., Shawker, T. H., Wagner, R. F., Bradford, M., and Res-sell, M. (1989). Quantitative ultrasonic detection and classification of diffuse liver disease comparison with human observer performance. *Investigate Radiology* **24**, 196-203.
- Hand, D. J. (1997). *Construction and Assessment of Classification Rules*. Wiley.
- Haralick, R. M., Shanugam, K., and Dinstein, I. (1973). Texture features for image classification. *IEEE Trans. Syst. Man. Cybernet* **3**, 610-621.
- Hartman, P. C., Oosterveld, B. J., Thijssen, J. M., Rosenbusch, G. J. E., and Berg, J. v. d. (1993). Detection and differentiation of diffuse liver disease by quantitative echography — a retrospective assessment. *Investigate Radiology* **28**, 1-6.
- He, D. C. and Wang, L. (1990). Texture unit, texture spectrum and texture analysis. *IEEE Trans Geosci Remote Sensing* **28**, 509-512.
- He, D. C., Wang, L., and Guibert, J. (1989). Texture discrimination based on an optimal utilization of texture features. *Pattern Recognition* **21**, 141-146.
- Korvin, G. (1992). *Fractal Models in the Earth Sciences*. Elsevier.
- Li, K. C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86**, 1991.
- Li, K. C., Lue, H. H., and Chen, C. H. (2000). Tree-structured regression via principal hessian directions. *J. Amer. Statist. Assoc.* **95**, 547-560.
- Lu, Z. F., Zagezebski, J. A., and Lee, F. T. (1999). Ultrasound backscatter and attenuation in human liver with diffuse disease. *Ultrasound in Medicine and Biology* **25**, 1047-1054.
- Mandelbrot, B. B. (1985). Self-affine fractals and fractal dimension. *Physica Scripta* **32**, 257-260.
- Momenann, R., Garra, B. S., Loew, M. H., Wagner, R. F., and Insana, M. F. (1988). Imaging of liver-disease using multiple quantitative tissue characterization parameters. *Ultrasonic Imaging* **10**, 71-72.
- Parker, R. J., Asztely, M. S., Lerner, R. M., Schenk, E. A., and Waag, R. C. (1988). In vivo measurements of ultrasound attenuation in normal or diseased liver. *Ultrasound in Medicine and Biology* **14**, 127-136.
- Pavlopoulos, S., Kyriacov, E., Koutsouris, D., Blekas, K., Stafylopatis, A., and Zoumpoulis, P. (2000). Fuzzy neural network-based texture analysis of ultrasonic images. *IEEE Engineering in Medicine and Biology* **19**, 39-47.

- Peitgen, H. O., Jurgens, H., and Saupe, D. (1992). *Chaos and Fractals: New Frontiers of Science*. Spriger-Verlag.
- Russ, J. C. (1990). Surface characterization: fractal dimensions, hurst coefficients, and frequency transforms. *Journal of Computer Assisted Microscopy* **2**, 249-257.
- Sahiner, B., Chan, H. P., Petrick, N., Wagner, R. F., and Hadjiiski, L. (2000). Feature selection and classifier performance in computer-aided diagnosis: the effect of finite sample size. *Medical Physics* **27**, 1509-1522.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- Specht, D. F. (1990). Probabilistic neural networks. *Neural Networks* **3**, 109-118.
- Sun, Y. N., Horng, M. H., Lin, X. Z., and Wang, J. Y. (1996). Ultrasonic image analysis for liver diagnosis. *IEEE Engineering in Medicine and Biology* **15**, 93-101.
- Wornell, G. W. (1996). *Signal Processing with Fractals: A Wavelet-Based Approach*. Prentice-Hall.
- Wu, C. M. and Chen, T. C. (1992). Statistical feature matrix for texture analysis. *Computer Vision, Graphic and Image Processing* **54**, 407-419.

Received February 27, 2008; accepted June 6, 2008.

Henry Horng-Shing Lu  
Yi-Ming Huang  
Jen-Shian Wu  
Institute of Statistics  
National Chiao Tung University  
Hsinchu, 300, Taiwan R.O.C.  
hslu@stat.nctu.edu.tw

Chung-Ming Chen  
Institute of Biomedical Engineering  
National Taiwan University  
Taipei, 100, Taiwan R.O.C.  
chung@ntu.edu.tw