Journal of Nondestructive Evaluation, Vol. 23, No. 4, December 2004 (© 2004)
DOI: 10.1007/s10921-004-0822-4

Pattern Recognition of Weld Defects in Preprocessed TOFD Signals Using Linear Classifiers

Elineudo P. Moura,[1,3] Romeu R. Silva,[2] Marcio H. S. Siqueira,[2] and João Marcos A. Rebello[2]

Received January 17, 2003; revised July 20, 2004

The TOFD (*"Time of Flight Diffraction"*) technique is being widely used for automatic weld inspection, especially in the petrochemical industry, where welding quality is essential to avoid productivity losses. Although it provides high speed inspection, high sizing reliability and low rate of false defect indications, the classification of defects using ultrasound signals generated by the TOFD technique is still frequently questioned, because it depends heavily on the knowledge and experience of the operator. However, the use of computational tools for signal preprocessing and pattern recognition, such as the artificial neural networks, improves the classification reliability of defects detected by this technique. In this present work, three kinds of defects: lack of fusion (LF), lack of penetration (LP) and porosity (PO) were inserted into the specimens during the welding process, generating pattern defects. The position, type and dimension of each inserted defect were recorded using conventional ultrasonic and radiographic techniques. The Fourier Transform and Wavelet Transform were used for preprocessing A-scan signals acquired during weld inspection by TOFD technique. This study was able to show the versatility of Wavelet Transform to preprocess these kinds of signals, since the correct scale in Continuous Wavelet Transform had been selected to supply a neural network. Hierarchical linear classifiers were implemented into the neural network in order to distinguish the main defects in welded joints detected by the TOFD technique. The results show the good success rate of welding defect recognition in preprocessed TOFD signals, mainly using Wavelet Transform. On the whole, the results obtained were very promising and could give relevant contributions to the development of an automatic system of detection and classification of welding defects inspected by the TOFD technique.

KEY WORDS: Non-destructive testing; TOFD; welding defects; pattern recognition; artificial neural networks; Wavelet Transform.

1. INTRODUCTION

The detecting and sizing of discontinuities carried out by ultrasonic techniques usually uses the amplitude of the echo obtained and relates such amplitude directly to the discontinuity dimension.

The TOFD technique is not based on echoes amplitude, but uses the travel time of a diffracted wave at the tip of a discontinuity to determine its depth. The principle of the technique was demonstrated by Silk[(1–3)] in the 70's. The technique utilizes two transducers, one as a transmitter and the other as a receiver in such a way as to cover the volume of material being inspected. The first echo to reach the receiver transducer corresponds to the surface wave. If there are no

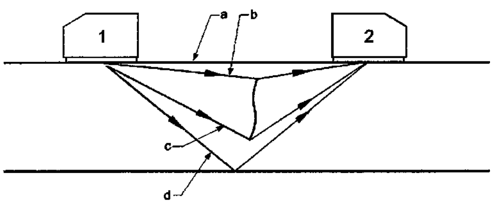
[1] Department of Mechanical Engineering, Federal University of Ceará, Campus do Pici, Centro de Tecnologia, Bloco 714, 60455-760, Brazil.
[2] Department of Metallurgical and Materials Engineering, Federal University of Rio de Janeiro (COPPE/UFRJ), PO Box 68505, CEP 21945-970.
[3] Corresponding author: E-mail: elineudo@yahoo.com

discontinuities the second signal will be the backwall echo. The importance of these two signals is that they can be used as a reference for measuring time from other waves. Any signals generated by discontinuities will occur between the surface wave and the backwall echo. A-scan mode is the most typical form of an ultrasound signal, which can be displayed on the screen of the ultrasound equipment as an amplitude-versus-time trace. A typical set-up and A-scan of the TOFD technique are shown in Fig. 1.[4]

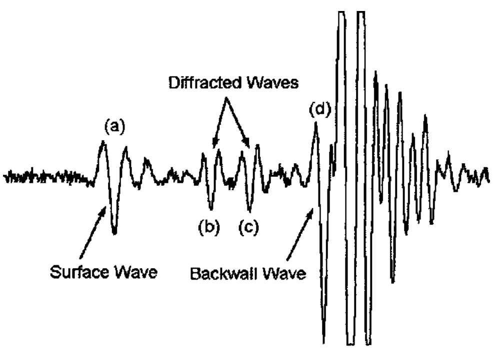
The acquisition of several A-scan signals consecutively recorded along a weld bead allows the formation of a D-scan image, where the positive and negative amplitudes are converted into gray scale.

The TOFD technique represents considerable progress since it allows the graphical recording of the weld bead inspection but despite the advantages of the ultrasound over the radiographic test, the TOFD technique does not provide precise information about the type of defect detected.[5] The defect classification based on its ultrasonic signals is still frequently questioned for being very subjective, since the analysis and the identification of defect types depend exclusively on the experience and knowledge of the operator.

The progress in computational techniques, specifically the development of neural networks, has greatly stimulated the research into the development



(a)



(b)

Fig. 1. (a). Typical set-up for TOFD. (1) Pulse probe, (2) receiver probe, (a) lateral wave, (b) diffracted wave by the upper tip defect, (c) diffracted wave by the lower tip defect, and (d) echo from the backwall.[4] (b). Typical A-Scan of the TOFD technique.

of automatic systems for the inspection and the classification of defects.[6–10] Neural networks use algorithms that learn functions, such as pattern recognition, creation of associations, signal processing and learning by experience or training. Although they are less complex than the human brain, the neural networks can process enormous amounts of data in a short period of time that typically could only be analyzed by specialists. One of the most important characteristics of the artificial neural network is the ability to be trained or learn by example, exactly like the human brain.[11,12]

In this study, ultrasonic signals were acquired using the TOFD technique during weld bead inspection with three (3) different kinds of defects: lack of fusion (LF), lack of penetration (LP) and porosity (PO). One class of signals from regions presenting no defect (ND) was also defined to evaluate the capability of these classifiers to identify signals from welds with defect or welds that presented no defects. Two different kinds of preprocessing (Fourier and Wavelet Transform) were applied to signals to be *a posteriori* distinguished by hierarchical and non-hierarchical linear classifiers. A neural network implemented these classifiers and their performance was then evaluated.

2. PREPROCESSING—FOURIER AND WAVELETS TRANSFORM

There are a lot of tools for signal processing, and Fourier Transform is one of the most important. The Fourier Transform can decompose a signal into constituent sinusoids of different frequencies, and this is enough for the characterization of a number of signals.

Mathematically, the process of Fourier Transform is represented by $F(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t}\, dt$, which is the sum over all the time domain of the signal $f(t)$ multiplied by a complex exponential. The results of the transform are the Fourier coefficients $F(\omega)$, which when multiplied by a sinusoid of appropriate frequency ω, yield the constituent sinusoidal components of the original signal.[13,14]

However, Fourier analysis has a serious drawback. When transformed to the frequency domain, time information is lost. If a signal does not change much over time—that is, if it is what is called a *stationary* signal—this drawback is not very important. However, most signals of interest contain numerous non-stationary or transitory characteristics. These characteristics are often the most important part of the signal, and Fourier analysis is not suitable to detect them.

The Wavelet Transform was developed especially to overcome these deficiencies. It is a windowing technique with variable-sized regions, which allows the use of long time intervals to obtain more precise low frequency information, and shorter regions where high frequency information is needed.[14,15]

Besides, the Wavelet Transform is capable of decomposing a signal into shifted and scaled versions of the original (or *mother*) wavelet. The Continuous Wavelet Transform (CWT) is defined as the sum over all the time domain of the signal, $f(t)$, multiplied by scaled and shifted versions of the wavelet function Ψ.

$$C(\text{scale, position}) = \int_{-\infty}^{\infty} f(t)\Psi(\text{scale, position}, t)\, dt$$

The results of the CWT are many wavelet coefficients C, which are function of scale and position. Multiplying each coefficient by the appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal.[14]

The coefficients produced on different scales in different sections of the signal constitute the results of a regression of the original signal performed on the wavelets. Figure 2a represents a graph for which the x-axis corresponds to the position along the signal (time), the y-axis represents scale and the color at each x-y point represents the magnitude of the wavelet coefficient C. A three-dimensional image would look something like Fig. 2b.

Notice that the scales (shown as y-axis labels) run from 1 to 60. The higher scales correspond to the most "stretched" wavelets. Thus, for higher scales a longer portion of the signal is compared with a more stretched wavelet. This allows the low frequency to be observed and thus slow changes and coarse features are measured by the wavelet coefficients. Similarly, for low scales, compressed wavelets are used, which allow the abrupt changes and details to be observed, that is, observation of high frequency.

3. MATERIALS

Inspections using the TOFD technique were performed on twelve test samples of steel plate AISI 1020, 20 mm in thickness, 300 mm in length, 50° V-bevel and welded by shielded process. As already mentioned previously different defects, such as lack of fusion (LF), lack of penetration (LP) and porosity (PO) were inserted into the test samples during the welding process, generating pattern defects. The position, type and dimension of each inserted defects were recorded by conventional ultrasonic and radiographic tests.

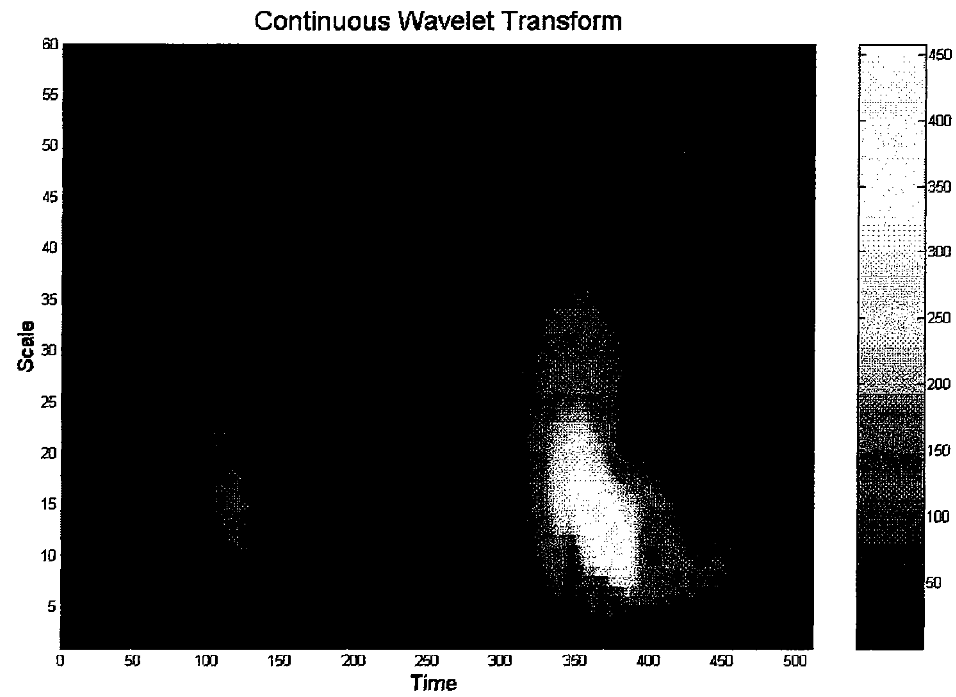
3.1. Acquisition of Signals

An automatic girth weld scanner was used in the acquisition of several A-scans, obtained by displacing the transducer along of weld bead. The automatic system of inspection is composed of a scanner unit designed to transport one pair of 5 MHz normal transducers with 60° wedges for longitudinal waves; a conventional ultrasound equipment; an A/D converter; a microcomputer (PC) and a program developed at the Non-Destructive Testing Laboratory—LABOEND/COPPE/UFRJ that controls the scanner unit, stores the signals from the inspected weld region and allows viewing of both A-scan and D-scan simultaneously. The signals, containing 512 points each, were acquired every 1 mm along the weld bead. Results obtained in the TOFD ultrasound tests were confirmed by X-ray tests.

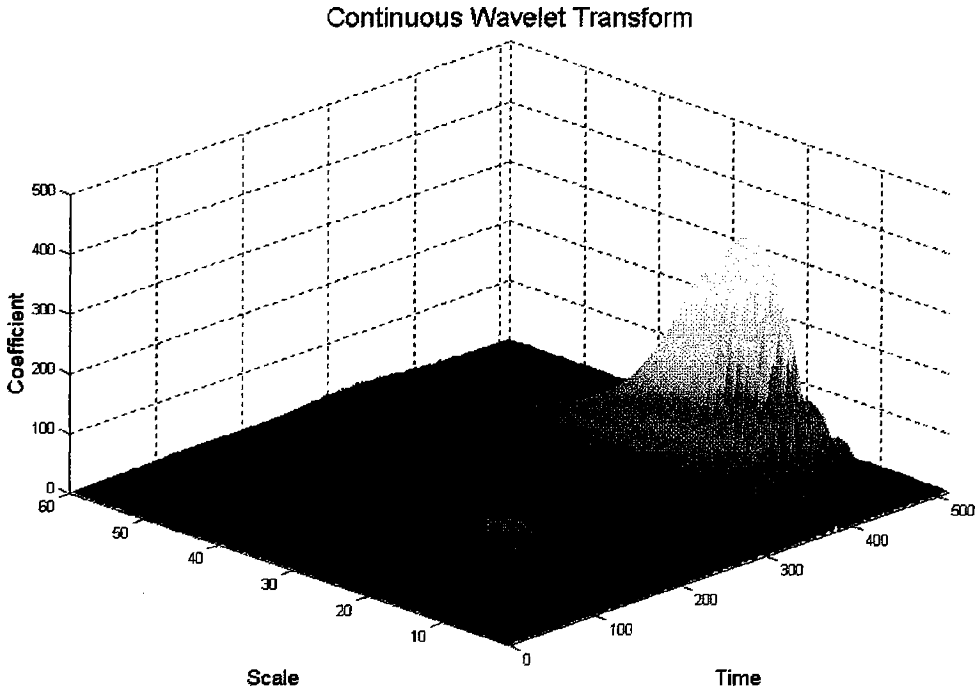
3.2. Selection of Signals

After the inspection of twelve test samples, a total of 240 signals (A-scan), equally divided into the four classes—lack of fusion (LF), lack of penetration (LP), porosity (PO) and one class of signals presenting no defect (ND)—were selected to be the input of the preprocessing by Fourier and Wavelet Transform before supplying the neural network. Each A-scan has 512 points. Thus, each input of the neural network originating from the Fourier Transform preprocessing is represented by a vector $\vec{x}$ of dimension 256 (256 freq. Fourier coefficients of each A-scan), or geometrically, by one point in a space of dimension 256, called space of inputs. In case of Wavelet Transform, the vector $\vec{x}$ has dimension 512 (512 Wavelet coefficients of each A-scan).

In both the preprocessing processes, from the 60 signals selected for each class, 40 signals were used during the network training stage, and 20 signals were reserved to test the capability of the classifier to identify signals not presented during the training process.



(a)



(b)

Fig. 2. (a). Wavelet coefficients of the A-Scan shown at Fig. 1. (b). Three-dimensional time-scale diagram of Fig. 2a.

4. HIERARCHICAL AND NON-HIERARCHICAL LINEAR CLASSIFIERS

In this study, each input of the neural network is represented by a vector $\vec{x}$, dimension: 512 (one signal with 512 points), or geometrically, by one point in a space of dimension 512, called space of inputs. A linear discriminator for the class C_j separates the inputs of this class from the others, through one linear inequation of the first order,

$$\vec{x} \in C_j \iff u_j > 0 \tag{1}$$

where

$$u_j = \sum_{i=1}^{512} w_{ji} x_i + b_j = \vec{w}_j^t \vec{x} + b_j \tag{2}$$

each class C_j has its own discriminator, defined by $\vec{w}_j$ and b_j.

In the input domain, the separator of the class C_j, that is, the geometric place of the points that fulfills $u_j = 0$, is a hyper plane which is perpendicular to the vector $\vec{w}_j$ and distant from the origin $-b_j/|\vec{w}_j|$, a distance measured in the direction of $\vec{w}_j$. Usually, it is regulated to $|\vec{w}_j| = 1$, adjusting the value of b_j not to alter the inequation (1). In this case, u_j measures the distance from the input $\vec{x}$ to the separator, and u_j is usually the success probability measurement of the classification for that specific input.(11,16)

The excellent linear discriminators are those that maximize the precision probability of the classification. One practical way to implement them is through an one-layer neural network, with only one neuron per class as described by Haykin.(11) This technique was used in this study. With a set of training available, networks with supervised learning were used, where the algorithm of the error backpropagation was used for training of the neural network.

The geometric display of the separators, which in the present case was difficult, due to the input space dimension (512), although in a bidimensional space it can be done easily. In Fig. 3, the shaded area shows the domain of hypothetical inputs of classes C_j and its respective separators S_j (that, in this case, are represented by the straight lines).

Each separator S_j divides the input space into two semi-spaces (in this case, two semi-planes), one with $U_j > 0$ and one with $U_j < 0$. The inputs belonging to the class C_j and that are classified incorrectly are represented by points in the semi-plane where u_j is positive. Notice that there are regions in the positive semi-space of two or more separators: one input in this region will be assigned to two or more classes. On the other hand, there might be regions in the negative semi-space of all separators: one input in this region will not be assigned to any class. In this situation, it can be considered that u_j is the probability measurement of one input belonging to the class C_j and "tie-breaking" the result, taking the class with the greatest u_j, which is the most probable one to include the input, as being the answer.

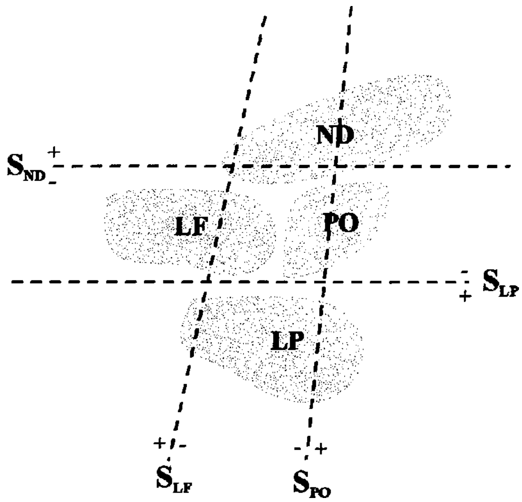


Fig. 3. The four classes C_j, j = LP, ND, PO and LF, and their respective non-hierarchical separators S_j with the polarities indicated. The most external classes LP and ND are almost perfectly separated, and the most internal classes PO and LF are imperfectly separated.

It can also be clearly observed that the most "external" classes are more easily separable, while the "internal" ones are hardly separable (Fig. 3). But, if these "external" classes are removed, other classes that had previously been "internal" will become "external," and then can be separated easily (Fig. 4). This introduces the concept of hierarchical classification, where the "external" classes are classified first, that is, those with high precision level, and only afterwards, the "internal" classes are classified.(17)

As Fig. 4b shows, there is a small region of ND that is not correctly classified by the separator S_{ND}, being situated in the negative semi-space. In this case, this set of signals regarding the class ND is not removed from the system, and the classifier of the class PO is found, not only to separate PO from LF, but also from such set of signals belonging to the class ND.

It is important to point out that those figures are simplified illustrations that describe the method used to obtain hierarchical and non-hierarchical classifiers.

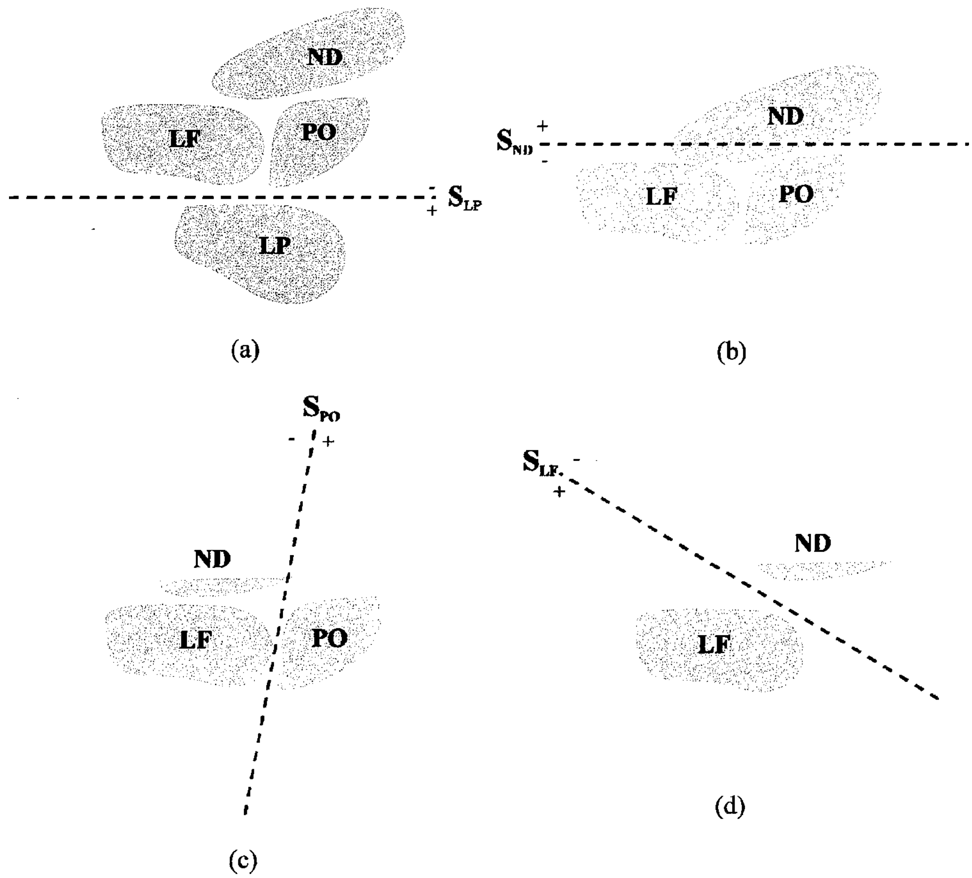


Fig. 4. (a) and (b): After the exclusion of the inputs classified as LP and ND, the other discriminators can assume positions that are much more efficient in separating the remaining classes, mainly S_{PO}, (c), and the same happens after the exclusion of the inputs classified as LF (d).

Their sole purpose is to show the operating principle of these classifiers and not to represent the exact condition of the signals used in this study.

The algorithms are constructed after obtaining the discriminators of each class. The flow diagram of a non-hierarchical classifier is shown in Fig. 5. The input vectors are multiplied by the vectors $\vec{w}_j$, of each class and added to the corresponding bias b_j, generating u_j. The result of this operation that is greater than zero corresponds to the selected class. In this situation, it is possible to have no class being indicated (when all outputs are negative), or the occurrence of more than one indication (when more than one output is greater than zero). In this case, one reclassification criterion, or a "tiebreak," can be used, in which the greatest value of u_j indicates the class. For both cases (with or without reclassification), the tables of defect-type confusion, precision and errors were constructed, based on the results obtained.[17]

Unlike the non-hierarchical classifier, the hierarchical classifier first classifies the most easily separable classes. The algorithm of this classifier is shown in Fig. 6.[17] Its performance was similar to the non-hierarchical classifier. The algorithms of the non-hierarchical and hierarchical classifiers are compared below, in terms of precision percentage.

5. RESULTS AND DISCUSSION

It was shown in a recent work[17] that the hierarchical classifiers normally present a better performance than the non-hierarchical ones and the reclassification criterion improves the obtained results. Using this criterion, the greatest output value of the classifier is taken into account as a tiebreak for the result. Therefore only the training and test results obtained by hierarchical classifiers with reclassification

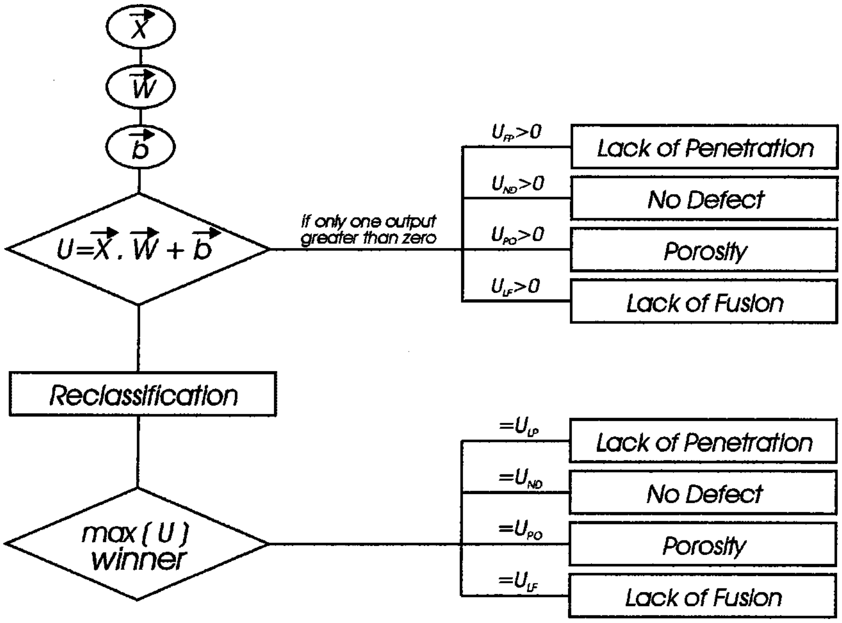


Fig. 5. Algorithm of the non-hierarchical linear classifier.

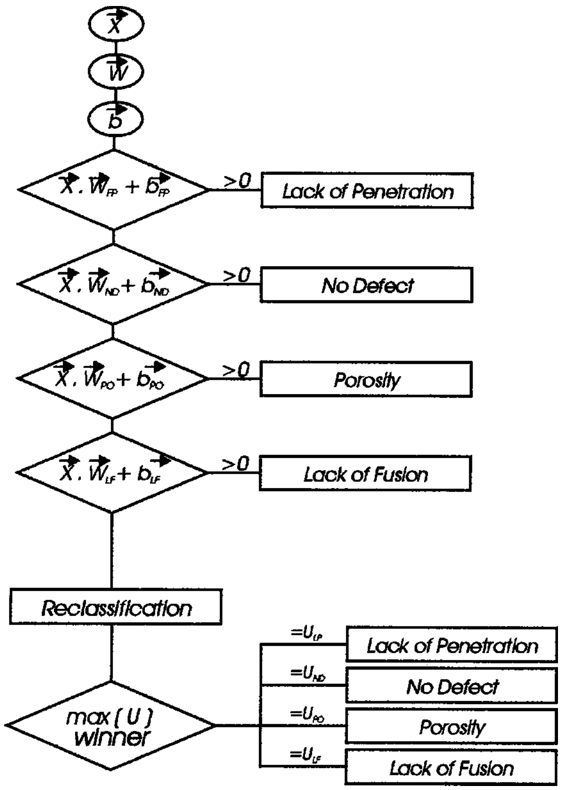


Fig. 6. Algorithm of the hierarchical linear classifier.

are shown. It must be stated that the test data are data used to evaluate the classifier performance and not used during training.

The order to distinguish classes (hierarchy) was chosen based on the results of the non-hierarchical process (not shown here). The first class is the one that presented the best rate of success in identification and the last class is the one that presented the worst rate.

The following tables show the rate of success for training and test data of the four A-scan classes. Tables I and II present the results obtained when the amplitudes of the frequency spectra were used to supply the classifiers. Tables III and IV present the ones obtained when the wavelet coefficients taken from the largest energy scale of the time-scale diagram were used.

From Tables I and III, it is possible to see that the overall performance of this kind of algorithm for training data classification (about 78% for the amplitudes of the frequency spectrum and about 99% for the wavelet coefficients) is superior to the overall performance for the test data classification (72.5% and 92.5%) as seen in Tables II and IV respectively. This is so because it is normally easier to classify a supplied signal to the neural network during the training process.

It must be pointed out that there is an influence of the type of preprocessing on the rate of success. It can be clearly seen that the results obtained with the

Table I. Table of Confusion, Successes and Errors—*Hierarchical Linear Classifier* - Amplitude of Spectrum—Training Data—With Criterion of Reclassification (%)

	Neuron Winner					
Class of signal	LF	LP	PO	ND	Success	Errors
LF	72.5	5	10	12.5	72.5	27.5
LP	0	85	5	10	85	15
PO	0	2.5	82.5	15	82.5	17.5
ND	5	0	22.5	72.5	72.5	27.5
Total					78.125	21.875

Table III. Table of Confusion, Successes and Errors—*Hierarchical Linear Classifier* - Wavelet Coefficients—Training Data—With Criterion of Reclassification (%)

	Neuron Winner					
Class of signal	LF	LP	PO	ND	Success	Errors
LF	100	0	0	0	100	0
LP	0	100	0	0	100	0
PO	0	0	97.5	2.5	97.5	2.5
ND	0	0	0	100	100	0
Total					99.375	0.625

wavelet coefficients (about 99% and 92.5% as shown in Tables III and IV) are considerably better than the ones obtained with the amplitudes of the frequency spectra both for training as well as for test (about 78% and 72.5% as shown in Tables I and II).

The worse performance of the classifier supplied with amplitudes of the frequency spectra of the A-scan signals when compared to the performance obtained by the wavelet coefficients can be explained in several ways. Firstly, all the time information is lost during the transformation of the signal to the frequency domain by the Fourier analysis and it is impossible to say when a particular event took place. This is important in the analysis of the *non-stationary* signals[14] and in particular to signals captured by TOFD technique that uses the time of flight of the diffracted wave. For the Wavelet transform, this does not occur. Secondly, it was observed that none of the defect classes has a well-defined frequency spectrum. Besides this, all of the classes present a great dispersion in spectra.

When the amplitudes of the frequency spectrum of the training data are supplied to the network (Table I) the class "lack of penetration" (LP) shows the best performance (85%) followed by the "porosity" class (82.5%) and the worst classes are "no defect" (72.5%) and "lack of fusion" (72.5%) with the same rate of success.

A change in the success rate and in the hierarchy of class separation of the original signal classes (LP; ND; PO; LF) was observed between signals preprocessed by Fourier analysis (PO; LP; LF; ND) and signals preprocessed by Wavelet analysis (ND; LF; PO; LP). This is seen by comparing Table I with Table III.

The "no defect" class was considered to be the easiest to be distinguished as it had the simplest signal, but this was not supported by the results. This is due to the fact that some noise signals were probably taken as defect signals.

Highlighted is the "lack of penetration" class (LP) which showed the best performance (85%) in Table I and maximum rate of success (100%) in Tables III and IV presenting a high level of separation.

Also the kind of the preprocessing should be pointed out as having an important influence on the success rates of 100% achieved for the classes LF and ND (Table III). This assures the efficiency of the wavelet transform as a preprocessing method and that the three classes are linearly discerned from the others. This result can be considered very good taking into account that a linear discerning criterion for classifying defects was used.

Analysing the errors during reclassification of all tables presented, it can be seen that none of the classes studied was systematically confused with another one.

Table II. Table of Confusion, Successes and Errors—*Hierarchical Linear Classifier* - Amplitude of Spectrum—Test Data—With Criterion of Reclassification (%)

	Neuron Winner					
Class of signal	LF	LP	PO	ND	Success	Errors
LF	80	0	15	5	80	20
LP	0	75	10	15	75	25
PO	5	20	65	10	65	35
ND	0	5	25	70	70	30
Total					72.5	27.5

Table IV. Table of Confusion, Successes and Errors—*Hierarchical Linear Classifier* - Wavelet Coefficients—Test Data—With Criterion of Reclassification (%)

	Neuron Winner					
Class of signal	LF	LP	PO	ND	Success	Errors
LF	100	0	0	0	100	0
LP	0	100	0	0	100	0
PO	0	5	80	15	80	20
ND	10	0	0	90	90	10
Total					92.5	7.5

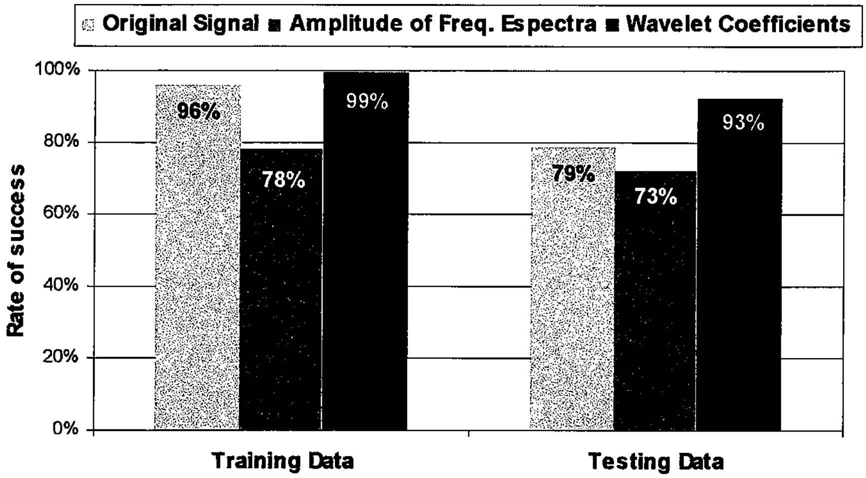


Fig. 7. Success rate of classifier supplied with three different preprocessing data.

Figure 7 compares the results obtained when the classifier is supplied with original A-scan signals and with A-scan signals after each of the different preprocessing methods. A minor rate of success achieved by the classifier supplied with the amplitude of the frequency spectra compared to the one supplied with the original signals was verified. Also an improvement in the results obtained by wavelet coefficients of the signals compared to the original ones can be seen. This is because the preprocessing revealed relevant information in the signal that is characteristic of the class and eliminated irrelevant information that could confuse their classification.

6. CONCLUSIONS

The linear pattern classifiers implemented by neural networks appear to be very efficient in recognition of defect ultrasonic signals obtained during inspection using the TOFD technique.

Furthermore, the wavelet transform used as preprocessing method for the signals made it easy to use these classifiers, improving considerably the performance of the classification. The fact that the test data are different from the training data confirms the ability of identification of the unknown signals in the classifiers studied.

As a whole, the results are very promising and could give relevant contributions to the development of an automatic system of detection and classification of welding defects inspected by the TOFD technique. The use of non-linear classifiers (the aim of future works) probably will further improve these results. Although other types of defect like cracks, undercut, slag, etc, were not studied, the same procedure will be used in a future work.

ACKNOWLEDGMENT

Authors wish to acknowledge CNPq and FAPERJ (Brazilian Agencies for Scientific Development) for their financial support.

REFERENCES

1. M. G. Silk and B. H. Lidington, "Defect sizing using an Ultrasonic Time Delay Approach," *Br. J. NDT* **March**, pp. 33–36 (1975).
2. M. G. Silk, "Sizing Crack-like Defects by Ultrasonic Means," In *Research Techniques in Nondestructive Testing*, Vol. 3, pp. 51–79, (1997), Academic Press, New York.
3. M. G. Silk, "The transfer of ultrasonic energy in the diffraction technique for crack sizing." *Ultrasonics* **May**, pp. 113–120, (1979).
4. British Standard BS7706, "Guide to Calibration and Setting-up of Ultrasonic Time of Flight Diffraction (TOFD) Technique for Defect Detection, Location and Sizing of Flaws," British Standard Institution, 1993.
5. A. A. A. Scheerder, "Equipment Condition Monitoring Using TOFD—Experiences at DSM," in *15th WCNDT*, 2000, Roma. Available from http://www.ndt.net/article/wcndt00/papers/idn070/idn070.htm
6. A. Masnata and M. Sunseri, *Neural Network Classification of Flaws Detected by Ultrasonic Means*, **Vol. 29**, No. 2, pp. 87–93, (1996), NDT & E International.
7. M. H. S. Siqueira, "Long Distances Inspection Using Ultrasonic Guided Waves and Artificial Intelligence," D.Sc. Thesis, Department of Metallurgical and Materials Engineering, COPPE/UFRJ, Brazil, 2002 (in Portuguese).
8. F. W. Margrave, K. Rigas, D. A. Bradley, and P. Barrowcliffe, "The use of neural networks in ultrasonic flaw detection," *Measurement* **25**, pp. 143–154 (1999).

9. J. Kim, P. Ramuhalli, L. Udpa, and S. Udpa, "Multidimensional Signal Processing for Ultrasonic Signal Classification," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti, **Vol. 20**, pp. 595–602, (2001), American Institute of Physics.
10. R. Polikar, L. Udpa, and S. S. Udpa, "Incremental Learning of Ultrasonic Weld Inspection Signals," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti, **Vol. 20A**, pp. 603–610, (2001), American Institute of Physics.
11. S. Haykin, *Neural Networks, A Comprehensive Foundation*, 1994, Macmillian College Publishing.
12. P. D. Wasserman, *Neural Computing Theory and Practice*, Van Nostrand Reinthold, 1989, New York.
13. B. P. Lathi, *Communication Systems*, (1968), Wiley, New York.
14. M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox for Use with MATLAB*, Version 1, 1996, The MathWorks Inc.
15. A. L. Graps, "An introduction to Wavelets," *IEEE Comput. Sci. Eng.*, **Vol. 2**, Number 2, pp. 50–61 (1995).
16. C. M. Bishop, *Neural Networks for Pattern Recognition*, 1995, Oxford University Press.
17. E. P. de Moura, R. R. da Silva, M. H. S. Siqueira, and J. M. A. Rebello, "Radiographics pattern recognition of welding defects using linear classifiers." *Insight*, **43** (10), pp. 669–674 (2001).