Automatic Defect Characterisation in Ultrasonic NDT

T. Stepinski and F. Lingvall, Uppsala University, Signals and Systems, Uppsala, Sweden

This paper reports results of the project concerned with development techniques for automatic defect characterisation based on ultrasonic data. The goal of the project was to develop a software tool facilitating interpretation of flaw signatures obtained from pulse-echo ultrasonic measurements. A neural network processing outputs from specialised feature detectors was used for the characterisation. Different ways of feature extraction and flaw position estimation are presented and discussed. In the experimental part ultrasonic measurements performed on welded carbon steel blocks including 36 different "natural" defects implanted into V-welds are shortly presented. The flaw population was divided in three major groups, sharp flaws (various types of cracks and lack of fusion), soft types of flaws (slag, porosity and over-penetration), and defect at the bottom of the weld. A large amount of B- and D-scan data was acquired using 6 different angle transducers. The evaluation of these measurements resulted in the conclusion that the signal variation within a certain class of defects was considerably larger than the corresponding variation found in signals from artificial and simulated defects. Steel block measurements also revealed that some of the defect types were hard to distinguish, particularly if only traditional features like, fall/raise times, pulse duration and echo-dynamics were used. To overcome this difficulty a more powerful feature extraction method was proposed, using discrete wavelet transform. In conclusion, difficulties encountered in automatic flaw classification are discussed and possible solutions are presented.

Introduction

Ultrasonic testing (UT) is one of NDT techniques that enable both defect detecting and sizing. However, the measurements typically obtained using UT are complex to evaluate since the response depends on many factors, such as lobe characteristics of the transducer, the transducer bandwidth and centre frequency, angle of inclination, depth and orientation of the defects etc. Today only very experienced operators are able to perform full evaluation of the results, especially for coarse-grained materials (stainless steel). Therefore, a considerable demand has been observed for tools that can support operators in such tasks as flaw sizing and classification.

The goal of the research presented in this paper was to develop a software tool, using signal processing and neural networks (NN) that would support an operator in making decisions concerning defect characteristics based on pulse-echo ultrasonic measurements. The idea was to study ultrasonic signatures extracted from the measurements, performed at our laboratory on a carefully chosen selection of "real" defects. A suitable set of signal processing tools should be then developed based on the characteristics of different flaw types, and the experience gained from the measurements. Previous research [1, 2], performed using simulated and artificial defects, has shown the feasibility of such an approach.

Due to practical reasons relatively few defects (36) of types similar to those commonly encountered in real V-welds were available for the measurements. A much larger number would have been desirable but the manufacturing cost of "real" defects is relatively high since the geometry of the used defects must be precisely known. One should also be aware that, due to physical restrictions, it was only possible to insonify the defects from a limited number of views. The inspections were performed from the front and the back surface of the test blocks, however, the access was limited by the weld.

The approach taken here consists in, first carefully analysing the ultrasonic responses, using different transducers, and then examining if there are features which are unique for the individual classes of defects, and hence, could be used for defect characterisation.

Methods and algorithms extracting these features in a format suitable for a classifier should be developed in the second step. Note that by "classifier" we mean either a human operator or dedicated software. If the software approach is chosen, then there is an apparent need of incorporating strong *a priori* knowledge since the number of examples is normally very limited. A reasonable goal, which was adapted here, was to select two classes, first containing sharp defects (various types of cracks) and second including soft defects (slag inclusion and porosity). Note that, even though only two classes are used, there are still very few examples of each class available for training.

Realistic Test Blocks

Four blocks, each with 9 various flaws, were manufactured by Sonaspection International Ltd. All blocks had dimensions 42 mm x 400 mm x 600 mm and consisted of two carbon steel plates, welded together (V-weld). The defect types and sizes manufactured in the blocks are summarised in Table 1 (see [6] for details).

Flaw type	Size in mm	No of flaws
Root crack	3 to 6	6
Lack of side-wall fusion	3 to 7	6
Side-wall crack	3 to 7	6
Centre line crack	3 to 6	6
Slag	3 to 6	3
Porosity	6 to 10	3
Over-penetration	3 to 5	3
Lack of penetration	2 to 25	3

Table 1. Specification defects used in the measurements

Test Block Measurements

The contact inspection of the blocks has been performed using a mechanised scanner and a digital ultrasonic system based on a Saphir PC board from Lecoeur Electronics. Two miniature screw-in transducers from Panametrics, with centre frequencies 2.25 MHz (type V539-SM) and 3.5 MHz (type A545S-SM) were used in the (shear wave) contact inspection. Both transducers had nominal element size 0.5" (13 mm) and were assembled by screwing directly into miniature angle beam wedges type ABWM-5T, also from Panametrics. Six different angle beam transducer configurations, (with angles 45, 60 and 70) were created in this way. The advantage of this solution is obvious – by using the same active element the obtained transducers have very similar characteristics. Since ultrasonic response of a particular flaw is determined both by the flaw type and by the transducer characteristics it is essential for defect characterisation to keep transducer characteristics as constant as possible.

The pulse-echo measurements were performed mostly in direct reflection mode from the upper block side; indirect reflection mode measurements and the measurements from the backside were also performed in some cases. B-scans and D-scans for each flaw were acquired and a database containing 2×133 (similar measurements for both frequencies) ultrasonic images was created.

The database was completed with some reference measurements, performed for comparison on two aluminium blocks used in a previous project containing artificial defects [4, 5]. In addition to all these measurements all flaws have also been subjected to manual inspections.

The detailed description of the acquired data for different flaw types and the examples of B-scans for 3.5 MHz transducer can be found in [6]. It appeared that the 2.25 MHz and the 3.5 MHz transducer syielded very similar results, with the exception that the 3.5 MHz transducer resulted in a higher resolution due to the shorter wavelength (both transducers had approximately the same bandwidth).

Signal Features and Feature Extraction

Generally, in each classification task employing self-learning scheme a trade-off between the number of examples available for training and the amount of *a priori* knowledge about the classes is encountered. The number of training examples must be compared to the complexity of the classifier (i.e., number of parameters in the classifier). If the training examples set is small and the classifier is complex, then the classifier will perform well on the training data and poor on the unseen data [3]. However, the number of parameters in the classifier can be reduced if the dimension of the input vector is reduced. The process of describing features in data in a compact way is known as feature extraction. To succeed well the features must be descriptive, so that differences vital for classification

are not lost. These aspects are illustrated in Fig. 1, where a fictitious example characterised by only two features is shown. In our application the number of training examples is very low and, hence, feature extraction is an essential part of the classification process.



Figure 1. Classification in two-feature space (a fictitious example).

Three different classes are shown in Fig. 1: first with dashed boundary (labelled A), second with dotted boundary (labelled B), and third with dash-dotted boundary (labelled C). Examples from each class are also indicated with different symbols. It is easy to see that class A and B are overlapping and they are represented by few examples only. This implies that it is difficult to design a classifier with a proper decision boundary based on those examples. Class C exemplifies the desired case with a sufficient number of examples and non-overlapping class boundaries. The two features used in this example are clearly not suitable for separating classes A and B. However, it is possible that for other feature set the classes may not be overlapping, and the problem can be solved.

In this study, only the envelope of the acquired ultrasonic data is used. This was also the strategy used in the previous study [3, 4]. The envelope is calculated by means of the Hilbert transform. The resulting data is also smoothed with a low-pass filter to reduce the measurement noise.

Flaw Position Estimation

The flaw position is used for both region of interest (ROI) selection the depth normalisation (which is required for the feature extraction performed later). It is therefore important that the position estimation is accurate and robust. The current method to find the flaw position is based on fitting an hyperbolic function to the flaw response in B-scan data, see [4]. The algorithm operation is illustrated at Fig 2. Consider the A-scan 48 mm from the centre of the weld, marked with a vertical line in Figure 2b (also included in the box in the same figure). The maximum response yielded by the algorithm is approx. r_{max} = 45mm.



Figure 2. Illustration of defect position estimation.

Depth Normalisation

Due to the lobe characteristics (cone-beam geometry) of the probe, a defect located close to the transducer will be seen in a fewer A-scans in a B-scan than a similar defect detected further away from the probe. Echo-dynamics of the flaw close to the transducer will have a narrower shape than the other echo-dynamics corresponding to the remote flaw. A simple way to normalise is to re-sample the echo-dynamics (or wavelet coefficients) in some angle interval. That is, the feature vector (or matrix) is re-sampled in an angular scale instead of the original linear scale. This is illustrated in Fig. 3, where the two horizontal arrows indicate the distances where the flaws f_1 , and f_2 , are inside the ultrasonic beam. The depth normalisation procedure consists of re-sampling the features for a suitable angular range given the depth of the flaw. This implies interpolating features from flaws located close to the probe, and down sampling features for flaws that lie further away from the probe.



Figure 3. Illustration of the effect of the cone beam geometry for two defects at different depths. Defect f_1 will be seen in fewer A-scans than defect f_2 .

ROI Selection

Region of interest (ROI) for defect characterisation is an important issue since all further processing relies on it. If the analysing window for ROI is not positioned correctly the features fed to the classier will vary between different measurements, giving inconsistent results. It appears however, that accurate positioning of the analysing window in ultrasonic B-scans is not a trivial task.

Ideally, a hyperbolic analysing window should be positioned around a flaw response in a B-scan, and the exact position of the window should be determined based on the estimated position of the flaw. The estimation procedure described above is, however, not accurate enough for the precise horizontal positioning required here. In the previous studies [4, 5] the echo-dynamics (max amplitude variation) of the flaw response was used for horizontal positioning.

However, for realistic defects the echo-dynamics curves may be skewed, have more than one peak, etc. The approach used previously was to smooth the echo-dynamics, with low-pass filtering, which partially solves the problem. This method was suitable for the simulated and artificial defects. Experiments have been performed using centre-of-mass calculations in order to find a robust estimate for the echo-dynamics centre point. This approach was, however, too sensitive to long tails with high amplitude (energy) in the echo-dynamics. Therefore, the previously used algorithm was adopted here as well. The algorithm includes: low-pass filtering and finding the maximum of the echo-dynamics in the first step, and selecting a number of A-scans centred around the A-scan corresponding to max amplitude in the second step.

Classical Features

The perhaps most commonly used features for classification of defects during ultrasonic inspection is the rise time, pulse duration and fall time [5, 6]. These three features are calculated from the A-scan envelope using time instants corresponding to the 10% and 90% amplitude levels.

These basic features are reliable provided that the ultrasonic pulses (echoes) are well defined. However, reliability of these features may be considerably impaired for realistic defects since they often result in pulses with envelopes that cannot be approximated by well-defined bell curves. Smoothing (low-pass filtering) partly alleviates these problems, but at the expense of some loss of information. Fig. 4 shows three examples of envelopes corresponding to different types of defects.

In spite of the very different shape of the waveforms, the rise time, pulse duration, and fall time are rather similar for all of the signals in Fig. 4. It is evident that more powerful features are needed if the classification should be feasible for this type of signals.



Figure 4. Envelopes of A-scans corresponding to three different flaws. The pulse duration and the raise- and fall times are similar despite the very different pulse shapes.

Feature Extraction using the Discrete Wavelet Transform

The discrete wavelet transform (DWT) has several interesting features for this application. The impulse-like nature and the locality of the basis functions in the DWT make it suitable for modelling ultrasonic signals. If an analysing window is centred on an ultrasonic pulse it is possible to examine at which position and scale this pulse has significant energy, which is reflected in the wavelet coefficients produced by the DWT.



Figure 5. (a) The echo-dynamics corresponding to a slag inclusion. (b) First 16 wavelet coefficients from the A-scans corresponding to (a).

It is worth noting that the envelopes of typical ultrasonic signals can be well described with only a few of the large-scale components (corresponding to lower frequencies).

Wavelet representation has good data compression ability. There are several different types of predefined mother wavelets available in common software packages, like the Wavelet Toolbox for MATLAB^M. The *Coiflet 2* mother wavelet, which is fairly smooth suits well for this application, [6]. The echo-dynamics, and the first 16 DWT coefficients, from he same (consecutive) A-scans are displayed in Fig. 5. It is clearly seen how the echo-dynamics is reflected in the wavelet coefficients. Note also that there is a significant energy (information) for more than one scale. Obviously, if only the echo-dynamics is used, a significant amount of information can be lost which then may impair the classification performance.

Defect Classes

After analysis of ultrasonic data acquired from the realistic flaws it became apparent that the characterisation task is much more complex for the realistic than for the artificial and the simulated defects. The variability of ultrasonic responses from the same type of defects appeared to be very large which resulted in a considerable overlapping of flaw classes in feature space. This obviously resulted in problems with the flaw classification.

A realistic goal is to categorise defects in sharp defects, like cracks and lack of fusion, and volumetric (or soft defects), like porosity and slag inclusions. Defects in the bottom of the weld are also easy to distinguish from other flaws since they all occur at the same position.

Sharp Defects vs. Volumetric Defects

Examples of A-scan envelopes and echo-dynamics for crack-like and volumetric defects are shown in Fig. 6 and 7, respectively. Analysis of Fig. 6 shows that pulse shape corresponding to the sharp defects is well defined but occasionally double echoes can be found.



Figure 6. (a) Examples of A-scan envelops from sharp defects. (b) Echo-dynamics in the range -5 to 5 degrees from sharp defects.



Figure 7. (a) Examples of A-scan envelops from volumetric defects. (b) Echo-dynamics in the range -5 to 5 degrees from volumetric defects.

Fig. 7 shows that pulse shape from slag inclusions is well defined in contrary to porosity characterised by multiple echoes. Unfortunately, no significant difference in echo-dynamics can be observed for both classes. At least three general observations can be made from the measured data:

- The ultrasonic responses from crack type defects exhibit a large variation of features. This is especially clear for centre cracks and lack of fusion defects.
- It is very difficult to distinguish side-wall cracks and lack of fusion from slag inclusions.
- It is difficult to draw any general conclusions from the echo-dynamics.

Generally, a larger number of examples from the sharp class of defects would be required, in order to see as many variations as would be needed to construct a fully automatic classifier based on training examples only. It is obvious that the classical features are insufficient for the classification, which implies that more sophisticated tool is needed for efficient feature extraction.

Defects at the Bottom of the Weld

The within-class variation seems much smaller for these defects than for the other classes. The class separation between the three different flaw types also appears larger than in the former case. Lack of penetration has a rather "clean" pulse shape, over-penetration is characterised by multiple pulses, and root cracks often result in a small pulse appearing slightly before the main pulse, which can be seen in Fig. 8.



Figure 8. Examples of A-scan envelops from defects at the bottom of the weld.

The classical features may be sufficient for separation of the three flaw types, at least if the multiple pulses characteristic for over-penetration, and the small initial pulse of the root cracks, are not separated too far from the main pulse. A large separation between the multiple pulses may result in that only the main pulse is used for feature extraction.

Realistic vs. Artificial flaws

As mentioned earlier B-scan data was also acquired for the aluminium specimens with artificial defects used in the previous studies [4, 5, 6]. In these studies the classical features, described above, were shown to be sufficient. For comparison with the present results some measurement were performed on the aluminium specimens. The pulse shapes obtained for the artificial defects are very "clean" compared to the ones in the steel blocks with realistic defects. No double echoes or irregular pulse shapes are present in the signals from the artificial cracks in the aluminium data. The defect characterisation (classification) task becomes much simpler since the within-class variation is much lower than for the realistic flaws.

The main implication of this is that the number of data needed "to span" the subspace of possible flaw signals is much lower for artificial defects than for the realistic defect counterpart. The features sufficient for the classification are also simpler, echo-dynamics, rise time, pulse duration, and fall time, work well for the artificial defects [5].

Conclusions

During the evaluation of ultrasonic data acquired from the V-welded steel blocks it became evident that the characterisation task is much more complex than for simulated and artificial flaw signals. The feature space of possible flaw signals is also considerably larger for the real defects than for the artificial counterparts, i.e., the variation of the ultrasonic signals within one type (class) of defects is much larger for real than for artificial defects.

Our goal was to separate soft (or volumetric) defects from the sharper ones (crack-like defects), but if one studies the echo-dynamics and the pulse shapes (i.e. the envelope) it becomes apparent that some sharp and soft defect types are very hard to separate. This implies that overlapping feature regions are encountered, especially when using classical features (fall/raise times, pulse duration and echo-dynamics). To avoid overlapping class boundaries, more powerful feature extraction algorithms are needed to achieve a good classification performance.

High variation of the ultrasonic signals also has two further consequences: flaw position estimation (needed for feature extraction) may be poor and the amount of data needed to construct reliable classifiers large.

The following is recommended during design of a self-learning classifier:

- Ensure that the measurements are informative enough to distinguish between different types of defects.
- The features must be representative. The features that are fed to the classifier must preserve the information needed for successful classification.
- The number of representative training examples must be sufficient. As a rule of thumb at least ten times as many examples as the parameters in the classifier are needed, to avoid that the classifier learns the training examples and performs poor on unseen examples.

The first recommendation may be not fulfilled using single B-scan measurements only. A common practice is to combine measurements from several transducers (with different angles, centre frequencies, etc.) and TOFD measurements. The second condition is clearly not fulfilled if only classical features are used and therefore more powerful methods are needed.

The third condition is a common problem in NDT, it can be hoped that mathematical modelling can contribute to the *a priori* knowledge and thus lead to decreasing the required number of examples. This knowledge can also take the form of expert knowledge acquired from experienced operators. A reasonable approach to defect characterisation is to concentrate on improving flaw imaging and leave the classification tasks to experienced operators.

Acknowledgement

This work is a part of project supported by the Swedish Nuclear Power Inspectorate (SKI).

References

- 1. C.G. Windsor, Can we train a Computer to be a Skilled Inspector, *Insight*, 37, January 1995, p. 36.
- 2. A. MacNab and I. Dunlop, Review of Artificial Intelligence Applied to Ultrasonic Defect Evaluation, *Insight*, **37**, January, 1995, p 11.
- 3. K. Fukanaga, Introduction to Statistical Pattern Recognition, Academic Press, 1990.
- 4. B. Eriksson and T. Stepinski, Ultrasonic Characterization of Defects, Part 2. Theoretical Studies, *SKI Report 95:21*, Swedish Nuclear Power Inspectorate, 1995.
- 5. B. Eriksson, T. Stepinski and B. Vagnhammar, Ultrasonic Characterization of Defects, Part 3. Experimental Verification, *SKI Report 96:75*, Swedish Nuclear Power Inspectorate, 1996
- 6. F. Lingvall, Efficient Data Representations for Eddy Current and Ultrasonic Applications, *Lic. Thesis*, Uppsala University, Signals and Systems, May 2000.