Characterization of Tissue from Ultrasound Images

Reza Momenan, Robert F. Wagner, Murray H. Loew, Michael F. Insana, and Brian S. Garra

ABSTRACT: This paper reviews and evaluates the application of a procedure for classifying tissue types from unlabeled acoustic measurements using unsupervised analysis. Unsupervised learning techniques are applied to the problems of detecting tumors within an organ and of discriminating between tissue types of two neighboring organs such as the liver and the kidney.

Introduction

Ultrasonic tissue characterization aims at quantitative identification of tissue type and its pathological state (i.e., healthy or diseased) from the echo signal of acoustic pulses transmitted into the body. Over the past 10 years, several investigators have proposed multiparameter approaches to this problem, in which acoustic properties of a tissue are extracted from radio-frequency (rf) echo signals [1] and B-scan images (two-dimensional cross sections, with reflectivity represented as intensity) [2]. Acoustic measurements that describe the structure of soft tissues are potentially very useful tissue signatures since disruption of the normal tissue architecture is one feature of diffuse and focal disease processes of many organs such as the liver. Second-order statistical properties extracted from speckle in clinical B-scans [4] have been shown to discriminate between normal liver and chronic active hepatitis [5]. Application of statistical pattern recognition methods to these second-order statistics has demonstrated their power to distinguish not only between normal and diseased tissue but also among different diseases [6].

A supervised classifier is designed under the guidance of a training set that consists of data features with known class membership. However, there exist many classification tasks in which sufficiently large training sets of labeled data, i.e., patient data with known physiological conditions, are not available or are too costly to acquire in sufficient detail. In addition, the diagnoses used to label data are often made from incomplete information and require subjective interpretation [7]. Obviously, any supervised classifier is only as reliable as its training set.

One important task in diagnostic ultrasound is the detection of signal changes in the data for a given patient. For example, lesion detection requires the identification of relative signal changes, e.g., identifying a lesion amidst the surrounding tissues, and may not be concerned with absolute signal changes, as would be used to separate normal from diseased tissue. Tissue characterization using unsupervised techniques such as clustering or distance measurement is an attempt to call attention to these subtle changes in relative signal strength [8].

Unsupervised pattern recognition techniques are an easily implemented approach to classification data. The purpose of this paper is to examine potential applications of unsupervised techniques to ultrasound image segmentation and tissue characterization of tumors and focal diseases. In particular, it is shown that the acoustic measurements used in this investigation naturally can separate tissue classes (normal versus abnormal, and disease A versus disease B). Furthermore, the possibility of detecting focal diseases (tumors) using unsupervised pattern recognition techniques is discussed.

Parameter (Feature) Extraction

As mentioned previously, in applying pattern recognition techniques to the tissue characterization problem, measurements are made of several acoustic properties of echo signals. In the investigation, four parameters of a tissue are estimated from statistical properties of rf echo signals [4], [5], [9], [10]. In determining these parameters, it is assumed that either the principal sources of ultrasonic scattering in the tissue are randomly positioned or the tissue contains both randomly positioned and regularly positioned (quasiperiodic) scatterers. The backscatter signal from many small, randomly positioned scatterers can be described as the sum of statistically independent scattering contributions (a random phasor sum) and is characterized by the diffuse incoherent backscatter intensity \( I_d \). The presence of regularly positioned specular scatterers, with mean backscatter intensity \( I_s \), adds a second phasor to the complex echo signal. This phasor is nearly constant with position if all of the specular scatterers have similar scattering properties and if the average spacing between specular scatterers is less than the resolution of the system. Otherwise, the magnitude of the second phasor will oscillate, and a corresponding mean intensity \( I \) and variance \( \sigma^2 \) can be defined [9].

Three scattering features that describe the tissue structure may be obtained directly from the statistics of the squared envelope (intensity data) rather than from the envelope (image data). These are \( d \), the average distance between regularly positioned specular scatterers; \( r = l / d \), the ratio of specular to diffuse backscatter intensities; and \( \sigma^2 = \sigma_0^2 \), a measure of the variability in the specular component normalized by the diffuse contribution. The fourth feature is the slope of the attenuation coefficient, \( \alpha_0 \), as a function of frequency. A calibrated time-domain method has been used to estimate \( \alpha_0 \) [11].

Techniques

Clustering

KSODATA is an iterative algorithm based on the nearest-neighbor criterion [in which a sample is labeled according to the label(s) of its nearest neighbor(s)], which combines essential features of the well-known K-Means [12] and ISODATA [13] techniques. This technique, developed by the authors, is called KSODATA, since it is a combination of K-Means and ISODATA. Figure 1 shows the flowchart of this technique. Cluster initialization is a heuristic method that divides the data set into \( K \) mutually exclusive partitions with equal numbers of samples. Initial cluster centers are determined from partition

samples that are farthest from samples in other partitions. In this way, the entire sample space is used to find the initial centers.

The procedure then continues in accord with the K-Means approach as follows: The algorithm checks whether two cluster centers are closer than a threshold established from the standard deviation of sample members about the sample-space mean. If the measurement is below the threshold, the two clusters are combined and a search is begun to find the cluster with the largest dispersion. This cluster is then split into two clusters with centers separated by two standard deviations in the direction of the largest standard deviation. The process continues until the clusters no longer change or until a performance criterion, e.g., cluster average similarity measure [14], [15], is optimized.

**Error Estimate**

A goal of this investigation is to determine whether there is any diagnostic information embedded in the clustering arrangement. For this reason, clinical data are clustered and the results are compared with those of a supervised method, the Bayes linear classifier [16]. The comparison measure is classification error, given by the following, where $n_i$ is the number of misclassified samples out of a total of $M_i$ from the $i$th cluster and $P(o_i)$ is the class prevalence.

$$E = \sum_{i=1}^{k} P(o_i) n_i / M_i$$

Note that $k$, the number of classes, does not necessarily equal $K$, the number of clusters. In fact, the clustering method should ensure that $k \leq K$ [15].

**Chi-Square Hypothesis Testing**

As previously mentioned, another goal of the application of unsupervised techniques is detection of focal disease within surrounding normal tissue. In this case, the interest is in detecting relative signal changes in the region of interest.

For this purpose, the region of interest is divided into several subregions and the four parameters $d, r, \sigma_r^2$, and $\alpha_0$ are calculated for each subregion. Then, using a hypothesis-testing approach, it is determined whether the subregions are of the same population (tissue type). In these investigations, the chi-square hypothesis-testing approach was used. Chi-square hypothesis testing, assuming independence of the features, is defined as the weighted sum of the squares [17]:

$$\chi^2(n) = (\bar{d}_a - \bar{d}_b)^2/\sigma_a^2 + (r_a - r_b)^2/\sigma_r^2 + \ldots$$

Here subscript $A$ denotes a reference subregion and $B$ denotes a test subregion. Using this hypothesis testing, the homogeneity or heterogeneity of these subregions can be determined with respect to the reference subregion, $A$, which leads to detection of tumors and focal lesions.

**Results and Discussion**

We have examined the applicability of unsupervised techniques, particularly KSO-DATA, for classifying human tissue types from acoustic measurements. We have compared the discriminability of the assigned clusters with a linear Bayes classifier [6].

The results are tabulated in the Table and some of them are illustrated in Figs. 2-4. In these comparisons, we have used 32 normal patients, 48 patients with chronic active hepatitis, 42 cases of Gaucher’s disease, and 11 cases of normal kidney.

The error estimates for the unsupervised method were fairly low and similar to the results of the supervised method. Although a more accurate discrimination might have been expected from the supervised technique (assuming the labeled data used to design the classifier are accurate), the assignment of the clusters to the classes uses the same a priori information as the classifier and, therefore, we are not surprised that both methods perform similarly.

Figures 5 and 6(a) are two $256 \times 256$ simulated images resembling ultrasound B-scans [18]. Figure 5 contains a background with random scatterers and no structure, the upper left-hand corner with random scatterers of higher intensity than that of background, and a structured region in the center of the image. The structured region has scatter-like structures with spatial periodicity of 6 pixels. Figure 6(a) includes the same non-structured background and the upper left-hand quarter with structures of 16-pixel spatial periodicity.
### Table

Comparison of Supervised Classifier and Unsupervised Clustering Technique

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Data Set</th>
<th>Features</th>
<th>Estimated Error (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Normal Liver vs.</td>
<td>$d, \sigma'$</td>
<td>16.5, 16.5</td>
</tr>
<tr>
<td>2</td>
<td>Normal Liver vs.</td>
<td>$r, \alpha_0$</td>
<td>31.6, 32.9</td>
</tr>
<tr>
<td>3</td>
<td>Normal Liver vs.</td>
<td>$\tilde{d}, r, \sigma'$</td>
<td>21.9, 16.9</td>
</tr>
<tr>
<td>4</td>
<td>Normal Liver vs.</td>
<td>$\tilde{d}, r, \sigma', \alpha_0$</td>
<td>32.9, 16.5</td>
</tr>
<tr>
<td>5</td>
<td>Normal Liver vs.</td>
<td>$\tilde{d}, \sigma'$</td>
<td>46.3, 28.1</td>
</tr>
<tr>
<td>6</td>
<td>Normal Liver vs.</td>
<td>$r, \tilde{d}$</td>
<td>9.52, 7.14</td>
</tr>
<tr>
<td></td>
<td>Normal Liver vs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>versus Gaucher's disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal Liver vs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal Kidney</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using these two images to detect localized tissue types, we find that the average scatterer distance, $d$, shows the most discriminability. This is also intuitively reasonable since we are differentiating between a background with no structure and an embedded tissue-like image with periodic structures. However, in Fig. 5, the upper left-hand quarter has been detected for its high value of the $r$ parameter, as expected. Figure 6(b) shows the color-coded image of Fig. 6(a) based on parameter $d$.

Our investigation, presently in progress, with respect to the detection of localized (focal) tissue type (i.e., lesions and tumors) and diffuse diseases, indicates the following:

1. In detecting tumors of various types, different weights may be given to each component of Eq. (2)—e.g., higher weight to $d$ and $\sigma'$ components in detecting diffuse diseases (e.g., chronic active hepatitis)—and higher weight to the $\alpha_0$ component in differentiating fatty versus nonfatty tissue. This is because different diseases affect the acoustic properties of tissue differently. Different weighting factors are under study.

2. As discussed earlier, our results indicate the ability of these features to discriminate between normal and diseased tissues. Figure 7(a) shows an image of a normal liver and a part of a normal kidney (arrow). Figure 7(b) shows the color-coded version of the same image. The

![Fig. 2](image1.png)

![Fig. 3](image2.png)

![Fig. 4](image3.png)

![Fig. 5](image4.png)
played. The results of the Table suggest that conventional imaging may be detected and displayed. Where subtle tissue embedded in every 16 pixels of upper left corner; (b) color-coded parametric image of a based on d.

colors are based on the result of chi-square testing. It is clear that the kidney and the diaphragm have been distinguished from the surrounding normal liver tissue.

(3) The results of our investigation with several simulated images suggest that the dimensions of subregions should be an order of magnitude greater than that of the spacing of the structure we are trying to detect. While this condition is easy to meet in large organs such as the liver, it dictates a minimum size for tumors to be detectable.

Conclusions

A prime application of cluster analysis is in image segmentation, where subtle tissue differences may be missed in conventional imaging may be detected and displayed. The results of the Table suggest that different tissue types naturally cluster in this four-dimensional feature space and, therefore, clustering may contribute in detection tasks in which relative tissue differences are important.

Our preliminary data suggest discriminability among the various normal tissue types as well as between normal and diseased tissues and among diseased tissues. The discrimination error of the clustering technique was comparable to that found with a linear classifier.

The application of chi-square hypothesis testing in discriminating subregions of different acoustic properties may be a powerful tool. While further clinical work is proceeding, detection of tumors and focal disease seems to be an easy task for this technique.

These results suggest the usefulness of unsupervised techniques in subclassification of unlabeled data and detection of homogeneity and heterogeneity of clinical ultrasound images and, hence, in tumor detection. As we have seen, the results are strongly dependent on the choice of subset of the feature space for the discrimination task at hand. Collectively, the four features $\alpha$, $r$, $\sigma^2$, and $\alpha_0$ display a powerful discriminability in the tissue characterization problem.

We are currently investigating the applicability of these unsupervised techniques to the display of color-coded B-scans. In these images, suspected disease regions (i.e., tumors) and regions of different acoustic properties are colored differently. This will bring the suspected area to the user’s attention for more detailed investigation of the physiological state of the tissue. Therefore, color-coded images of this type may be powerful visual aids to diagnosticians in the near future.

References

ternpcrienccd in
Applications 01 \igiial
iiig
255-3433. bridge. MA 02139: Phone: (617) 253-2216, 3930.

C. Richard Johnson. Jr.. School of Electrical
eering. Cornell Univcrsit)
Hall. Ithaca. NY 14853-5301: Phone: (607) chuwts Institute of Technology. Cam-
lanta. GA

and

vision
recog
ition. ancl linage
Washington. D
Catholic Uiii
ersit),
the B.S.E.E.
and Ph.D. candidate mapring
1976 to 1978. Senior Engineer at Adaptron- and

From 1972 to 1976. he

and

Dr. Gama's research

from Villanova Univer-

rently. At BRH

sion and image display: and statistical pattern rec-

at the

R. Mommen-

the B.S. degree in

Pennsylvania, in 1938. He

received the B.S.E.E. de-

Pennsylvania, in 1959. the M.A. degree

Physics. Technion. Israel

nuclear physics from

1969. respectively. From 1970 to 1972. he

held a Postdoctoral Fellowship at Ohio University

and

and

in

received the B.S. degree

1954. the M.A. degree

in

from Drexel Institute

nuclear physics from

1965. and the M.S.E.E.

from 1976

and

Computer Science at The George Washington

University, Washington, D.C., where he is now a
Professor. His teaching and research interests are
in the areas of pattern recognition, computer
vision, and medical imaging and signal processing.
He is also co-founder of the Institute for Med-
ical Imaging and Image Analysis at the university.
In 1986-1987, he was a Lady Davis Visiting As-


Robert F. Wagner was

born in Philadelphia,
Pennsylvania, in 1938. He

received the B.S. degree in

electrical engineering from

Villanova University in

1959. the M.A. degree

from Augustinian College

in Washington, D.C., in

1965. and the M.S. in

physics and the Ph.D. in

nuclear physics from

Catholic University, Washington, D.C., in 1965

and 1969, respectively. From 1970 to 1972. he

held a Postdoctoral Fellowship at Ohio University

at Athens, Ohio, taught modern physics, and did

research on photo- and electromolecular

interactions.

In 1972. he joined the laboratories of the Bureau

of Radiological Health (BPH) in Rockville. Mary-

land, where he has been Chief of the Radiologic

Imaging Section for 12 years. At BRH (now

CDRH. Center for Devices and Radiological

Health) he has

published by the Amer-

ican Association of Physicians in Medicine.

Murray H. Loew received the B.S.E.E.
degree from Drexel Institute

of Technology, Philadel-
phia. Pennsylvania, in 1965, and the M.S.E.E.
and Ph.D. degrees from Purdue University,

Lafayette, Indiana, in 1967

and 1972, respectively.

From 1972 to 1976, he

was a Member of the Technical Staff at Ana-

tlytic Services, Inc., Falls Church, Virginia, and

from 1976 to 1978, Senior Engineer at Adaptron-

ics, Inc., McLean, Virginia. In 1978, he joined

the Department of Electrical Engineering and

Computer Science at The George Washington

University, Washington, D.C., where he is now a
Professor. His teaching and research interests are
in the areas of pattern recognition, computer
vision, and medical imaging and signal processing.
He is also co-founder of the Institute for Med-
ical Imaging and Image Analysis at the university.
In 1986-1987, he was a Lady Davis Visiting As-


53