Vector Quantization Distortion of Medical Ultrasound Features

Brian Krasner, Shih-Chung B. Lo, and Seong Ki Mun

Pruned-tree structured vectored quantization (PTSVQ) was applied to the lower five gray scale remapped bits of normal and fatty ultrasound liver images. The upper bits were compressed reversibly. This combination of techniques is termed PTSVQ with splitting. The effect of the compression on the difference in texture between normal and fatty liver images was studied at different compression rates and distortions. The changes in texture were measured by changes in the principal components of the covariance matrix of image vectors. The vectors were the same size as those used in the compression technique. There were clear differences in the components of normal and fatty liver images. These differences were largely removed by the PTSVQ with splitting technique even at average single pixel distortions several times smaller than the image noise. These results suggest that the effect of compression on second order statistics should be measured when evaluating algorithms in addition to the first order average distortion.

KEY WORDS: ultrasound liver images, data compression, tissue characterization, principal components, vector quantization.

SMALL VECTOR quantization errors could cause degradation in features derived from and used to classify medical ultrasound images of normal and fatty livers, according to an earlier report.[1] The quantization errors were small relative to the noise present in the images. Preprocessing of the images that reduced the vector quantization (vq) error actually increased the degradation.[1] The data thus posed the problem of finding the component of error that causes feature degradation given the nonmonotonic relationship between the average quantization error and the feature degradation. Related to this problem, the component of the overall average error that leads to feature degradation may be relatively small. This may due to either spatial or statistical localization of error that leads to the following questions: 1) Can edge regions be found in ultrasound images at which a disproportionately large vq error occurs? If so, what is their frequency of occurrence? and 2) Alternatively, is there a small component in the covariance relationship between neighboring pixels that is disproportionately distorted by vq? In each question, it is assumed that the small component is responsible for the feature used in classification.

This study concentrated on statistical localization of the error. The following experiment, designed to find statistical localization, provided interesting results. The covariance matrix is estimated for pixels in an encoding block. The principal components are calculated. The relationship of different components to feature classification is measured. The selective effect of vq errors on components is determined.

The following results and conclusions were determined: There are principal component (pc) statistics that are related to previously used classification features. More importantly, the pc statistics can be used as classification features. The principal components generating these features are in fact small relative to the main components and are selectively degraded by vq, even with small overall normalized mean squared errors.

These results indicate that vq effectively represents the major correlations between pixels in a block but not the minor components. The major correlations are enough to represent significant classification information that is stored in pixels in different blocks. However, useful classification information within a block is lost. A vq method that explicitly encodes pixel statistical components instead of values may be a useful method of resolving this problem.

METHODS

The methodology in this article can be summarized as follows:

1. An ultrasound radiologist collected ultrasound images of normal and fatty livers;
2. The ultrasound radiologist selected re-

From the Department of Radiology, Georgetown University Hospital, Washington, DC.

Supported in part by a Whitaker Foundation Grant and National Institutes of Health grant 1R29CA59763-01.

Address reprint requests to Brian Krasner, PhD, Department of Radiology, Georgetown University Hospital, 3800 Reservoir Rd, NW, Washington, DC 20007.

0897-1889/93/0603-0001$3.00/0

gions of interest (ROIs) on the liver for analysis;
3. The images were compressed/decompressed at varying bit rates with pruned-tree structured vectored quantization (PTSVQ) with splitting;
4. Features were measured from each image's ROI;
5. For each class of liver images at different compression rates, eigen matrices and values were calculated for the covariance matrix of the ROI vectors of that class and compression rate;
6. Additional features based on the eigen matrices and values were calculated; and
7. For each compression level for each technique, the ability of a statistical discriminator to distinguish fatty and normal liver images was tested.

Note that the general approach in the methods discussed below is to simulate the design and use of a compression technique in a realistic environment, and then test the effects of this technique with a tissue characterizer experiment designed to reveal any significant statistical effects of the compression. These steps are described in more detail below.

Image Acquisition and Description

Volunteers provided the normal liver images. Fatty livers were selected retrospectively from previously diagnosed cases. For the purposes of this initial study, only obviously fatty livers were selected. Ultrasound images were acquired on three ultrasound machines, an ATL UltraMark-9, an ATL UltraMark-8, and an ATL UltraMark-9 HDI (Advanced Techn Laboratory, Bothel, WA). Details are described in a previous study.[1] The video output from the ultrasound machines was digitized and acquired using a CommView medical imaging network (Philips, Shelton, CT).

For the data compression codebook design (see below), 14 normal liver images, four fatty liver images and two hepatitis images were chosen in approximate proportion to their occurrence in the radiology department examinations. This selection sampling followed the goal of trying to simulate the use of compression in a realistic setting. The normal liver images were taken from five subjects, the fatty liver images from two subjects, and the hepatitis liver images from two subjects. Images taken with both 17- and 12-cm fields of view were included.

To facilitate statistical analysis of the effect of the compression technique on the tissue characterizer, this study used approximately equal numbers of fatty and normal liver images taken at one field of view, 17 cm. Twenty-two normal images from nine subjects and 24 fatty liver images from seven subjects were used. One ROI was taken from each image giving 22 normal liver ROIs and 24 fatty liver ROIs. The ROIs were 1- to 2-cm square regions taken from as near as possible to the center of the liver image and chosen to avoid any deterministic structures such as blood vessels.

Vector Quantization Technique

Vector quantization is a technique for representing a block of image values or vector, by the vector in a codebook that is closest to the original vector. Splitting is a technique for decomposing image pixel values into the high and low values. The high values are compressed reversibly while the low values are compressed via PTSVQ.[2]

The use of PTSVQ to compress medical magnetic resonance images has been described in detail.[3] Briefly, there are two stages of PTSVQ, a codebook design step and the image compression step. The codebook design takes as input a series of input images and uses these to construct a tree structure coder. Coding is accomplished by comparing an input vector to both the left and right node vectors of a binary tree and going to the node whose vector is closest to the input vector. This continues until a leaf of the tree is reached, at which point a code is assigned equal to the binary representation of the path used to reach the leaf. Initially a complete tree is obtained, which codes a vector with the number of bits equal to the number of levels in the tree. If branches are removed from this tree (pruned), when a leaf is reached fewer bits are needed to describe the traversal but at a cost of more distortion in the final representation. Test images are encoded by the complete tree to gain statistics on the average number of times a node is used and the distortion that would be produced if the node's vector was used to represent the image vector. A series of

subtrees are then generated by successively removing the node that produces the least added distortion per reduction in bits needed to represent the image.

The vector size used in this study was 4, corresponding to 2 × 2 blocks of pixels.

Compression rates and distortion were varied by choosing successively smaller size trees produced by the PTSVQ technique.

The error metric used to quantify the error produced by the compression technique was the normalized mean squared error (NMSE). The NMSE is calculated by dividing the average squared difference between the original pixel values and the processed pixel values by the sample variance of the pixel values. The average and sample variance are taken over the whole image.

Generation and Use of Whitening Matrix

A matrix was generated so that a 2 × 2 block of pixels transformed with this matrix would have 0 estimated non-diagonal covariance values. Such a matrix is referred to as a whitening matrix. The values of a vector transformed by this matrix are called components. The process of generating this matrix is shown in Fig 1. The eigenvectors and eigenvalues were determined by the Jacobi method.[4] The eigenvector corresponding to the largest eigenvalue is termed the first eigenvector and so on. The eigenvectors were ordered according to the size of the eigenvalue. Therefore, the first component refers to the dot product of a vector with the first eigenvector and so on (Fig 1).

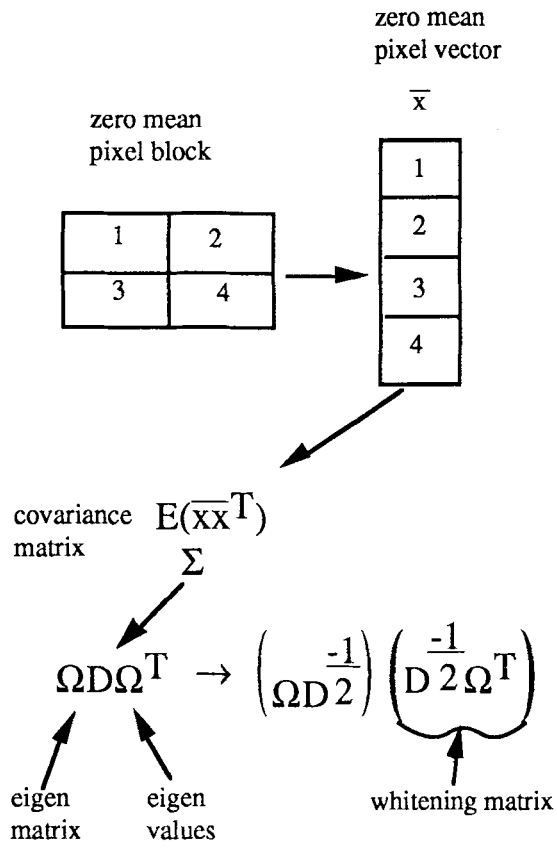

Fig 1. Generation of whitening matrix.

Feature Extraction

Three commonly used features in ultrasound tissue characterization can be derived from co-occurrence matrix, fractal dimension, and run length statistics.[1,5] Of those features, this article will use the long run length emphasis, lrem, for comparison purposes with the principal component statistics, since lrem was found to be the most powerful feature for distinguishing fatty and normal livers. lrem is defined: a histogram of run length frequencies is determined and then the long run length emphasis is determined by

$$\text{lrem} = (1/N) * \Sigma \text{ length}^2 * \text{histogram (length)}$$

where *lrem* is the long run length emphasis, *length* is the run length, *histogram (length)* is a count of the run of the given length and N is the total number of pixels. By multiplying the histogram value by the length squared, a nonconstant relationship between lrem and (length*frequency) is established where the joining of several shorting runs into one longer run will increase the long run length emphasis. Runs were taken in the vertical direction after dividing the pixel values by four.

Features were also derived from the covariance matrices of the ROI vectors. A whitening matrix was calculated from the combined normal ROI vectors as discussed earlier. When classifying a normal ROI, the whitening matrix was recalculated with that ROI's vectors left out. Each ROI's vectors were transformed by the whitening matrix, the resulting components squared, and, finally, the average squared components were calculated. The average squared second component, evar2, and average squared fourth component, evar4, were used as features.

Each ROI's covariance matrix and corresponding eigenmatrix and eigenvalues were also calculated. The first eigenvector of the covariance matrix was selected. Similarly the first eigenvector of the combined normal ROI covariance matrix was selected. The angle between these two vectors, eang1, was used as a feature.

Statistical Analysis Using Hotelling Trace

The Hotelling Trace Criterion, symbolized by J in this section and as the *feature distance* in the rest of this article is used as a measure of the ability of a set of features to distinguish between classes. It is a multivariate generalization of the univariate Student t test statistic. The sample covariance matrix for a class i is defined as

$$Si = E\,[(x - u)(x - u)^T]$$

where E stands for the expected value, u is the sample mean feature vector and x is the observed feature vector value. Then the Hotelling trace criterion, J, is defined as:

$$J = tr\,\{S_w^{-1}\,Sb\}$$

where L is the number of classes, S_w is the within-class scatter matrix,

$$S_w = \sum_{i=1}^{L} P(w_i)\,S_i,$$

and S_b is the between-class scatter matrix,

$$S_b = \sum_{i=1}^{L} P(wi)(ui - uo)(ui - uo)^T.$$

uo is the sample mean for the mixture of all classes. *P(wi)* is the a prior probability for class *wi* and *ui* is the sample mean for class wi. In this study, the number of fatty and normal livers tested for diagnostic accuracy was artificially kept equal so that P(wi) can be left out of the above equations. An important property of J, or the "feature distance," is that for normal distributions it predicts the total diagnostic accuracy for a receiver operating characteristic curve based on the use of those features, where diagnostic accuracy is defined as the percentage of correct choices on a two-alternative (disease, no disease) forced-choice test. The relationship is:

$$Az = \int_{-\infty}^{\sqrt{2J}} \exp(-t^2/2)\,dt$$

where *Az* represents the total diagnostic accuracy. To give an idea of this relationship over the range of feature distances seen in this article, a feature distance of 1 gives an Az of 0.91, whereas a feature distance of 0.4 gives an Az of only 0.82.[6]

RESULTS

Difference Between Overall and ROI NMSE

The NMSE value, as stated earlier, is calculated from the whole image. If the distortion and image variance is calculated just for the ROI regions themselves, the NMSE is much higher. For example, when the overall NMSE is 0.0004, the NMSE for the ROIs is 0.004 for normal liver images and 0.19 for fatty liver images. However, given that a large percentage of the variance is due to noise, the compression distortion even for the ROIs is still several times less than the noise. For comparison with previous articles, the results will be given using the overall NMSE.

Differences in Eigenvectors and Eigenvalues of Normal and Fatty Livers at Different Compression Rates

The eigenvectors and eigenvalues of the covariance matrices for normal and fatty liver ROI vectors are shown in the Table 1 with the eigenvectors shown in order of decreasing size of their corresponding eigenvalues. Table 2 shows similar eigenvectors and eigenvalues for

Table 1. Eigenvalues and Eigenvectors for Uncompressed Fatty and Normal Liver Images

	C1	C2	C3	C4
	Normal Liver			
Eigenvalues	811.4	113.9	28.2	12.0
Eigenvectors	0.50	−0.43	−0.57	0.49
	0.49	−0.57	0.43	−0.49
	0.50	0.55	−0.42	−0.52
	0.51	0.42	0.55	0.50
	Fatty Liver			
Eigenvalues	860.7	75.6	56.5	35.1
Eigenvectors	0.50	−0.40	−0.52	0.57
	0.50	−0.56	0.28	−0.59
	0.49	0.63	−0.45	−0.39
	0.50	0.49	0.67	0.42
|Angle| between eigenvectors	0.02	0.12	0.21	0.20

C1-C4, components 1 through 4.

Table 2. Eigenvalues, Eigenvectors for Fatty and Normal Liver Images at NMSE, 0.0004

	C1	C2	C3	C4
	Normal Liver			
Eigenvalues	810.6	113.4	29.0	12.7
Eigenvectors	0.50	−0.43	−0.57	0.49
	0.49	−0.57	0.44	−0.50
	0.50	0.55	−0.42	−0.52
	0.51	0.43	0.55	0.50
	Fatty Liver			
Eigenvalues	915.4	76.9	25.1	10.6
Eigenvectors	0.50	−0.46	0.53	0.49
	0.50	−0.52	−0.48	−0.49
	0.50	0.52	0.46	−0.51
	0.50	0.48	−0.52	0.50
|Angle| between eigenvctors	0.02	0.08	0.08	0.01

C1-4, components 1 through 4.

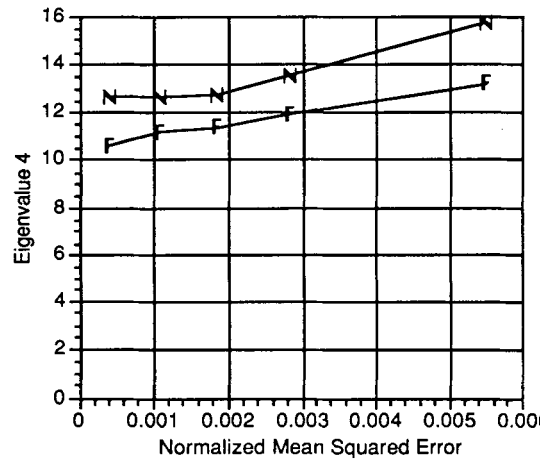

Fig 2. Plot of smallest eigenvalue of covariance matrix versus normalized mean squared error. (—N—, normal liver; —F—, fatty liver).

images compressed by PTSVQ with splitting with a NMSE of 0.0004.

The tables show that there are large differences in uncompressed normal and fatty liver ROI eigen vectors and values. Even a small amount of compression error greatly reduces these differences.

Using the principle component transformation (whitening matrix) derived from the combined normal ROIs, the average squared components of the combined fatty liver ROIs were determined. These component values for uncompressed images were 1.06, 0.67, 2.00, and 3.06. Only the first average squared component value was close to the expected value of 1.0 for a normal ROI vector.

Vector quantization even at the lowest distortion and highest bit rates removes this difference in principal components. The average squared component values with this small error are 1.13, 0.68, 0.88, and 0.87, so that the third and fourth components are now much closer to the expected values for normal ROI vectors.

This elimination of differences in the smallest components is maintained at higher compression rates.

In addition to these immediate effects of even a small amount of vector quantization compression, there are changes that occur as the compression rate increases as can be seen in Figs 2 and 3. For normal ROIs the smallest eigenvalue increases and the largest decreases. There is also an increase in the variance along the direction specified by the smallest component of the uncompressed, normal whitening matrix. Similar changes occur for fatty ROIs.

Use of Components and Eigenvector Angles as Features

Table 3 shows the values of the evar4, evar2, and eangl features for normal and fatty liver images at different compression errors. At 0 compression error, the differences between normal and fatty livers in the evar4 and evar2 features are statistically significant ($P < .05$ using a standard t test). The difference in eangl is not statistically significant. The eangl feature is included because even though by itself it could not distinguish between normal and fatty liver, in combination with the other two features it increases the classification accuracy. With

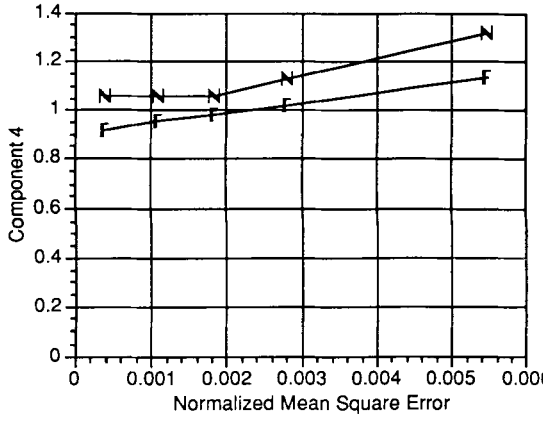

Fig 3. Plot of smallest average squared component of covariance matrix versus normalized mean squared error. (—N—, normal liver; —F—, fatty liver)

Table 3. Effect of PTSVQ With Splitting on Component Features

NMSE (bits/pixel)	evar4		evar2		eang1		Hotelling Trace
	Fat	Normal	Fat	Normal	Fat	Normal	
0.0/8.0	3.13	1.00	0.64	1.00	0.22	0.07	1.52
	(1.99)	(0.38)	(0.28)	(0.40)	(0.20)	(0.04)	
0.0004/3.8	0.85	1.00	0.65	1.00	0.06	0.07	0.34
	(0.32)	(0.35)	(0.27)	(0.39)	(0.05)	(0.03)	
0.005/1.8	0.82	1.00	0.64	0.99	0.06	0.09	0.32
	(0.44)	(0.39)	(0.27)	(0.43)	(0.05)	(0.06)	

NMSEs of 0.0004 and 0.005, the difference in evar4 is no longer statistically significant, while the difference in evar2 remains statistically significant ($P < .05$). The differences in eang1 are not statistically significant at any compression level tested.

As can also be seen in Table 3 the combination of evar4, evar2, and eang1 is a powerful classifier of fatty and normal livers with an estimated accuracy of 96%. The table shows that even a small amount of compression produces an immediate large reduction of classification accuracy, though the residual accuracy is maintained at higher compression rates.

The usefulness of the principal components in distinguishing fatty and normal liver ROIs can be better visualized by examining a scatter diagram of the second and fourth components (Fig 4). Figure 4 shows that only 4 of 24 liver images would be misclassified by a simple linear classifier. However, the scatter diagram in Fig 5

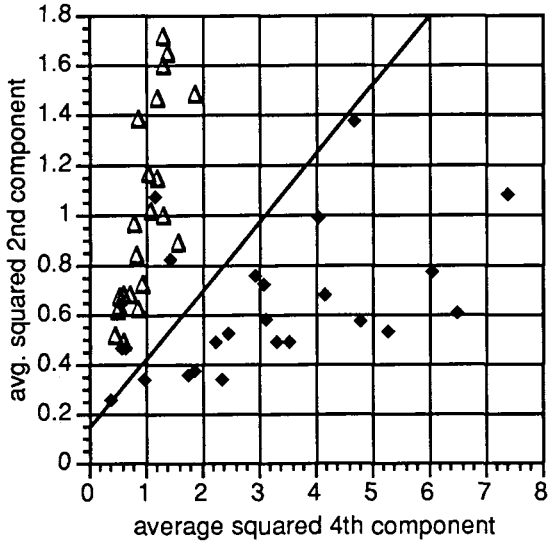

Fig 4. Scatter diagram of normal versus fat liver ROI components for uncompressed images. (△, normal; ◆, fat)

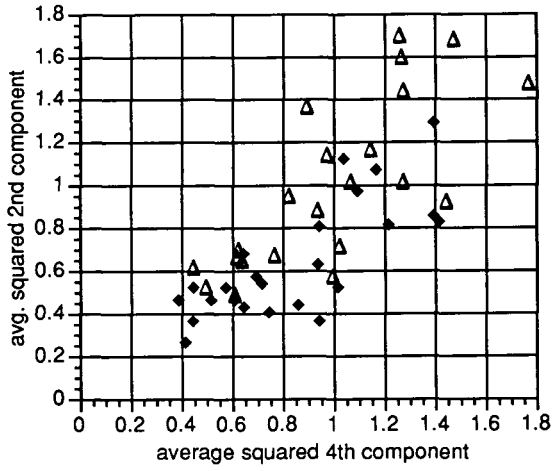

Fig 5. Scatter diagram of normal versus fat liver ROI components for images compressed using PTSVQ with splitting at a normalized mean squared error of 0.0004. (△, normal; ◆, fat)

shows that even a small amount of compression error produces a large amount of overlap between the fatty and normal liver ROI components, which greatly reduces their value as diagnostic features.

CONCLUSION

Relationship of PTSVQ With Splitting to Covariance Components

PTSVQ with splitting did not represent accurately the minor components of the covariance matrix of fatty liver images. Because there is no inherent reason why vector quantization can not represent such patterns, the reason for this failure was probably the relative absence of fatty liver images in the training sets. The minor components of the normal liver images were accurately represented.

Practically, the need to include large number of images of every class in the training set to represent minor components is a serious flaw. The important classes of images may not be known or available during the design phase. Also, representing less frequently occurring but important classes equally well will degrade the overall average performance of the compression technique.

A possible solution would be to code the components directly rather than the pixels. For instance, the image could be divided into 64 × 64 blocks. For each block, the covariance matrix and mean vector for each block would be stored

requiring 56 bytes for a 2 × 2 image. From the covariance matrix in both the compressor and decompressor, a transformation matrix could be determined, which would allow determination of components in the compressor and retrieval of pixel values in the decompressor. The minor components would be compressed at low bit rates, but the direction and overall magnitude would be recovered from the covariance matrix.

Relationship of Component Values, Features With Previous Feature Results

Comparing Fig 6 with Figs 2 and 3, increases in compression error lead to similar increases in vertical long run length emphasis and the smallest eigenvalue of normal liver images. In addition, there is a small, parallel decrease in the second largest eigenvalue. The increase in vertical long run length emphasis reflects a smoothing of the image in the vertical direction (longer vertical runs). The changes in the eigenvalues reflect a similar smoothing in the vertical direction. The eigenvectors corresponding to the smallest and second largest eigenvalues can be rewritten in two-dimensional form (Table 4).

The second largest eigenvalue eigenvector clearly corresponds to high frequency variations (or edges) in the vertical direction, while the smallest eigenvalue eigenvector corresponds to high frequency variations in a diagonal direction. The decrease in the second largest eigenvalue with an increase in the smallest eigenvalue reflects a shift from vertical high frequency components to diagonal high frequency components. The decrease in vertical high frequency components reflects a smoothing in the vertical direction, which parallels the change indicated by the increase in vertical long run length emphasis.

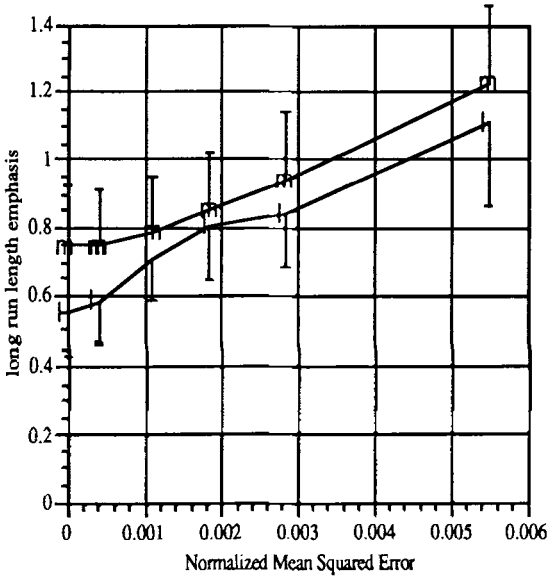

Fig 6. Plot of the long run length emphasis feature versus normalized mean squared error using PTSVQ with splitting. (+, normal liver; —m—, fatty liver)

Table 4. Two-Dimensional Representation of Eigenvectors

Smallest Eigenvalue Eigenvector		Second Largest Eigenvalue Eigenvector	
0.49	−0.49	−0.43	−0.57
−0.52	0.50	0.55	0.42

However, the agreement in changes is not complete—the changes in long run length emphasis start before the changes in the eigenvalues. This may be due to a change in the covariance between vectors earlier than the change in covariance within vectors. A component analysis could also be done to study this possibility.

Use as a Feature

The components seem to be useful for distinguishing fatty and normal liver, but given the directionality of ultrasound images, rotation in invariance is an issue. It might be necessary to determine normal components at different liver orientations, determine the orientation of test liver ROIs, and then apply the closest oriented set of component transformations to achieve good results.

Error Metrics for Evaluating Compression Techniques

The global, first order error metric, NMSE, is not an adequate measure for evaluating the effect of compression techniques on medical images. The error in medically significant ROIs could be much higher than the overall error. Additionally, components of the covariance between neighboring pixels could be significantly distorted even though the first order error is small relative to the first order variance. The following alternative method of measuring error is suggested: 1) divide the image into small blocks (eg, 64 × 64); 2) calculate image vector principal components for each block; and 3) measure and report on percentage error in each component for each block.

REFERENCES

1. Krasner B, Lo SB, Garra B, et al: Effect of Vector Quantization on Ultrasound Tissue Characterization. SPIE Proceedings, Medical Imaging VI: Image Capture, Formatting and Display, 1653:360-373, 1992

2. Lo SB, Shen EL, Mun SK, et al: A method for splitting digital value in radiological image compression. Med Phys 18:939-946, 1991

3. Riskin EA, Lookabaugh T, Chou PA, et al: Variable rate vector quantization for medical image compression. IEEE Trans Med Imaging 9:290-298, 1990

4. Press WH, Flannery BP, Teukolsky SA, et al: Numerical Recipes. Cambridge, England, Cambridge, 1986, pp 335-380

5. Raeth U, Schlaps D, Limberg B, et al: Diagnostic Accuracy of Computerized B-Scan Texture Analysis and Conventional Ultrasonography in Diffuse Parenchymal and Malignant Liver Disease. J Clin Ultrasound 13:87-99, 1985

6. Insana MF, Wagner RF, Garra BS, et al: Pattern recognition methods for optimizing multivariate tissue signatures in diagnostic ultrasound. Ultrasonic Imaging 8:165-180, 1986