

Pipeline Flaw Detection with Wavelet Packets and Gas

Stephen W. Kercel^{*a}, Raymond W. Tucker, Jr.^b, Venugopal K. Varma^b
Oak Ridge National Laboratory

^aAddress correspondence to: 2 Brian Drive, Brunswick ME 04011

^bP.O. Box 2008, Oak Ridge TN 37831

ABSTRACT

This paper is concerned with the detection of physical flaws on pipe walls in gas pipelines. The sensor technology is EMAT (electromagnetic acoustic transducer), a non-contact ultrasonic technology. One EMAT is used as a transmitter, exciting an ultrasonic impulse into the pipe wall. Another EMAT located a few inches away from the first is used as a receiving transducer. This paper reports on the identification of flaw signatures in the receiver output. The first step in flaw characterization is to perform wavelet analysis of the signature. Being non-shift-invariant, an array of coefficients of a discrete wavelet transform of a signal is not directly suitable as a pattern recognition feature. However, comparing composite properties of the signal on different scales is useful, because the mode conversion caused by a flaw, changes the composite properties of the signal in wavelet space. For EMAT data, the useful information projects onto five mutually orthogonal wavelet scales. This paper reports on the use of a robust 17-dimensional feature vector that consistently distinguishes “flaw” signatures from “no-flaw” signatures in a substantial collection of experimental data.

Key words: wavelet, flaw detection, feature vector, pattern recognition, Mahalanobis distance, pipeline, EMAT, data compression, real-time processing, process monitoring

1. NON-DESTRUCTIVE TESTING OF GAS PIPELINES

Approximately 30% of the energy produced in the United States is derived from natural gas. Accordingly, the integrity of the natural gas supply system is of prime importance. Natural gas is supplied to users through a vast pipeline network that consists of over a million miles of pipeline.¹ Pipeline companies have an impressive safety record due to the proactive role they have taken in establishing standards and in the inspection of pipelines. Many of the pipelines are aging, and there is a great need for a way to identify cracks, corrosion, and other defects that can potentially cause problems.

A gas pipeline can fail due to many causes. Some of the most common failure modes are corrosion, pitting, stress corrosion cracks, seam weld cracks, dents, and other flaws induced by external impact from earth-moving equipment. Ideally, it would be desirable to detect all of the above flaw types with a single inspection technique. Unfortunately, there is no one inspection technique that is ideally suited to detect all of the possible flaw types. Hence, the gas industry uses a combination of techniques to ensure the safety margin for their operation. Probabilistic approaches have been used for estimating pipeline integrity.² Probabilistic methods attempt to predict safety using crack rate growth data, inspection frequency data, and the operating parameters of the pipe. Probabilistic methods, however, require valid statistical data on flaw rate occurrences and distributions to be of any real use. Hence, the need persists to be able to collect accurate data on the actual condition of the pipelines in service.

There are two main methods of testing the integrity of pipelines; destructive inspection and non-destructive inspection (NDI.) The destructive inspection procedure generally uses a hydrostatic technique to verify that the pipeline integrity is within the safety margin for operation. The procedure does not, however, locate defects that are just below the threshold of safety. In addition, destructive testing disrupts the pipeline’s normal operation. For this reason it is not the preferred method. Generally, such techniques are good for the initial inspection of pipelines before they are put into use. On the

* kercel1@suscom-maine.net; phone 1 207 729-5155; fax 1 207 729-6226

other hand, NDI techniques can detect developing flaws that can cause failures in the future. Thus, NDI provides a quantitative measure of the integrity of the pipeline as well as a measure of its current safety margin.

Some types of defects that can occur in pipelines are a serious threat to the safety of the pipeline. Stress corrosion cracking, for example, can occur at any time in the life of the pipeline, and it occurs under a broad range of field conditions. This type of defect is usually oriented along the axial (lengthwise) direction of the pipe. If not detected early, the cracks may grow and/or coalesce, eventually resulting in a leak or rupture of the pipe. Not all defects that develop in pipelines threaten the integrity of the pipeline. Benign, internal inclusions are common and do not pose a serious threat to the integrity of the pipeline. NDI systems are urgently needed that can (1) provide early detection of the more serious defects; (2) differentiate between the serious defects and benign inclusions; and (3) characterize the type and size of the defects for repair or replacement management.

A promising new sensing technology is the use of guided ultrasonic waves for detecting the defects that occur in pipelines. One major benefit of ultrasonic guided waves is their ability to travel longer distances in the pipe wall than the signals from other sensing technologies, and this enables them to inspect the structure line-by-line instead of point-by-point. As the guided waves travel through the pipe walls, they are affected by the features they encounter. In particular, the mode structure of the propagating ultrasonic wave is modified in specific ways depending on the specific flaw type encountered by the wave.³ The received signals contain much information regarding the nature and sizing of the features encountered. However, defect detection, classification, and sizing using guided ultrasonic waves is still a major challenge under investigation due to the complexity of the wave propagation characteristics. A major shortcoming of many of the previous investigations into the use of ultrasonic guided waves has been the lack of a suitable signal analysis method to adequately identify the mode conversion produced in the ultrasonic wave by the flaw.

This paper summarizes a preliminary investigation into the use of the discrete wavelet transform as the underlying methodology for flaw detection using ultrasonic guided waves. This paper describes the generation and detection of ultrasonic waves using electromagnetic acoustic transducers (EMATs). It then describes the experimental apparatus used to collect the data. It then discusses the characteristics of wavelets that make them particularly suitable for addressing the problem at hand, and then reports the results obtained to date using wavelet analysis. It concludes by suggesting a possible way to extend the analysis to include the classification of flaws by type.

2. ULTRASONIC WAVE GENERATION AND DETECTION USING EMATS

A common method for inducing ultrasonic guided waves in pipe walls is to use piezoelectric transducers. Since piezoelectric transducers require an incompressible liquid medium (oil, honey, etc.) to transmit the sound energy into the pipe wall, it is impractical for situations when the sensors are required to travel along the length of the pipe. A non-contact method for generating ultrasonic waves is required for this purpose. Since the EMAT does not have to be in physical contact with the pipe wall to induce ultrasonic waves, it is particularly well suited for pipeline inspection. The EMAT consists of a strong permanent magnet with a coil of wire located between the magnet and the pipe wall. The permanent magnet creates a static electromagnetic field that extends in a direction perpendicular to the surface of the pipe wall. When a gated, high-current sinusoidal pulse is passed through the EMAT coil, a dynamic field is created that is parallel to the pipe wall and perpendicular to the static field. The interaction of the static field and oscillating dynamic field generates magnetostrictive forces within the surface of the pipe. The forces acting on the surface can generate lamb, shear, and longitudinal ultrasonic waves in the pipe wall, depending on the configuration of the magnet and the characteristics of the current flow.⁴

These waves can propagate over relatively long distances in the pipe wall. Figure 1 illustrates the use of EMATs for pipeline inspection. In this configuration, one EMAT serves to generate the ultrasonic wave, while the other acts as the receiver (sensor). Figure 1 illustrates the two EMATs in a *transmission mode* configuration, where the ultrasonic wave travels from a source EMAT through the flaw on its way to a separate EMAT used as a sensor. It is also possible to perform measurements of this type in a *reflection mode* using a single EMAT as both source and sensor. In this mode of operation, the received signal is actually a partial reflection of the ultrasonic wave created by the interaction of the wave with the flaw. All measurements used in this work were obtained using the transmission mode configuration. By using

the through transmission mode whereby the distance between the EMATs is fixed, better correlation was obtained for the effect of flaws.

There have been many studies devoted to understanding the generation and propagation characteristics of the different ultrasonic wave types in media such as pipe walls.^{5,6,7,8,9} Shear horizontal (SH) waves have demonstrated their suitability for use in pipeline inspection because of their simpler dispersion characteristics and their tolerance of the damping caused by the protective coating present on the inner surface of many gas pipelines.⁵ It is also desirable to avoid generating multiple wave modes, since this greatly complicates both the interaction of the waves with the defects and the signal analysis required to interpret the results. Most of the tests for this study were performed using the n1 mode SH wave. The particular wave mode induced is dependent upon the frequency of the current signal that is passed through the EMAT conductor.

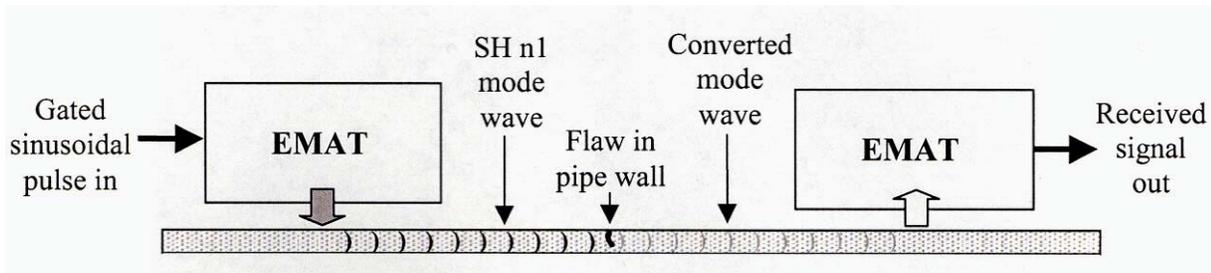


Figure 1. Illustration of through transmission mode ultrasonic wave generation using EMATs

3. EXPERIMENTAL SETUP

Six-foot lengths of 10-inch and 12-inch pipe were obtained for use in collecting the experimental data. To simulate the presence of cracks in the pipe wall, 0.25-inch-, 0.125-inch-, and 0.006-inch-wide grooves of varying depths were machined in both the axial and circumferential directions in the pipe walls. The simulated flaws were separated spatially from the end of the pipe and from each other by several times the spacing between the transmitter and receiver EMATs. This was done to allow collection of ultrasonic wave signatures containing a minimum number of reflections from pipe wall features other than the immediate flaw of interest. In addition, a short length of used 10-inch pipe containing stress corrosion cracking was obtained to allow the collection of data representative of that flaw type.

The EMAT transducers used were designed to fit the inside radius of both the 10-inch and 12-inch pipe. Hardware fixtures were designed and fabricated for allowing the EMATs to be positioned and traversed along the inside walls of the pipe. The mounting fixtures also had provisions to vary the gap between the pipe wall and the EMATs. The EMAT/fixture assembly was held in place by the strong permanent magnet within the EMAT. Figure 2 shows the transmitter/receiver EMAT pair in both the circumferential and axial geometries. The circumferentially-oriented EMAT pair shown on the left was used to detect flaws oriented in the axial direction, while the axially-oriented pair shown on the right was used to detect flaws and weld defects in the circumferential direction.

The EMAT excitation signal was provided by a programmable amplifier/receiver module capable of delivering a gated, high voltage RF pulse to the transmitter EMAT. The amplifier/receiver module was also equipped with a broadband receiver amplifier which was used to amplify the signal from the receiver EMAT. The resulting signal was lowpass filtered and sampled using an analog-to-digital converter. The received signals were recorded on disk and analyzed offline. Sampling rate was 5-million samples/second. For each trial, ten sets of either 2048 or 4096 samples were collected, and the ten sets were summed to average out noise. A 512-sample time window was applied to the summed signal to capture the burst to be wavelet analyzed.

Data were collected on sections of the pipe known to be free of flaws as well as on the sections containing the flaws. Samples of the signals conditioned as described above for three “flaw” observations and three “no-flaw” observations

are shown in Figure 3. These are bad cases compared to the experimental data as a whole, but the algorithm must correctly classify the bad cases as well as the easy cases. Figure 3 indicates something of the difficulty of the pattern recognition problem. To the eye, the “no-flaw” class does not look very different from the “flaw” class.

The signal is strongest when there is no flaw. The signature is reduced in amplitude when it encounters a flaw. Despite the fact that such changes are obvious in laboratory data collected under carefully controlled conditions, changes in the received signal amplitude in the time domain data are *not* a useful feature for identifying flaws in less controlled field conditions. In actual pipelines, there are other mechanisms besides a flaw that can lead to attenuation of the signal. The necessary features for flaw detection are those that depend on the mode conversion imposed by the flaw as observed in the shape of the received signal. Thus, the goal of the signal processing effort was to identify features that can differentiate between the “flaw” and “no-flaw” cases based on the shape rather than the amplitude of the received pulse.



Figure 2. EMAT Test Fixtures Used to Collect Experimental Data

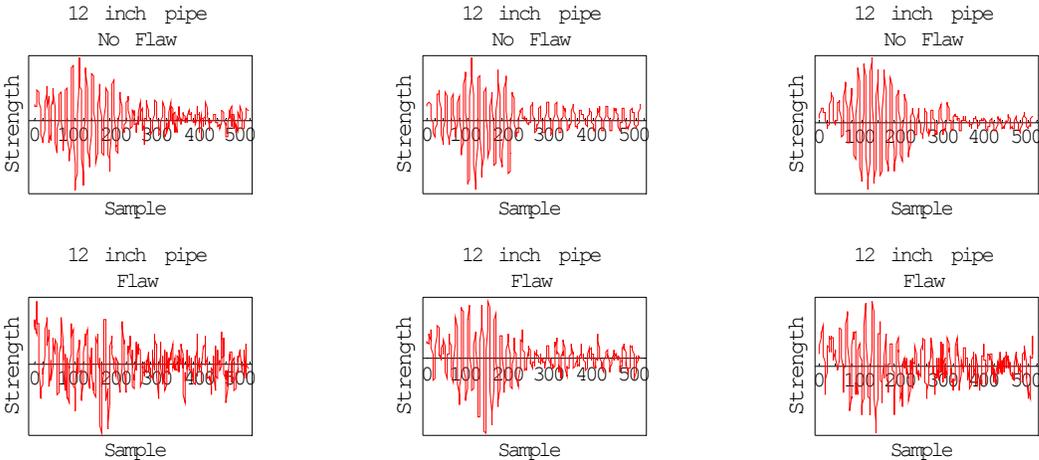


Figure 3. “Flaw” and “no-flaw” signatures

4. WAVELET-BASED CLASSIFICATION

A single wavelet-based feature vector consistently separates “flaw” from “no-flaw” samples for all the experimental data for both 10” pipe and 12” pipe. The wavelet function is a 58-coefficient Least Asymmetric wavelet. The plot of the wavelet function and the corresponding scaling function are indicated in Figure 4.

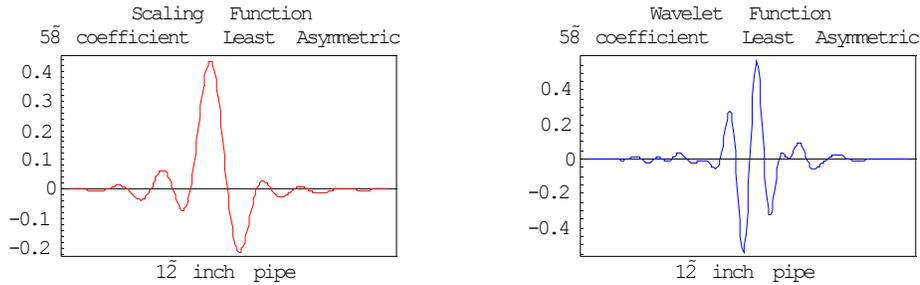


Figure 4. Wavelet Basis Function

For the 12-inch pipe data this particular function is the basis in that it provides the maximum compression of a set of eighteen samples of “no-flaw” data collected for a variety of axial and circumferential displacements between the transmit and receive EMATs. Compression is defined as the number of wavelet coefficients required to represent 50% of the signal energy and 90% of signal energy. The sum of the coefficients is taken across all eighteen samples. This compression computation was repeated for a wide range of Daubechies wavelets, Daubechies least asymmetric wavelets and Coiflets.

The power of this algorithm seems to be in the extremely tight compression obtained with this particular wavelet. The time domain data set consists of 512 samples, and the corresponding wavelet transform also has 512 samples. However, for the 12-inch pipe signals over 50% of the signal energy is contained in the *three* largest wavelet coefficients, and over 90% of the signal energy is contained in the *thirteen* largest wavelet coefficients. In other words, the wavelet is very similar to the waveform of the actual signal being observed.

The same wavelet was also the most effective compressor for the 10-inch pipe data. 26 samples of the 10 inch “no-flaw” data were analyzed using the same wavelet technique. For the 10-inch pipe signals over 50% of the signal energy is contained in the *two* largest wavelet coefficients, and over 90% of the signal energy is contained in the *seven* largest wavelet coefficients. It is remarkable that the same wavelet is the best compressor across all the data for both pipe sizes, various excitation frequencies, and assorted variations in the sensor configuration. This suggests that the experimental apparatus is collecting very consistent data, and that the wavelet basis is a robust choice. The wavelet is very similar to the waveform of the actual ultrasonic signal being observed.

The signal is a 512 sample list with the window centered on the approximate center of the peak of the signal of interest. This is typically in the range of the 350th to the 861st samples in the overall 2048 or 4096 sample data set. Each list of samples is obtained by summing the ten repetitions for each sample. The wavelet transform of each signal consists of five wavelet levels. The coarsest scale (256 samples) is noise and can be ignored. There is characterizing information on each of the other four scales.

The feature vector is 17-dimensional, with four features extracted from each scale plus an error feature. The rationale for the feature extraction is the idea that the transmitting EMAT launches a single mode into the pipe wall, and if no flaw is encountered, that single mode is all that arrives at the receiving EMAT. This single mode is relatively tightly concentrated in time and frequency.

If the signal encounters a flaw, the change in boundary conditions launches several new modes. These modes will have slightly different frequencies, dispersion (or chirp) rates, phases and propagation rates. Indications of these effects can be seen in the different wavelet scales.

One feature is an energy feature. This is computed by computing the energy of the signal on a given scale, normalized by the total energy in the signal. In the “no-flaw” case the single mode will have a higher concentration of energy on a dominant scale. In the “flaw” case, there will be several modes with frequency spectra spread out over a wider range. This is seen as a slightly smaller fraction of energy in the dominant scale, and larger fractions of energies on other scales. Four of the seventeen features are the fraction of energy on each of the four relevant scales.

Another feature is an entropy feature. This is computed by taking the Shannon entropy on each scale. In the “no-flaw” case the single mode travels at a particular rate and appears relatively compact in time. This compactness can be measured by the entropy. Smaller entropy implies a more compact signal. In the “flaw” case, there will be several modes each traveling at slightly different velocities. This is seen at the receive sensor as a longer time duration in the signal, and is indicated by a higher entropy than the “no-flaw” case. Four of the seventeen features are the entropies on each of the four relevant scales.

Another feature is a frequency feature. Due to the multiple modes in the “flaw signal” it will have a slightly shifted frequency spectrum compared to the “no-flaw” signal. For 12-inch pipe data, this is seen in Figure 5. Here, the red traces are the magnitude of the discrete Fourier Transform at wavelet level 2 for each of the 23 “no-flaw” signals, and the blue traces are the magnitude of the discrete Fourier Transform at wavelet level 2 for each of the 33 “flaw” signals. Four of the seventeen features are the peak frequency bins in the left-hand of the Fourier spectra of the wavelet coefficients on each of the four relevant wavelet scales.

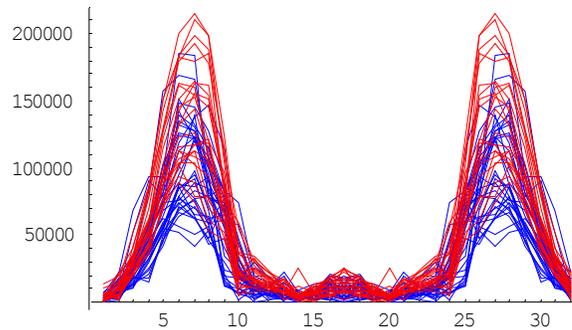


Figure 5. Magnitude of Fourier Spectra at Wavelet Level 2

Another feature is a phase feature. Phase data for 12-inch pipe are seen in Figure 6. Here, the red traces are the argument (phase angle) of the discrete Fourier Transform at wavelet level 2 for each of the 23 “no-flaw” signals, and the blue traces are the magnitude of the discrete Fourier Transform at wavelet level 2 for each of the 33 “flaw” signals. The different modes have slightly different dispersion (chirp) rates. The chirp causes a change in the phase shift of the received signal. Four of the seventeen features are the slope of the phase at the peak frequency bins in the left-hand of the Fourier spectra of the wavelet coefficients on each of the four relevant wavelet scales.

The sixteen-dimensional feature vector mentioned above was sufficient to completely separate the 12-inch pipe data. However, it did not reliably separate the 10-inch pipe data. One additional feature was required. The feature is an “error” feature. Figure 7 illustrates the idea. It shows the ensemble average for the “no flaw” signals in blue and the “flaw” signals in red averaged across the entire ensemble of “no flaw” and “flaw” samples respectively. Figure 7 shows these ensemble averages across the 12-inch pipe data at wavelet level 1.

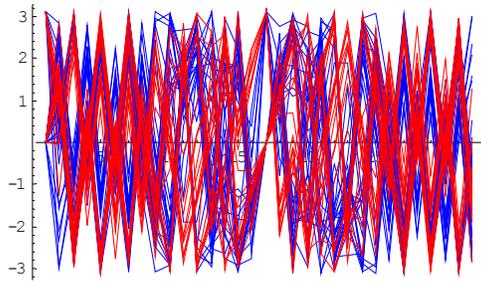


Figure 6. Phase of Fourier Spectra at Level 2

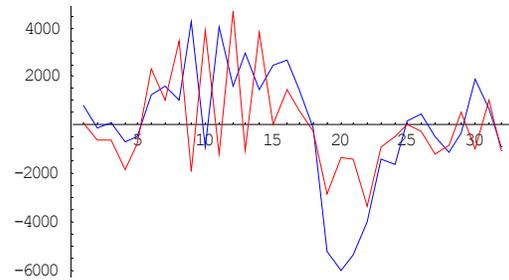


Figure 7. Average Signal at Level 1

The classification feature is based on the notion that a “no flaw” signal is expected to look more like an idealized exemplar (computed as the average “no flaw” signal) than does a “flaw” signal. In the actual classification algorithm, the exemplar is the ensemble average of all the “no flaw” signatures at wavelet level 2. Then, for each signal, the first step in the error calculation is the subtraction of the exemplar from the level 2 coefficients. Then the absolute value of each member of the resulting list is taken. Then the list is summed. Then, to mitigate conditioning problems with the covariance matrix in the Mahalanobis distance calculation, the sum is divided by 10^6 . This quotient is the “error,” a measure of how much the signal *does not* look like the exemplar. This constitutes the 17th feature.

Classification of each feature vector is determined by the Mahalanobis distance of the feature vector from the cluster center of the good samples. The cluster center is the mean value of the feature vectors of all the good samples. Computing the Mahalanobis distance from the cluster center of the bad samples turns out not to add any more useful information.

The 17-dimension classifier was applied to all the 12 inch pipe data, just as the 16-dimension classifier had been used previously. The reason for trying this was to assure that the 17th feature actually added information rather than noise to the 12 inch pipe classification problem. This proved to be the case. In every instance the 17-dimension classifier slightly outperformed the 16-dimension classifier.

Like the 16-dimension classifier, the 17-dimension classifier is remarkably robust, considering that some of the data were collected at 270 kHz and some at 250 kHz. It should also be noted that the data used were collected in three separate experiments conducted at different times during 2002 and 2003. This was necessary because limitations in the mounting hardware prevented collecting axial and circumferential data at the same time. Integrating all the data from the two separate axial flaw tests and the circumferential flaw data collected results in 23 “no-flaw” samples and 33 “flaw” samples. For the “no-flaw” samples the Mahalanobis distances from the good cluster center range from 3.08 to 4.55. For the “flaw” samples the Mahalanobis distances from the good cluster center range from 5.27 to 119.27. In other words, for the integrated data set, this classifier completely separates the “flaw” from the “no-flaw” signals.

The same classifier was applied to all three data sets separately. The separation between “flaw” and “no-flaw” signals for the smaller data sets are quite dramatic. For each of the three data sets, for the “no-flaw” samples the Mahalanobis distances from the good cluster center are on the order of 2 to 5. For the “flaw” samples the Mahalanobis distances from the good cluster center are on the order of 10^8 to 10^9 .

The same classifier also completely separates the corrosion data sets. The separation between “flaw” and “no-flaw” signals for this data set is also quite dramatic. For the “no-flaw” samples the Mahalanobis distances from the good cluster center are on the order of 2 to 5. For the “flaw” samples the Mahalanobis distances from the good cluster center are on the order of 10^8 to 10^9 .

Since the corrosion data were collected at 310 kHz, no attempt was made to integrate it with the 250-270 kHz data discussed above. In practice, if this algorithm were used in an on-line real-time system, it would be necessary to calibrate the algorithm by collecting a set of “no-flaw” data and computing a good cluster center for each combination of frequency, pipe diameter, and pipe wall thickness.

The 16-dimension classifier almost, but not completely separates the flaw from no flaw data for 10-inch pipe. However, the 17-dimension classifier completely separates the 10-inch pipe data. A composite data set of all the 10-inch pipe data, except for those “no flaw” signals where the tone burst overlapped the signal of interest was tried first. For this composite, across all flaw types for the 10 inch data, the algorithm completely separates “flaw” from “no flaw” signatures. Flaws were of all types: corrosion data collected in 2002, stress corrosion crack data collected in 2002 and 2003, axial flaw data collected in 2003, and 2003 circumferential flaw data collected in 2003.

The algorithm was also tried on each of the four flaw types in isolation. Given “flaw” and “no flaw samples” for the 10-inch pipe circumferential flaw data collected in 2003, the 17-dimension algorithm completely separated “flaw” from “no flaw” signatures. Given “flaw” and “no flaw samples” for the 10-inch pipe axial flaw data collected in 2003, the 17-dimension algorithm completely separated “flaw” from “no flaw” signatures. Given “flaw” and “no flaw samples” for the 10-inch pipe stress corrosion crack data collected in 2003 combined with the 10-inch pipe stress corrosion crack data collected in 2002, the 17-dimension algorithm completely separated “flaw” from “no flaw” signatures; this speaks to the robustness of the algorithm since the 2002 data were collected at 270 kHz excitation, and the 2003 data were collected at 250 kHz excitation. Given “flaw” and “no flaw samples” for the 10-inch pipe corrosion data collected in 2002, the 17-dimension algorithm completely separated “flaw” from “no flaw” signatures; in this case the “no flaw” samples had the tone burst overlapping the signal but the classifier worked anyway.

5. SOPHISTICATED SIGNAL PROCESSING

It might be asked if there is anything to be gained by using more costly signal processing algorithms than wavelet analysis. The data were examined in wavelet packet space. The “best basis” wavelet packet is considerably more costly than the wavelet transform because the signal is projected onto a variety of wavelet packet bases and the basis that produces the minimum entropy is taken as the “best basis.” It was found that the compression of the “no flaw” signals by the “best basis” wavelet packet for 12-inch pipe data was only slightly superior to the wavelet transform. For the 10-inch data, the wavelet transform itself (a degenerate case of the more general wavelet packet) turned out to be the best basis. In other words, the 58-coefficient least-asymmetric wavelet already looks so much like the signal under test that it is not worth the considerable added computational cost to find a wavelet packet function that looks slightly more like the signal.

Another powerful way to analyze ultrasonic signals is by the use of Bayesian parameter estimation. It can be highly useful in situations for which wavelet-based techniques are not adequate. This is especially true for situations in which the signal of interest is weak and it is biased by a very strong signal. This is the case for laser-generated Lamb waves, for which the large biasing signal can be subtracted out using Bayesian parameter estimation without disturbing the signal of interest.¹⁰

This situation also arises in EMAT data when the transmitter and receiver are very close together. In this case the very strong tone burst overlaps the much weaker flaw signature, and results in a signal that can confuse a wavelet-based classifier. Bayesian parameter estimation is very effective at identifying the tone burst and allowing it to be removed from the test data so that the residual can be classified with a wavelet-based classifier.

This technique is very computationally costly. It requires a global optimization of an objective that is pockmarked with local minima. For representative signals with the tone burst overlapping the flaw signature, several global optimizers were applied to the Bayesian algorithm. The Bayesian algorithm fits the experimental data to a Gaussian-windowed linear-chirped sinusoid, and provides four parameter values, carrier frequency, chirp rate, window width, and location of window peak.

The most reliable optimizer was the simple genetic algorithm (GA). It typically requires approximately 40,000 function evaluations (40 generations and a population of 1000) to obtain convergence. A 32-bit string is used, and each parameter is allocated eight bits. For this problem, the population was considered to have converged if the median fitness value is at least 97% of the maximum fitness value. In other words, convergence is defined as at least half the population having fitness values within 3% of the fitness value of the fittest member.

As is typical of simple GAs, it is crucial to use a large population if the convergence is not to be sensitively dependent on the tunable parameters used in the GA operators. A population of 100 never converged reliably. With a population of 2000, and tournament selection, reliable convergence was obtained for selection pressures from 4 to 14, and crossover probabilities of 0.4 to 0.75. Probability of mutation was 0.03. Eight bits per parameter leads to too much quantization noise to find the global minimum, but it consistently finds the global basin of attraction. Taking the result of the eight bits per parameter GA and using it as the starting point in a gradient descent algorithm is a relatively economical way of obtaining the global optimum.

6. CONCLUSIONS

In practice, it is not a good idea to combine signals collected at different excitation frequencies or signals with the tone burst overlapping, as done above to test the robustness of the algorithm. In the interest of reliability the full robustness of the algorithm should not be pressed to its limits; it is the same principle as not always driving a car at 120 mph just because it is occasionally possible to go that fast. The EMATs should be arranged such that the tone burst does not overlap the signature being classified.

The tone burst can be removed with Bayesian parameter estimation, but it is unfeasible to do so in real time. Using a compiled Mathematica implementation of the Bayesian algorithm on a 1.2 GHz Pentium IV processor, the evaluation of the fitness function in the global optimization required 6 ms per function call. A highly efficient C++ implementation might improve the speed to perhaps 0.5 ms per call. If the average optimization requires 40,000 calls to converge, it requires 20 seconds for the Bayesian algorithm to produce a result. This is about three orders of magnitude too slow for real time.

Also in practice, the algorithm should use different calibration data for each combination of pipe diameter, excitation frequency and sensor geometry. The calibration data set should be derived from at least 25 “no flaw” signatures, without the tone burst overlapping, for each combination. The 17-dimension feature vector is calculated for each sample. The calibration data is the statistical properties, specifically the 17-dimension mean vector (or cluster center) and the 17×17 inverted covariance matrix, of the “no flaw” cluster.

The decision of whether or not a given signature is a “flaw” or “no flaw” signature is based on the Mahalanobis distance of the 17-dimension feature vector of the signature from the “no flaw” cluster center. Thus, the other necessary calibration datum is the Mahalanobis distance from the cluster that is considered to be the decision boundary. This must be determined by computing the Mahalanobis distances from the “no flaw” cluster center for a large number of samples, at least 25 each for “flaw” and “no flaw” signatures, and placing the boundary between them.

The “no flaw” signatures tend to form a tight cluster. The “flaw” signatures tend to form a loose and wildly varying cluster; it tends to exaggerate the effect of noise. Thus, the Mahalanobis distance from the “flaw” cluster does not reliably separate “flaw” from “no-flaw” signatures and cannot be used as a reliable classifier. No computational energy need be wasted computing it.

The advantage of this 17-dimension algorithm is that it is feasible to apply it inexpensively in real time. Using a similar algorithm in another project, implementing it as a highly efficient C program on an 800-MHz Pentium III, computation of a wavelet-based feature vector and its Mahalanobis distance from a cluster center required approximately 4 milliseconds. With a 2.8-GHz Pentium IV, and the improved efficiency of the C++ compiler in Microsoft Visual Studio, it should be feasible to implement a version of the 17-dimension classifier that makes the classification decision for a given signature in 2 ms or less. Real-time implementation of this algorithm does not require a dedicated DSP chip.

However, it may be useful to use a control computer with a multiprocessor mother board where one Pentium IV is dedicated to this process. Windows 2000 and XP support multi-processor programming.

This research has found a single feature space in which flaws can be readily distinguished from no-flaw signatures. Although it is feasible to use different techniques to extract multiple features vectors as described above to identify different flaw types, having a single classifier lightens the computational load. It is feasible to classify each trial in real-time, and record only those signatures that are classified as flaws.

Ideally, classification by flaw type will also be performed in real-time by the inspection system on board the pig that serves as a platform for the EMATs. Different flaw types induce conversions of transmitted ultrasonic wave modes in predictable ways. The idealized time domain signature of a given mode is known. Once the presence of the flaw is determined by wavelet analysis, cross correlation can be performed between the idealized exemplar and the measured signal. This calculation should be easily realized in real time using Fourier-based cross correlation, and provide reliable classification by flaw type. There is also a possibility that this analysis will yield specific flaw parameters such as depth of crack, but that would be much more difficult than merely classifying by type.

7. ACKNOWLEDGEMENTS

This work was supported by the DOE National Energy Technology Laboratory (NETL), Strategic Center for Natural Gas. The work was performed at Oak Ridge National Laboratory, which is managed and operated by UT-Battelle, LLC., for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

The authors thank Regina Ferrell and Roger Kisner of Oak Ridge National Laboratory for their review and criticism of this paper. The author would also like to thank Dr. Joseph Rose of The Pennsylvania State University for his guidance in understanding the physical phenomena that occurs during ultrasonic wave propagation.

REFERENCES

1. Posakony, G.J. and Hill, V.L., "Assuring the integrity of natural gas transmission pipelines," Topical Report, Gas Research Institute, GRI-91/0366, November 1992.
2. Khaleel, M.A. and Simonen, F.A., "Effects of alternative inspection strategies on piping reliability," *Nuclear Engineering and Design*, Vol. 197, pp. 115-140, April 2000.
3. J.L. Rose, *Ultrasonic Waves in Solid Media*, Cambridge University Press, New York, pp. 44-46, 1999.
4. Thompson, R.B., "Physical Principles of Measurements with EMAT Transducers," in *Physical Acoustics*, W. P. Mason, Ed., Academic Press, Vol. XIX, pp.157-199, 1990.
5. Hirao, M. and Ogi, H., "An SH-wave EMAT technique for gas pipeline inspection," *NDT&E International*, Vol. 32, pp. 127-132, 1999.
6. Gauthler, J., Mustafa, V., Chabbaz, A. and Hay, D.R., "EMAT generation of horizontally polarized guided shear waves for ultrasonic inspection," in *Proceedings of the International Pipeline Conference*, Vol. 1, ASME, June 1998.
7. Sawaragi, K., Salzburger, H.J., Hubschen, G., Enami, K., Kirihigashi, A. and Tachibana, N., "Improvement of SH-wave EMAT phased array inspection by new eight segment probes," *Nuclear Engineering and Design*, Vol. 198, pp. 153-163, May 2000.
8. Kwun, H., Hanley, J.J. and Holt, A.E., "Detection of corrosion in pipe using magnetostrictive sensor technique," in *Nondestructive Evaluation of Aging Maritime Applications*, Richard B. Mignogna, Editor, Proceedings of SPIE Vol 2459, pp. 140-148, (1995).
9. Kwun, H., and Hanley, J.J., "NDE of steel Gas pipelines using magnetostrictive sensors," Topical Report, Gas Research Institute, GRI-95/0362, October 1995.
10. Kerckel, S.W., Klein, M.B. and Pouet, B., *In-process Detection of Weld Defects Using Laser-based Ultrasonic Lamb Waves*, ORNL/TM-2000/346, November 2000.