

Classification of Real Flaws Using Ultrasonic Signals

Carlos Correia, Roberto Otero, Gustavo Sulbarán

carlosc@fii.org, rotero@fii.org, gustavos@fii.org

Fundación Instituto de Ingeniería. Ministerio de Ciencia y Tecnología
Caracas, Venezuela.

Abstract.

The problem of flaw identification in fusion welded joints has been addressed using different methods and principles; nevertheless, we think this is a still open field, not solved completely yet. The deal amount of data that is possible to record with the new available technology today, makes automatic flaw classification an important engineering task. In the present work, we used two approximations to show other possibilities to tackle this unsolved problem. The first one is based on the properties of the covariance matrix, calculated from the level 2 approximation wavelet coefficients and obtained from a B-scan image with a defect. In the second procedure, the kurtosis and skewness of the continuous wavelet coefficients of five scales (taken from several ultrasonic signals) are used to create a feature space to train a Fisher Linear Classifier to discriminate common defects in welding joints. The behavior of the classifier was tested to differentiate discontinuities and the preliminary results are presented.

Keywords: *automatic flaw classification, ultrasonic signal processing*

1. Introduction.

The problem of ultrasonic signals identification and classification in non-destructive testing of welding joints has been explored in different ways and with different methods. The most common approach is the echodynamics identification that is based on the observation of changes shown on the amplitude of A-scans presentation as a function of position of the transducer in the piece's surface. This is the classical method used in level II courses for inspectors⁽¹⁾; nevertheless, it is relatively unreliable and it is very dependent of inspector's experience and skills.

Since the development of computers after 1980, and the implementation of a great deal of signal processing software, the use of the Fourier Transform Representation to evaluate different material properties and characteristics were extensively reviewed ⁽²⁻⁵⁾; nonetheless, the frequency representation was not appropriate enough to achieve the signal classification task. Nowadays computers and software allows high velocity and a lot of operations for signal processing, it is possible to change the domain representation, transforming, filtering, decomposing, denoising, compression and many others unsuspected operations. At the same time, statistical and neural network routines have been created and extensively used to different applications^(6,8). Neural Networks have proved to be very efficient in recognizing different real flaws using ultrasonic signals ⁽⁹⁾. Generally, the supervised methods have shown better results; however, the central problem for a good identification is not the method of classification by itself, the behavior of a method is a function of the quality of the selected features to represent the signal. In the present work, the continuous (CWT) and discrete wavelet transform (DWT) are used. The DWT is used to compress a B-scan image and the covariance matrix is presented as flaw characteristic representation. On the other hand, the kurtosis and skewness of the coefficients related to a set of scales of the CWT are selected as features and employed to train and test a Fisher Classifier. Four of the most commons welding discontinuities signals, acquired by means of pulse echo technique, namely *pores cluster, non metallic inclusions,*

ID connected crack and lack of fusion, are used to represent the main discontinuities found in welds. The pores cluster and nonmetallic inclusions are considered as volumetric flaws; while, ID connected crack and lack of fusion are planar flaws. Both were used to train a statistical classifier to separate planar from volumetric flaws.

2. Theoretical Review.

Wavelets

The wavelet transform is a useful tool that allows many operations over the signals, the most common, denoising and compression⁽¹⁰⁾. This transformation is essentially a representation of the correlation between the signal and a special function known as wavelet that is scaled and translated along the signal. The operation is synthesized in the following expression:

$$C_w(a, k) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) w\left(\frac{t-k}{a}\right) dt \quad (\text{Eq. 1})$$

Where a is the scaling factor, k is the translation, t is the time, $f(t)$ is the time domain function of interest, and C_w are the coefficients for each value of scale and translation. The continuous wavelet transform is a similar operation to view the signal at many different levels of magnification. It can be seen too, as a projection of the signal in a space of a higher dimension, where the spatial frequency components are easily detected.

The wavelet transform can be operated with dyadic scales and translations, down sampling the coefficients after each step obtaining the discrete wavelet transform. This is an extremely useful tool for compression. The level of decomposition used, produce a set of coefficients related with a frequency band in the signal. The higher level of decomposition the fewer coefficients are produced. The scaling and wavelet functions operate recursively over the signal producing approximation and details coefficients.

Statistical Functions & Operators

In the next lines some classical definitions are presented⁽¹¹⁾. The *variance* is given by:

$$S_j = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n} \quad (\text{Eq. 2})$$

Where x_{ij} is the value i of a set of values of j variable and \bar{x}_j is the mean of these values, n is the number of elements for the j variable.

Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data are spread out more to the left of the mean. If skewness is positive, the data are spread out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero. The Skewness is given by:

$$A_j = \frac{1}{n} \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^3}{s_j^3} \quad (\text{Eq. 3})$$

Kurtosis is a measure of how outlier-prone a distribution is. The more outlier-prone present greater values and vice versa. The kurtosis is given by:

$$K_j = \frac{1}{n} \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^4}{s_j^4} \quad (\text{Eq. 4})$$

The *covariance matrix* is a square and symmetric matrix that has the variances of each variable in its diagonal and the covariance within the variables outside the diagonal. A very common way to calculate this matrix is by means of the following operation:

$$Cm = \frac{1}{n} \overline{X}' \overline{X} \quad (\text{Eq. 5})$$

Where \overline{X} is data centered matrix obtained subtracting the mean of each variable from the data matrix, the notation (') represents the transposition operation and n is the number of elements observed.

Fisher Linear Discriminant

For the case of two sets of data, each one built with elements $x_i \in \mathfrak{R}^n$, and labeled as $y_k \in \{1,2\}$, the associated training vector is as $T = \{(x_1, y_1), \dots, (x_l, y_l)\}$. The Fisher Linear Discriminant algorithm uses the training vector to create a plane that divides the space in two regions, one associated with the class 1 and the other with the class 2. To find the appropriate plane, the algorithm computes the vector that maximizes the class separability function⁽⁷⁾.

3. Materials and Methods.

The procedure followed in the present work is schematized in Figures 1 and 2.

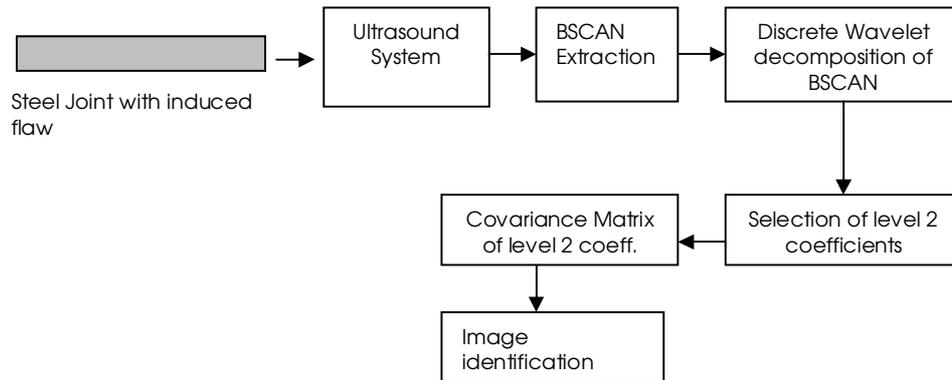


Figure 1. Schematic representation of the procedure to obtain a covariance matrix image from the level 2 discrete wavelet coefficients.

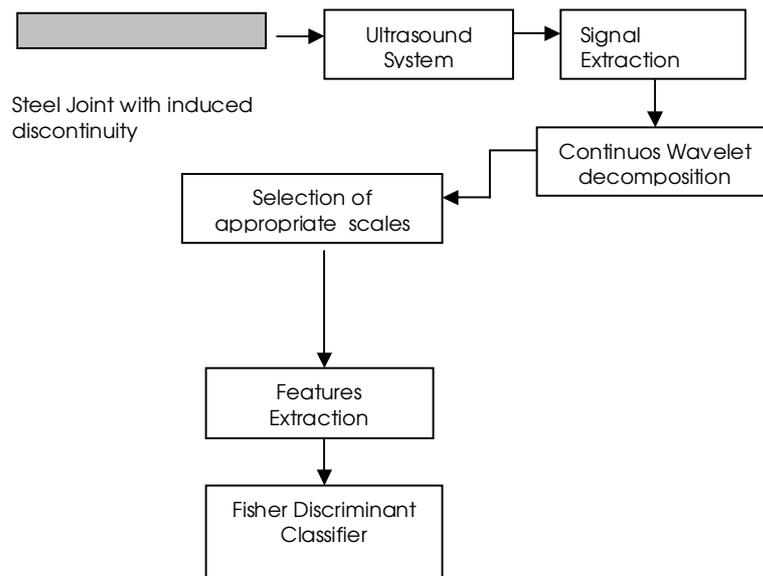


Figure 2. Schematic representation of the procedure to develop the Fisher Linear Discriminant Function.

The shear wave signals were obtained from different steel samples (10 mm thick) with induced flaws, from a kit manufactured by SONASPECTION®. The selected flaws were: (1) pores cluster, (2) non-metallic inclusion, (3) internal surface connected crack and (4) lack of fusion.

The radiography of each discontinuity is shown in Figures 3 and 4 just for graphical purposes.

In the first procedure (Figure 1), the level 2 coefficients of each A-scan were obtained to work with a reduced dimension. The new reduced B-scan representation was used to compute the covariance matrix. The images of the covariance matrix were obtained from a set of 50 B-scans experimentally obtained from each kind of flaw. The resultant patterns were analyzed.

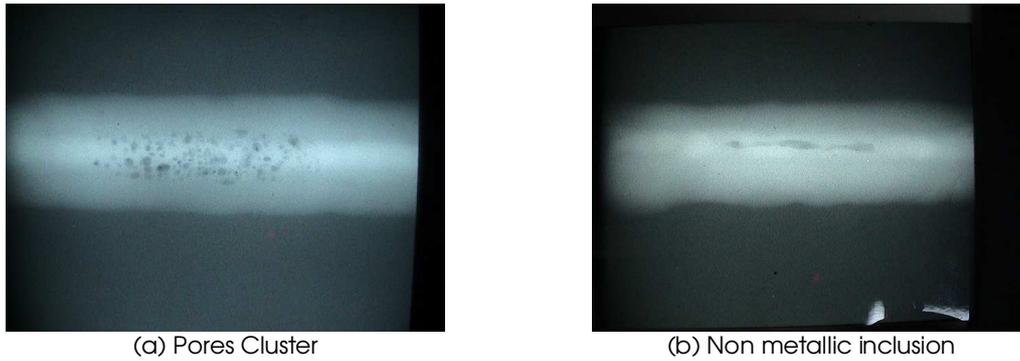


Figure 3. Volumetric flaws. (a) Pores Cluster, (b) Non metallic inclusion.

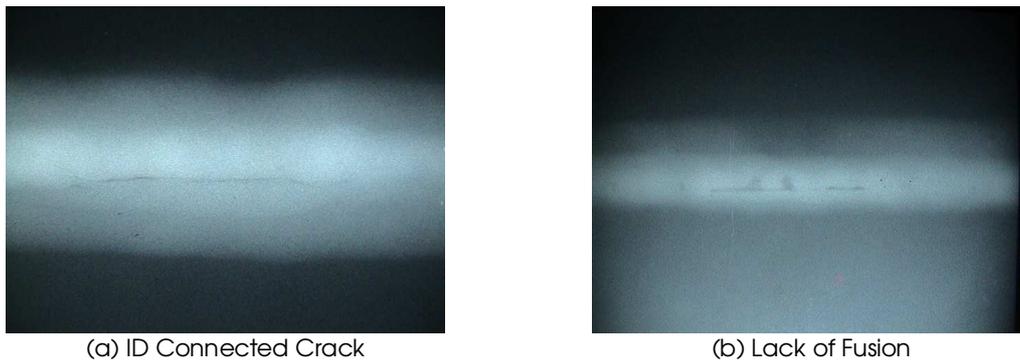


Figure 4. Planar Flaws. (a) ID Connected Crack, (b) Lack of Fusion

In the second procedure (Figure 2), 250 signals were recorded from each type of reflector using commercial ultrasonic equipment. The refracted angle was 60° shear wedge and the transducer was a Krautkramer Branson®, high damped (alpha series) with a frequency of 5 MHz. Those signals were obtained with a slight variation in the position of the transducer on the surface of the sample to introduce different angular positions between the flaw and the ultrasonic beam. The continuous wavelet transform was obtained for each signal using the Daubechies 4th degree wavelet. The decomposition was made using 48 scales, and the coefficients from the scale 15 to 20 were selected for all the defects. The kurtosis and skewness were calculated for each scale and applied as features of the signals. In the features space, the signals were described by a set of five values in \mathfrak{R}^2 of the form:

$$O_i = (s_i, k_i) \quad (\text{Eq. 6})$$

Where s_i, k_i were the skewness and kurtosis of each selected scale.

The Fisher Linear Discriminant function was built with a set of 100 signals, and 150 signals were used to test the classifier.

The following characteristics of the excitation pulse were used:

- Double spike, double polarity
- Pulse length of 15 μs
- Maximum peak to peak voltage 130 V

4. Results and Discussion.

Procedure 1. Covariance Matrix Image.

A typical B-scan obtained under the conditions established in the Section 3, is presented in Figure 5. As the ultrasonic beam scans the volume of interest the signals are collected and finally displayed. In this B-scan format, the image shows a center region with the flaw presence, and the rest without relevant indications. A compressed B-scan image is shown in Figure 6.

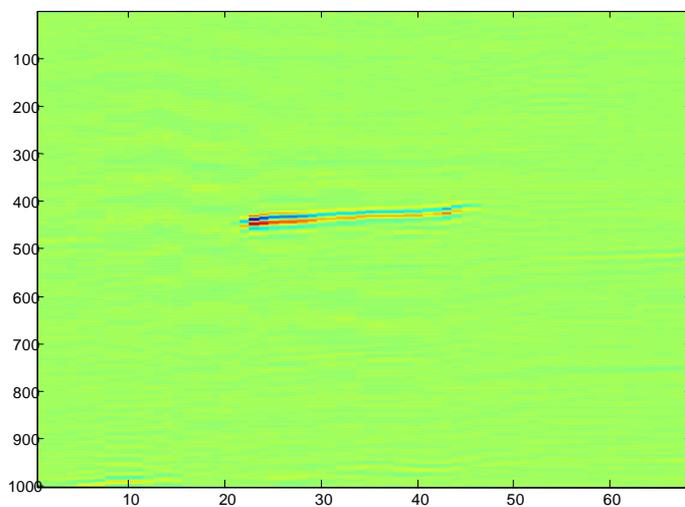


Figure 5. B-scan image obtained from the test-piece with a side bevel lack of fusion.

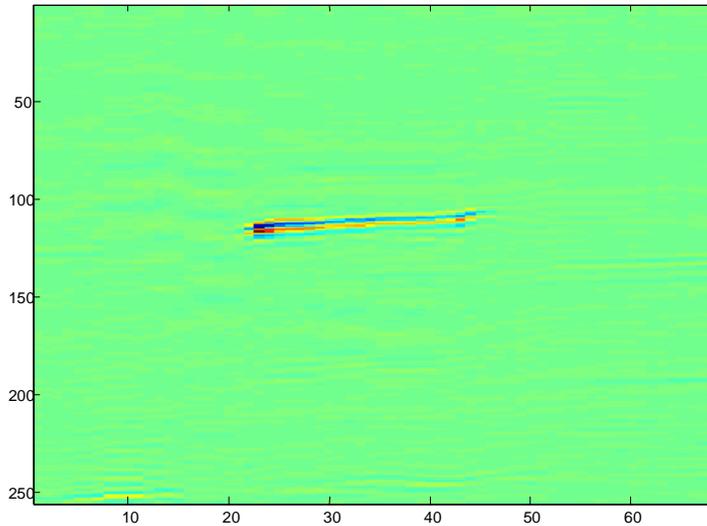


Figure 6. B-scan compressed image obtained from the test-piece with a side bevel lack of fusion.

Both images in Figure 5 and 6, contains visually the same information. The presence of a reflector is clearly presented at the center of the image, approximately; but the image in Figure 6 has only 26 % of points compared to Figure 5, with a compression of 74 %.

In a multivariate system there are a set of elements and a related set of variables measured over each element. This array is classically described by a data matrix with the elements located in the matrix's rows, and the variables in the columns. In the present work the B-scans were arranged with vertically A-scans. This system can be understood as a multivariate problem, with the difference that there is only one variable (the signal amplitude), but measured in different positions. The variable is always the amplitude for a given time value, but the amplitude changes as the transducer scans the surface of the sample. In other words, this is a spatial approximation of the ultrasonic signal evolution.

With this array, the covariance matrix will show in the diagonal the variance of each B-scan row, the equivalent to the variance of spatial amplitudes. Outside the diagonal, the covariance matrix will show the correlation between each pair of spatial amplitudes. The image of covariance matrix and the corresponding diagonal elements, could be used as a reference pattern associated with the nature of the flaws spatial amplitude distribution. This approach can exploit the spatial features instead of time or frequencies characteristics. The general appearance of covariance matrix for the four types of reflectors and the variance values of spatial amplitudes are presented in Figures 7 to 10.

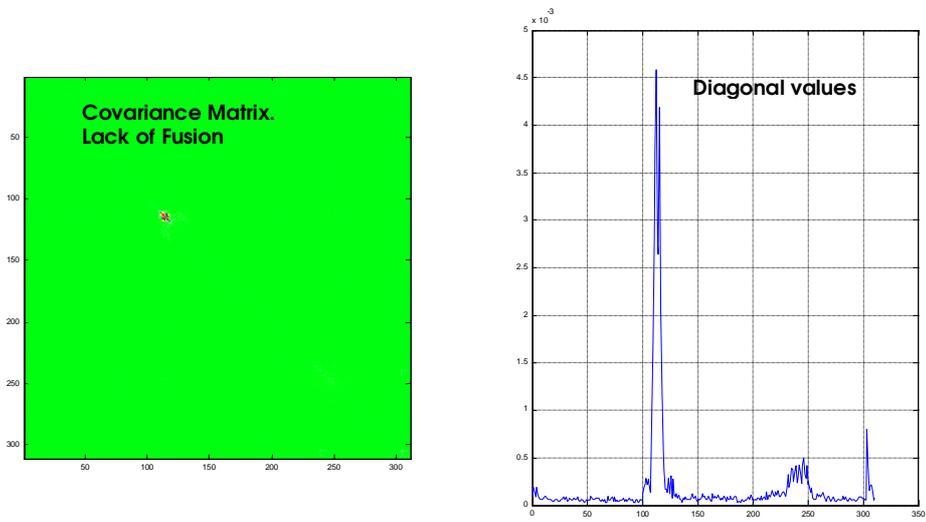


Figure 7. Lack of Fusion. Covariance Matrix of 2 level decomposition DWT, and the diagonal values.

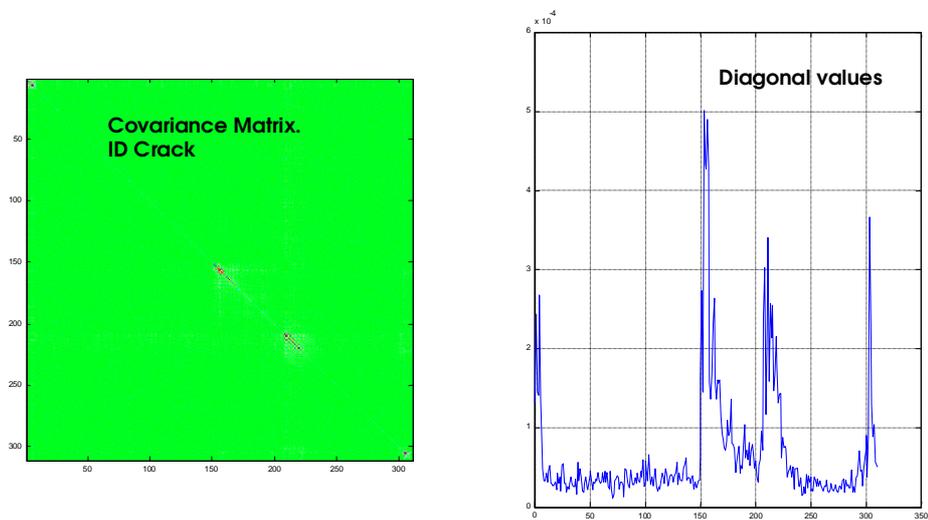


Figure 8. ID connected crack. Covariance Matrix of 2 level decomposition DWT, and the diagonal values.

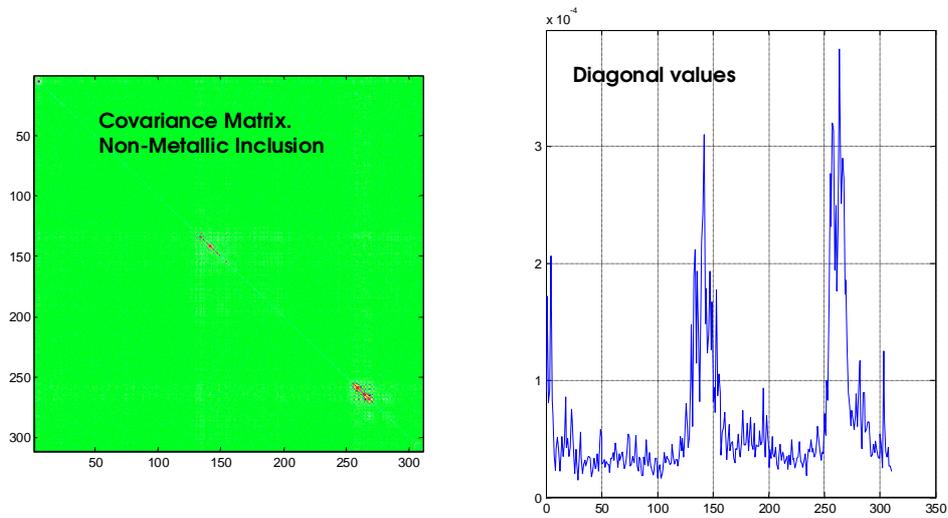


Figure 9. Non-metallic inclusion. Covariance Matrix of 2 level decomposition DWT, and the diagonal values.

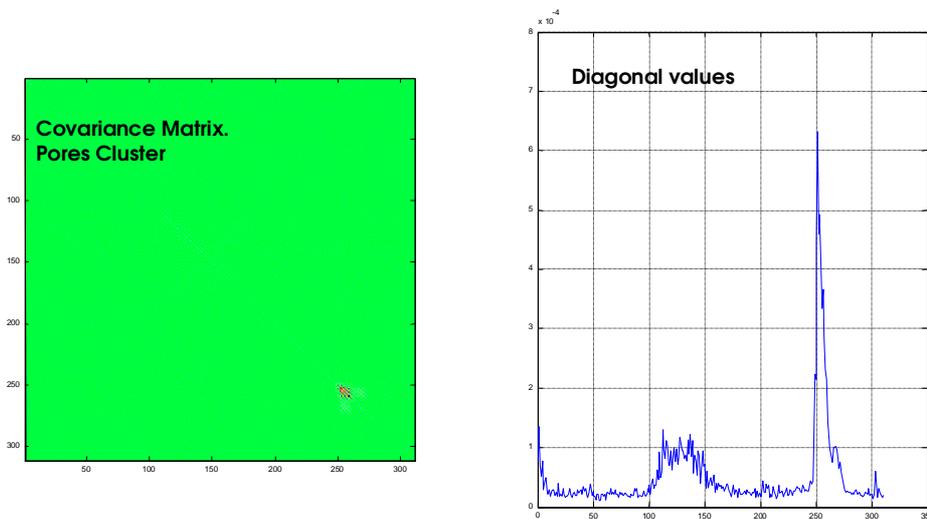


Figure 10. Pores Cluster. Covariance Matrix of 2 level decomposition DWT, and the diagonal values.

As can be seen in Figure 7, the lack of fusion, shows a region of strong variance concentrated at the center of the covariance matrix and a secondary region of low variance at the end right corner. The geometrical situation of this reflector seems to produce a strong concentration of energy in a specific region of the signals. The second

region with a mitigated variance (second peak in the diagonal) at the image right corner, indicates the presence of perturbations in the signal produced perhaps by geometrical reflectors.

The covariance matrix for the ID connected crack in Figure 8, shows two high variance regions (the two peaks in the diagonal plot). These two peaks so close one to another, could be produced by the toe and the tip crack signals. The distance between them are correlated with the crack height, and it is important to take into account this value because the signal coming from the non metallic inclusion, in Figure 9, have the same pattern, but differs essentially in the peak separation distance.

In Figure 10, the covariance matrix for the pores cluster, shows an inverted variance distribution in comparison with the lack of fusion in Figure 7. The first peak (the first zone of high variance) is of less amplitude than the second one. This is a characteristic of pores signals. Generally the variance of pores signals is low in comparison with other signals flaws. Many reflectors acting as emitting sources have the effect to spread out the energy, showing a homogeneous amplitude distribution along the signal.

As can be shown in Figures 7 to 10, the lack of fusion and pores cluster are clearly different observing the shape of covariance matrix; nevertheless to distinguish between non metallic inclusion and ID connected crack it was important to take into account the distance between the peaks of maximum variance. The patterns shown in Figures 7 to 10 were observed in a proportion summarized in Table 1.

Table 1. Percentage of coincidence with Figures 7 to 10 from all experimental B-scans data.

Side Bevel Lack of Fusion (Figure 7)	ID Connected Crack (Figure 8)	Non-metallic inclusion (Figure 9)	Pores Cluster (Figure 10)
90%	80%	70%	100%

The high values in the diagonal indicates the variance of variables have higher values in comparison with the covariance between variables. The presence in the matrix of a middle zone with high values indicates this is the zone of B-scan containing useful information. The covariance matrix image or the associated diagonal could be used to train an automatic classifier and it could be used also as a visual tool to the inspector in future instrumentation developments.

Procedure 2. Linear Fisher Classifier.

As it was mentioned in the Section 3, 250 signals were obtained from each kind of reflector, a set of 100 signals were used to train a Linear Fisher Classifier and 150 signals were used to test the classifier. The selected features were the kurtosis and skewness from a set of CWT coefficients from the same five appropriated scales in each case, showing a good signal representation. The results are summarized in Table 2, and some examples are shown in Figures 11 to 13.

Table 2. Results obtained with a trained Fisher Linear Classifier.

Group 1	Group 2	Success Percentage	Figures
Pores Cluster	Lack of Fusion	100 %	Figure 11
Non-metallic Inclusion	ID Connected Crack	60 %	Figure 12
Non-metallic Inclusion	Lack of Fusion	70 %	Figure 13
Pores Cluster	Non-metallic Inclusion	100 %	-
Pores Cluster	ID Connected Crack	100 %	-

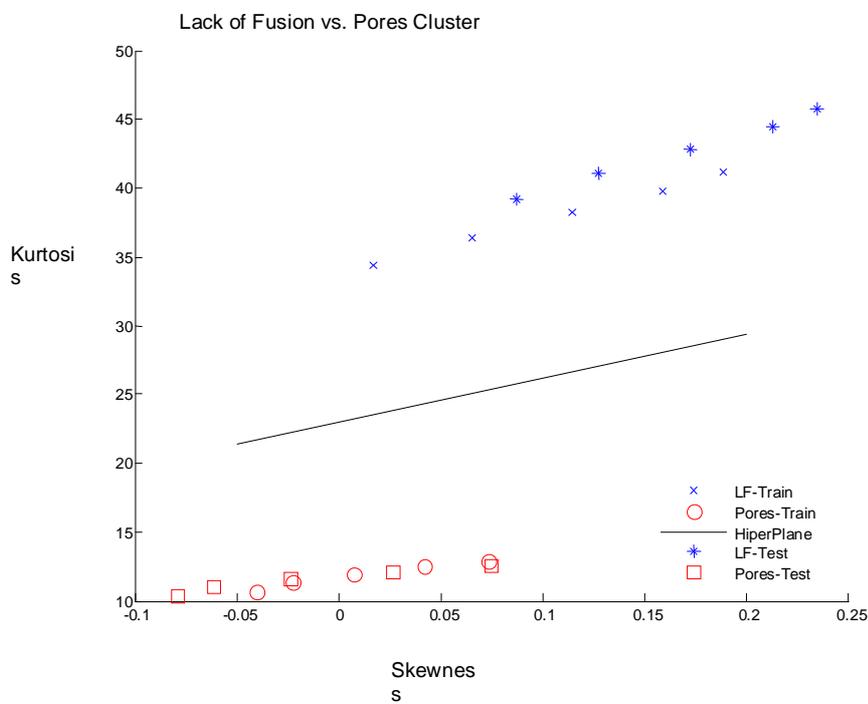


Figure 11. Classification of a test group in a system Pores cluster vs. Lack of Fusion (LF), SPRT00L-image.

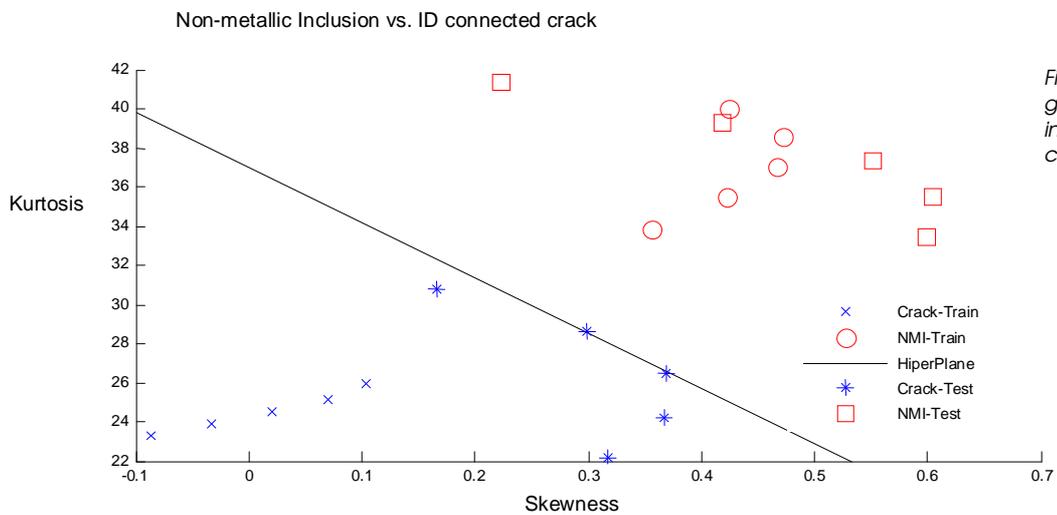


Figure 12. Classification of a test group in a system Non-metallic inclusion (NMI) vs. ID connected crack . SPRT00L- image.

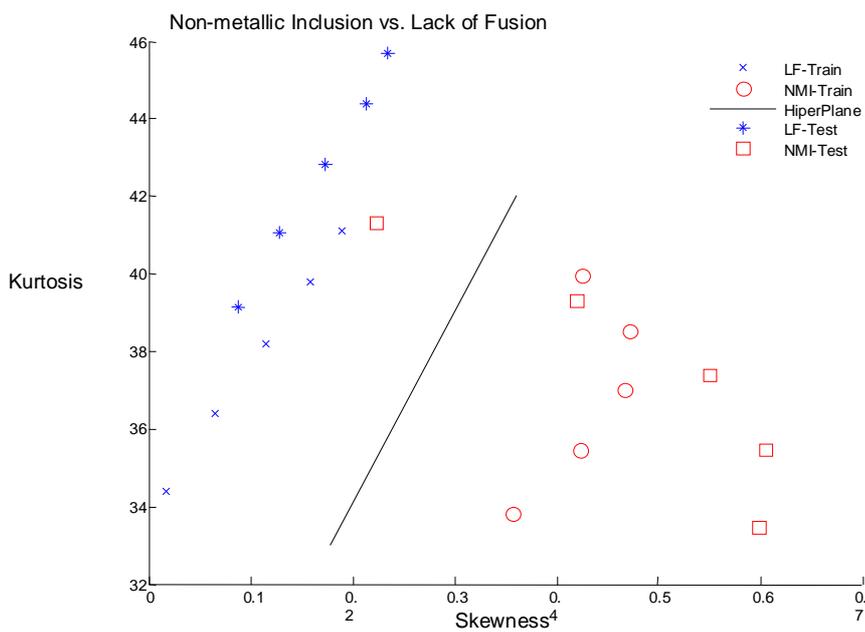


Figure 13. Classification of a test group in a system Non-metallic inclusion (NMI) vs. Lack of fusion (LF). SPRT00L- image.

As can be shown in Figures 11 to 13, kurtosis could be a good feature to classify pores cluster from other discontinuities. A possible explanation could be in terms of how the energy is described in the characteristic pulses. It seems the energy in the pores cluster signals is more spread out than others discontinuities. In those cases, the kurtosis value had a strong tendency to be low.

The success of classification, for non-metallic inclusions, lack of fusion and ID crack was relatively low because the selected features (kurtosis and skewness) have closely values for these discontinuities.

5. Conclusions.

The matrix covariance is a tool that exploits the A-scans spatial characteristics and shows interesting properties which can be used as features for an automatic signal classifier or even as a visual guide to the inspector. The most easily revealed flaws using this approximation were pores cluster and lack of fusion. The non-metallic inclusion and ID connected crack could not be easily differentiated using the matrix covariance approximation. The identification of pores cluster was also well determined using the Fisher Linear Discriminant function (our second procedure). The experimental results obtained for lack of fusion, non-metallic inclusion and ID connected crack were limited; however, the 100% of pores cluster signals were positively identified and this is encouraging for future approximations on automatic flaw classification, using TOFD images from raw data, for example.

6. Acknowledgements.

We want to acknowledge to the Engineering Institute Foundation of Science and Technology Ministry from Venezuela, which supported this research. We would also like to express our gratitude to Mr. Robert Hutchinson, from Central University of Venezuela, for his help in the English version of this paper.

7. Bibliography.

- (1) General Dynamics, Non Destructive Testing, Ultrasonic Volume III, 1981.
- (2) J. Saniie and N.M. Bilgutay, "Quantitative Grain Size Evaluation Using Ultrasonic Ultrasonic Backscattered Echoes", Journal of The Acoustical Society of America, Vol. 80, pp. 1816-184, Dec., 1986.
- (3) T. Wang, J. Saniie, and X. Jin, "Analysis of Low Order Autoregressive Models for Ultrasonic Grain Signal Characterization", IEEE Trans. UFFC, Vol. 38, pp. 116-124, Mar., 1991.
- (4) J. Saniie, T. Wang, and N.M. Bilgutay, "Analysis of Homomorphism Processing for Ultrasonic Grain Signal Characterizations", IEEE Trans. UFFC, Vol. 36, pp. 365-375, May., 1989.
- (5) J. Saniie, "Analysis of Order-Statistic CFAR Threshold Estimators for Improved Ultrasonic Flaw Detection", IEEE Trans. UFFC, Vol. 36, pp. 618-627, Sept., 1992.
- (6) R. Otero, C. Correia, C. Ruiz ; Statistical Characterization from Ultrasonic Signals Using Time-Frequency Representation. www.ndt.net/article/v08n05/otero/otero.htm , 2003.
- (7) Vojtech F., Hlavac V. Statistical Pattern Recognition Toolbox for MATLAB (SPRTOOL). User's Guide. http://cmp.felk.cvut.cz/cmp/cmp_software.html
- (8) Duin R.P.W., Juszczak P., et al, A Matlab Toolbox for Pattern Recognition (Prtools4), <http://prtools.org/>
- (9) Polikar R., Udpa L., Udpa S. and Taylor T., Frequency Invariant Classification of Ultrasonic Weld Inspection Signals, IEEE Trans. UFFC, Vol. 45, No. 3, pp. 614-625, 1998.
- (10) G. Strang, T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, USA, 1996.
- (11) Peña D., Análisis de Datos Multivariantes, Edit. Mc Graw Hill, España, 2002.