Pattern recognition techniques are currently being applied to many signal interpretation problems in nondestructive testing. Simulearning technology combines various aspects of wave propagation analysis, pattern recognition philosophy, and signal processing theory in such a way as to outline procedures and establish guidelines for solving many problems in flaw classification. A portion of this paper will be used to present flaw classification problem statements and potential solution techniques along with simple data and analysis techniques.

Emphasis in the paper will be placed on a work description and analysis associated with a flaw classification problem of discriminating between ultrasonic signals that have been reflected from elliptical and circular side drilled electro discharge machined slots in a steel block. The flaw types used in this experiment are several elliptical holes with eccentricities, e, from 0.15 to 1.0. The signals are sampled at a 100 MHz rate and quantized with an 8 bit word length. The signal processing is performed on a PDP 11/05 minicomputer.

Items discussed in this paper also include aspects of computational efficiency, waveform averaging, adaptive quantization, and Wiener filtering to suppress the effects of measurement and quantization noise. A novel deconvolution procedure was considered for removing the effects of the transmission medium and transducers and to enhance the discrimination between the flaw types. Feature extraction and pattern classification techniques that were used include the Fischer linear discriminant function, a declustering algorithm, and nearest neighbor classifiers.

Results obtained thus far indicate that for minor diameter to major diameter ratios e in excess of 0.7, discrimination between elliptical and circular flaws is very difficult. For e less than 0.3, discrimination is easy. Consequently, the feature extraction and pattern classification techniques have been concentrated on e in the range 0.3 to 0.7 in order to establish the efficiency of the research protocol.

With the increasing importance of nuclear power plant inspection, pressure vessel inspection for the energy industry, and rail inspection for the transportation industry, there has become an urgent need for the development of reliable and precise flaw classification techniques. Emphasis has recently been placed on studying ultrasonic response variations as a function of flaw type, shape, size and orientation. Considerable attention is currently being placed on quantitative aspects of NDE, as illustrated by many of the research programs on scattering theories and flaw characterization work carried out by Rockwell International and the various sub-contractors in the ARPA/AIML NDE research program. Emphasis is being placed on obtaining an improved understanding of the physics and mechanics of wave interaction with a flaw.

The purpose of this paper is to review aspects of flaw classification work, but with the emphasis being placed on pattern recognition and signal processing, rather than detailed physics and mechanics. Physics and mechanics is used in a qualitative sense to improve data acquisition systems and to gain insight into potential signal processing and feature extraction techniques for solving critical problems in flaw classification. Frequency analysis has demonstrated some potential for solving problems of this type. The state of the art on this subject, however, is progressing very slowly because of the large number of parameters generally associated with flaw characteristics. Other transform signature techniques are also being studied, but it is becoming quite evident that computer search and analytical techniques are required because of the signature complexities and computational efficiency required to obtain satisfactory correlations and/or solution paths.

A brief review of two research papers is presented in the following paragraphs, followed by a sample problem of side drilled elliptical hole eccentricity classification. Concepts presented in the first two papers serve as background information in the development of the elliptical hole classification problem. The two papers are:

1) "Disk and Spherical Inclusion Classification Concepts" by J. Rose, Drexel University; Phil Mast, Naval Research Laboratory; and Phil Walker, Krautkramer-Branson, Inc. This paper was presented at the Fall Meeting of ASNT in Atlanta, Georgia, in 1975 and has been submitted to Materials Evaluation for publication.

2) "Flaw Classification Techniques in Ultrasonic Inspection" by J. Rose, Drexel University; L. Niklas, Krautkramer GMBH; P. Mast, Naval Research Laboratory. This paper is included in proceedings of the Eighth World Conference on Nondestructive Testing held in Cannes, France, September 1976.

The subject of simulearning is outlined next and is presented in references (1) and (2) above. Many signal interpretation procedures have been studied to date that rely on such simple data reduction techniques as peak amplitude analysis or arrival time analysis, or more sophisticated transform "signature" analysis. For many complex problems in ultrasonic inspection and flaw characterization, however, these simplified approaches to signal interpretation and classification are not adequate. The simulearning technique presented in this paper provides us with a procedure for obtaining complete experiences associated with data acquisition along with methods of complete analysis through signal processing in a computationally efficient fashion. Techniques are presented that
enable us to obtain a reasonable solution technique for material or flaw characterization, provided a solution is possible at all. The technique of simulearning is a hybrid concept integrating various aspects of analytical mechanics, wave propagation analysis, learning machine philosophies, mathematical pattern recognition and signal processing techniques, and finally, human judgment. A simulearner can best be described as a logic system activated by a parametric input that searches for classifier parameters for solving specific flaw, material, or system classification problems in a computationally efficient fashion. Parameters related to this technique are fed into a numerical computation scheme or model that generates data representative of many real world flaw characterization problems. Large numbers of data sets are obtained either analytically, experimentally, or generated by some combined analytical-experimental technique. The amplitude-time signatures of the simulated flaw situations are then subjected to a class of fast linear (tensor) transforms, such as Fourier, Mellin, etc. This set of data forms the domain of definition for non-linear maps, the range being a pattern space. For example, amplitudes at $N$ specified frequency coordinates of the Fourier spectrum may be used to generate a column vector or pattern with $N$ entries. The simulearner will generate patterns of this nature using a class of known useful characteristics, such as 6dB down points, the maximum amplitude over an interval, etc. The simulearner will sequence through these "pure" patterns and evaluate each derived set as to its separability into classes and the relevancy of these class divisions to the particular problem at hand. The simulearner will then investigate the utility of hybrid patterns, that is, column vectors whose entries are disjoint in the sense that each is obtained from a different linear transform, followed by the non-linear feature extraction process mentioned above.

A proposed simulearning computation procedure is shown in Table I. Variations on the proposed scheme certainly exist. Discussions, definitions, and interpretations of the various items contained in the chart could be carried out with both enthusiasm and controversy. The chart does, however, provide us with one logical approach of solving many complex problems in flaw classification. Only portions of the chart are required for obtaining solutions to some problems. On the other hand, careful attention to every block may not solve some of the more complex problems in flaw classification.

**Table 1 - A Proposed Simulearning Computation Procedure**

**Disk and Spherical Inclusion Classification Concepts**

Potential applications of pattern recognition and simulearning in flaw classification and ultrasonic inspection analysis are reviewed in (1). The sample problem of disk and spherical inclusion classification is reviewed. Analytical procedures for generating ultrasonic response function data sets for the spherical and disk inclusion in a fluid are presented along with amplitude time profiles, selected transform signatures, and finally the resulting index of performance values for the simulearning computation.

A computer program is reviewed in (1) to calculate the ultrasonic field pressure variations in a fluid resulting from ultrasonic wave interactions with arbitrarily shaped air type flaws in the fluid. The flaw is divided into segments of approximately equal surface area from which a spherical wave is propagated from each segment on the surface of the flaw. Although the problem of studying ultrasonic wave reflections from an air-filled inclusion in a fluid is not totally realistic, the data sets generated from this kind of problem allows us to evaluate qualitatively the concepts of simulearning, feature extraction from a data set, pattern recognition details, etc.
Data sets considered in the study consisted of two amplitude-time profiles, one for normal wave scattering at the sending transducer and one for normal wave scattering received at the receiving transducer located at a position \( x_2 \).

Ultrasonic pulse echo signals representing response echoes from either spherical or disk type flaws were generated as a series of sample data sets. Certain features of the ultrasonic response functions were chosen to be stored in the simulearned. Then, test data representing spherical and disk flaws of unknown size were compared with the prototypes in the simulearmer in order to make a flaw classification prediction. Features selected for this comparison were Fourier transform amplitude, phase angle \( \phi \), Laplace transform magnitude, and Mellin transform magnitude. The comparison of test data with the prototype data was based on the minimum distance classification technique.

In the first problem of sphere and disk classification, all sphere training points clustered nicely and were separated easily from the disk training points, regardless of the transform selected for the study. In this particular problem, there was no need for more sophisticated analysis utilizing either decision surface adjustment techniques that force the data into the correct class or the selection of some other transform. The straightforward procedure of transform selection and prototype selection based on average training data produces for us an index of performance in all cases of 100%.

Let us now consider the problem of disk size classification. This problem illustrates the values of transform selection in that a 100% index of performance value is obtained for only 2 of the 4 transforms studied. The Fourier phase angle and Laplace transform approaches produce index of performance values which were not acceptable.

Let us consider, for example, the problem of sphere size classification, a summary of which is outlined in Table 2. In this particular case, no index of performance value was 100%. The values shown, however, do indicate the best possible values for the case of the transform selected in combination with the minimum distance classifier since of 4 training sets considered in the study, 4 prototype points were considered, therefore forcing all of the training set information to appear in the proper class. The index of performance results shown in the table could perhaps be improved by considering some other transform or discriminant function type, or possibly by classification adjustment if engineering knowledge of the subject permits such action. As an example, if we were to combine classes 1 and 2 in Table 2, as a result of some engineering study, the index of performance would be 100% for the Fourier transform amplitude situations.

<table>
<thead>
<tr>
<th>Class</th>
<th>Prototype Used</th>
<th>Test Set Classification Technique</th>
<th>Laplace Phase Angle</th>
<th>Fourier Phase Angle</th>
<th>Index of Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>10 to 20 spheres</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.03, -0.01, -0.00</td>
<td>-0.39</td>
</tr>
<tr>
<td>Class 2</td>
<td>0.1 to 0.8 spheres</td>
<td>-0.39</td>
<td>-0.39</td>
<td>-0.03, -0.01, -0.00</td>
<td>-0.39</td>
</tr>
<tr>
<td>Class 3</td>
<td>0.5 to 0.7 spheres</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.03, -0.01, -0.00</td>
<td>-0.39</td>
</tr>
<tr>
<td>Class 4</td>
<td>27 to 30 spheres</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.03, -0.01, -0.00</td>
<td>-0.39</td>
</tr>
<tr>
<td>Index of</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.03, -0.01, -0.00</td>
<td>-0.39</td>
<td></td>
</tr>
</tbody>
</table>

Although specific details of the complete simulearning computation procedure have not been carried out in (1), the subject of transform selection and its utility in varying the index of performance has been illustrated quite well. The concept of classification selection and adjustment is also illustrated quite well.

Flaw Classification Techniques in Ultrasonic Inspection

The work reported in Ref. (1) is theoretical in nature. In order to consider the more realistic experimental problem with such parameters as "noise" components, instrumentation variations, transducer effects, etc., it was decided to conduct a 10 flaw sorting study, the goal of the study being to separate the 10 test flaws into as many groups as possible. Flaw types considered in the study are presented in Table 3. The flaws were all manufactured by electrode discharge machining in steel blocks. Characteristics of the various flaws are presented in Table 4. The data acquisition technique considered in this study is illustrated in Fig. 1. Transducer 1 was considered as the sending transducer to the flaw machined in the test specimen approximately 25mm from the sending transducer. An angle beam transducer was used in position 2 to receive scattered normal and shear waves. The transducers used in this study were of 5 MHz center frequency with a 6dB down bandwidth of 3 MHz. The angle beam receiving transducer was rated at 45° in steel. The data was recorded with a Biomation 8100 analog to digital converter and stored in a PDP 11/05 minicomputer. The data points were stored along with the corresponding Fourier transform and phase angle.
TABLE 3 - ELECTRO-DISCHARGE MACHINED SIDE DRILLED FLAW TYPES

<table>
<thead>
<tr>
<th>Type</th>
<th>Diagram</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td><img src="image1" alt="Diagram" /></td>
<td>Normal Sharp Edge Defect</td>
</tr>
<tr>
<td>Type 2</td>
<td><img src="image2" alt="Diagram" /></td>
<td>Inclined Sharp Edge Defect</td>
</tr>
<tr>
<td>Type 3</td>
<td><img src="image3" alt="Diagram" /></td>
<td>Cylindrical Defect</td>
</tr>
<tr>
<td>Type 4</td>
<td><img src="image4" alt="Diagram" /></td>
<td>Rectangular Edge Defect</td>
</tr>
<tr>
<td>Type 5</td>
<td><img src="image5" alt="Diagram" /></td>
<td>Elliptical Defect</td>
</tr>
</tbody>
</table>

TABLE 4 - TEST SPECIMEN FLAW CHARACTERISTICS

<table>
<thead>
<tr>
<th>FLAW NO.</th>
<th>TYPE</th>
<th>L₁ (mm.)</th>
<th>L₂ (mm.)</th>
<th>θ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>4.76</td>
<td>-</td>
<td>0°</td>
</tr>
<tr>
<td>2.</td>
<td>1</td>
<td>3.18</td>
<td>-</td>
<td>0°</td>
</tr>
<tr>
<td>3.</td>
<td>4</td>
<td>3.18</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>1</td>
<td>2.38</td>
<td>-</td>
<td>0°</td>
</tr>
<tr>
<td>5.</td>
<td>2</td>
<td>3.18</td>
<td>-</td>
<td>30°</td>
</tr>
<tr>
<td>6.</td>
<td>2</td>
<td>3.18</td>
<td>-</td>
<td>45°</td>
</tr>
<tr>
<td>7.</td>
<td>3</td>
<td>4.76</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>3</td>
<td>3.18</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>3</td>
<td>1.59</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>5</td>
<td>3.18</td>
<td>2.38</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Data acquisition technique.

Combinations of the parameters A₁, A₂, T₁, A₃, A₄, and T₂ were studied in detail. After several attempts at classification, it was found that the 10 flaws could be separated nicely by considering the parameters defined below:

- T₁ = principal separation and the sharp edge surface area parameter
- A₁/A₂ = sharp edge defect inclination parameter
- A₄/A₃ = sharpness parameter
- A₄/A₁ = cylindrical surface area parameter

The final sorting procedure for this problem is illustrated in Table 5.
Additional work is currently being carried out that examines various flaw cluster groups, etc. A problem encountered quite early, however, in the new work was that of classifying various elliptical shapes. The purpose of the sample problem, therefore, presented in the next section, is to study the elliptical eccentricity classification problem in detail.

Elliptical Eccentricity Classification Study

The subject of elliptical cavity eccentricity classification is reviewed in this section. As indicated earlier, emphasis will be placed on noise aspects of the classification problem.

The signal processing techniques described below were employed to reduce random variations, or noise in the received ultrasonic signal. Although these techniques are necessary for complex operations such as deconvolution, they can also greatly improve the performance of simple classification algorithms. In preliminary studies, the discrimination between signals reflected from circular flaws versus those reflected from elliptical flaws has been poor, especially for elliptical flaws with eccentricities greater than 0.5. Three elliptical flaws and a circular flaw, all with the same major axis diameter, were assembled as shown in Fig. 2. Typical signals arising from each flaw type are shown in Fig. 3, utilizing data acquisition pro-
The shear wave amplitude, normalized by the longitudinal wave amplitude, appears to increase with decreasing eccentricity. This feature, termed the shear strength, performed unsatisfactorily as evidenced by the overlapping probability density functions in Fig. 4. The density function estimates are based on approximately 120 points for both $\epsilon = 1$ and $\epsilon = 0.75$, and approximately 50 points for each of the other two classes. The probability of error in discriminating between $\epsilon = 1$ and $\epsilon = 0.75$ is approximately 30%, between $\epsilon = 0.75$ and $\epsilon = 0.5$ approximately 3%, and negligible between $\epsilon = 0.5$ and $\epsilon = 0.3$. By employing noise reduction procedures, this performance can be greatly improved as indicated in the following sections.

Figure 4. Probability density functions of shear strength before signal processing.

The primary sources of noise contaminating the ultrasonic return pulse are the placement noise, the measurement noise, and the quantization noise, as shown in Fig. 5. The placement noise includes the effects of the material, the coupling, and varying transducer placement. This noise term affects the signal in a complex manner, and was minimized by positioning both the sending and receiving transducers to maximize the energy in the reflected acoustic signal.

Measurement noise, sometimes called thermal noise, arises mainly from the wideband amplifiers used to amplify the ultrasonic signals. One way to suppress measurement noise effects is to increase the signal energy. Since the received signal is repetitive, a simple averaging procedure is also possible. The variance of the measurement noise term using the latter method is reduced by a factor inversely proportional to the number of waveforms averaged.

However, in this problem the signal plus measurement noise is quantitized by an 8 bit analog to digital (A/D) converter. At each sampling instant the signal is assigned to one of 28 or 256 quantum levels. This step can be modeled as the addition of a quantization noise term whose probability density function is uniform. With the biometry 8100 A/D converter used in this project, the quantization noise has a mean of $+Q/2$ for positive signals and $-Q/2$ for negative signals, where $Q$ is the quantum step size. The variance is proportional to $Q^2$. The effect of the quantization can be neglected when the signal is large compared to $Q$.

However, normally the range of the A/D converter is set to accommodate the largest signal encountered in the received signal. The low level signals therefore are severely degraded by the quantization. In particular, the shear signal may only reach the 8th or 10th quantum levels. The result is that the shear strength can only be measured approximately due to the quantization noise.

The obvious way to decrease the quantization effects is to increase the number of quantum levels. However, since this option wasn't available, an alternate scheme, termed adaptive quantization, was implemented. The signal is stored in the computer in the usual manner. The ultrasonic signal's amplitude is next increased by a fixed amount. Some portions of the signal are now clipped, but these can be detected since they reside in either the lowest or highest quantum levels. The remaining sample values, which now span a greater number of quantum levels, can be rescaled to their correct values in the computer. Thus the quantum step size is effectively reduced for the lower amplitude signals, improving the signal to quantization noise ratio.

To test these signal processing procedures, each signal was stored in the computer, amplified by 20 dB and rescaled. This process was repeated 6 times, and the results averaged. An example of a signal before and after processing is shown in Fig. 6.

Figure 5. Primary sources of noise.
The shear strength parameter measured from the processed waveform now discriminates between the various flaw classes as shown in Fig. 7. The estimated probability of error for approximately 20 points per class is about 5% for discriminating between the circle and the ellipse with $e = 0.75$, 2% between $e = 0.75$ and $e = 0.5$, and negligible between $e = 0.5$ and $e = 0.3$. The main source of error is due to careless transducer placement and not quantization or measurement noise.

Further signal processing may be necessary for more complex feature extraction. For example, the averaged and adaptively quantized signal still contains high frequency components due to noise. Some type of low pass filtering is required, for example, if deconvolution is attempted, since this emphasizes the high frequencies. Preliminary tests indicate that the waveform averaging and adaptive quantization described in this paper, followed by filtering out all frequency components above 15 MHz, provides a signal whose major variation is due solely to placement noise. The low pass filtering had little effect on simple features such as shear strength, and, hence, wasn't incorporated into the present test.

Further improvements in performance can be effected by improving the classification algorithm. For example, additional features can be used to increase the reliability of the classifier, or to permit the assignment of waveforms into classes other than circular or elliptical.

One approach to incorporate additional features is to implement a Bayes decision rule. The Bayes rule is the optimal decision strategy in the sense of minimizing the probability of error. Also, the Bayes classifier can be used to evaluate feature sets to determine which features are needed for discrimination and which can be discarded. Although the Bayes approach requires knowledge of the multivariate probability density functions for each pattern class, these can be estimated from test waveforms.

Figure 6. Typical signal before and after processing.

Figure 7. Probability density functions of shear strength after signal processing.
In conclusion, it has been shown how a simple signal processing scheme comprised of waveform averaging and adaptive quantization can improve the performance of a pattern classification system.

Acknowledgement

We would like to thank Krautkramer Branson Inc. for their motivation and minicomputer support of various phases of this program of study.

DISCUSSION

DR. PAPADAKIS: Are there any questions?

MR. PAT RYAN (DOT): Could you select the quantizationized problem with logarithmic compression before quantizing or would that louse something else up?

DR. CARSON: I don't know for sure.

DR. TIEMANN (General Electric): I know the answer to that. The problem is that the wave form crosses zero and so you can't really, and it goes negative; so, you can't take logarithms.

DR. PAPADAKIS: Any others?

DR. SY FRIEDMAN (Naval Ship R and D Center): The word shear strength in the presentation - I'm just wondering how your measurement related to shear strength?

DR. CARSON: This is the shear wave--reflected shear wave.

DR. FRIEDMAN: Oh, not shear strength but reflected shear wave. Thank you.

DR. PAPADAKIS: Good. Thank you very much.