PULSED EDDY CURRENT AND TIME-FREQUENCY ANALYSIS IN DETECTING THE CORROSION DISTRIBUTION IN A MULTILAYER ALUMINUM PLATE

A. Lakis(*)1, S. M. Slaeh Hosseini1, M. Toorani2, S. Sassi3

1Mechanical Engineering Department, Ecole Polytechnique of Montreal, Montreal (Quebec), Canada
2Mechanical System Engineering, Conestoga College, Cambridge (Ontario), Canada
3Department of Mechanical and Industrial Engineering, Qatar University, Doha, Qatar

(*)Email: aouni.lakis@polymtl.ca

ABSTRACT

Pulsed Eddy Current (PEC) is a non-destructive testing method used to detect corrosion and cracks in multilayer aluminum structures which are typically found in aircraft applications. Corrosion and metal loss in thin multi-layer structures are complex and variable phenomena that diminish the reliability of pulsed eddy current measurements. In this paper, the Pulsed Eddy Current signals are processed to improve the accuracy and reliability of these measurements. PEC's results (time domain data) are converted by Time-Frequency Analysis (Rihaczek distribution) to represent data in three dimensions. The Time-Frequency approach generates a large amount of data. Principal component analysis (PCA) is applied as feature extraction to reduce redundant data to provide new features for classifiers. K-means Clustering and Expectation-Maximization are applied to classify data and automatically determine corrosion distribution in each layer. A set of tests has been carried out to detect the defects located at different depths and different positions of a multilayer aluminum plate.

Keywords: Multilayer Aluminum Structures, Pulsed Eddy Current, Time-Frequency.

INTRODUCTION

Pulsed Eddy Current (PEC) is one of the new methods that provide accurate and reliable measurements when used to inspect aircraft body components for corrosion and cracks (Moulder et al. 1996), (Sophian et al. 2001), (Tian et al. 2005), (Giguere et al. 2001), (Lee et al. 2007), (Edwards et al. 2009), (Sicard et al. 2008), and (Abidin et al. 2009). Pulsed eddy current can provide accurate and reliable measurements when used to inspect the aircraft body components for corrosion. The pulsed eddy current technique as one of the non-destructive tests (NDT) is capable to characterize the defects in the components without removing them from service, or causing damage to the sample during the testing process. The broadband nature of pulsed eddy current allows more penetration than other NDT methods. Response signals of pulsed eddy current provide good information describing the condition of interior structures. Corrosion, cracks and other defects can be detected by applying signal processing methods on PEC signals. This paper outlines the applicability of the time-frequency analysis method to treat Pulsed Eddy Current (PEC) data to detect the corrosion defects at different depths in different positions of a multilayer structure. The PEC data are converted from time domain to time-frequency domain and then the data are analyzed in an enhanced form to detect the defects with various sizes in each layer. The time-frequency analysis method provides a three-dimensional representation of a signal and allows detection of the size and location of defects in a multilayer system. Maximum variances of PCA (Principal Component
Analysis) are selected as features, and a classifier is used to separate and classify these new features. A pulsed eddy current testing system (consisting of a function generator, magnetic reflection probe, preamplifier, digitizer converter card and a computer system with signal processing software) has been designed to detect the defects of a multi-layer aluminium sheet.

**THEORY**

**Pulsed Eddy current**

A Pulsed Eddy Current system usually consists of four main parts which are; function generator, probes, preamplifiers and an analog-to-digital converter (A/D). The function generator creates a square wave which excites the probe. Probe excitation produces an electromagnetic field that affects the conductive surface and induces a current on the surface called an eddy current. This eddy current generates an electromagnetic field which penetrates the metal bulk. The electromagnetic responses of the eddy current electromagnetic field contain information describing features of inner structures. Its amplitude is related to type of material, metal thickness, defects and cracks that exist in the material. The electromagnetic response is received by another probe. This Pick-Up probe absorbs two types of electromagnetic fields; the electromagnetic field from the driver probe and the eddy current’s electromagnetic response.

**Time-frequency analysis**

Common results of the pulsed eddy current technique provide a set of data in the time domain. Noise sources (less amounts of defects, lift-off and interlayer gap), cause close and noisy curves in this domain. It is important to use alternative methods for multilayer structures when defects and corrosion occur on the surfaces. Time-Frequency analysis methods provide three-dimensional representations of signals (time, frequency and amplitude) to detect defects with higher accuracy. In this paper, a Rihaczek distribution is applied in a time-frequency analysis method to present the data in three dimensions. During this study, different types of time-frequency analysis methods were implemented and applied. The Rihaczek distribution provides best results of three-dimensional representation and also the most clear defect separation during training of classification. Using the real part of energy is the main advantage of the Rihaczek Distribution over other time-frequency analysis methods. The presence of cross-terms in time-frequency analysis results reduces the efficiency of signal processing. Applying a Kernel function to the Rihaczek distribution reduces cross-term effects. Therefore the cross-term effect in Rihaczek distribution images is less than Wigner-Ville distribution images and other time-frequency analysis methods.

**Rihaczek Distribution**

In Time-Frequency Analysis, the Rihaczek distribution allows the energy of a complex deterministic signal over finite range of time and frequency to become infinitesimal. The obtained result is called a complex energy distribution. The interaction energy between a signal $x$ restricted to an infinitesimal interval $\delta_t$ centered on $t$, where $x$ passes through an infinitesimal band pass filter $\delta_f$ centered on $v$, can be approximated by the following expression:

$$\delta_t \delta_f [x(t)X^*(v) e^{-j2\pi vt}].$$ (1)
Equation (1) leads to the following definition of the quantity of energy density function:

\[
R_j(t,v) = x(t)X^*(v)e^{-j2\pi vt}
\]  

(2)

The Rihaczek distribution can be interpreted as a complex energy density at point \( (t,v) \) with elements of Cohen's class written as \( f(v,\tau) = e^{j2\pi v\tau} \). The Rihaczek distribution is a bilinear time–frequency distribution (TFD) and is a member of Cohen’s class. It is covariant to shifts in time and frequency. In this paper, a Rihaczek distribution is used for discrete-time, non-stationary, harmonizable, zero-mean time series pulsed eddy current signals. Many of time-frequency analysis methods abandon any sort of analogue to physical phenomena with negative and complex values. Complex energy density in the Rihaczek distribution overcomes the problem of missing phenomena by using only the real part of the energy and trends the energy to the standard and real amount.

**Feature Extraction and Classification**

Time-frequency methods provide very large amounts of data and it is necessary to remove redundant data and reduce the dimensions of the data. Principal component analysis (PCA) is a method that creates new features and reduces the size of data.

**Principal Component Analysis (PCA)**

Principal Component Analysis (PCA) is also known as the Karhunen Loeve (KL) transforms or the Hotelling transform. It is based on factorization techniques developed in linear algebra. PCA uses factorization to transform data according to statistical properties to optimize and reduce the amount of redundant data. This data transformation is particularly useful for classification and compression. The main advantage of PCA is generation of independent features. In PCA processing, covariance of input data \( \Sigma_X \) is calculated as:

\[
\Sigma_x = E\{(x - \mu_x)(x - \mu_x)^T\}.
\]  

(3)

Eigenvalues \( \lambda \) and Eigenvectors \( W_i \) are identified using covariance of the time-frequency results \( \Sigma_X \).

\[
(\lambda I - \Sigma_x) = 0.
\]  

(4)

\[
(\lambda I - \Sigma_x)W_i = 0.
\]  

(5)

Where, \( I \) is an identifier matrix of the same order as \( \Sigma_X \). Transform features \( C_Y \) are obtained using \( C_XW^T \). The new features are linearly independent.
**K-Means Clustering and Expectation-Maximization Algorithm (EM)**

K-means is one of the unsupervised learning algorithms that solves the clustering problem. The K-means algorithm is based on the use of squared Euclidean distance as the measure of dissimilarity between a data point and a prototype vector. K-means separates \( n \) observations into \( K \) clusters, for which each observation (data point) belongs to the cluster with the nearest mean. The K-means algorithm consists of two major parts; (1) review all the data and assign data to the closest selected centers (centroid), (2) make a loop and recalculate the center points to find the nearest centroids. The K-means algorithm partitions a set of \( n \) observations \((x_1, x_2, ..., x_n)\) into \( k \) clusters \( S = [S_1, S_2, ..., S_k] \) to minimize the square error function:

\[
\arg\min \sum_{i=1}^{k} \sum_{x_j \in S_i} |x_j - \mu_i|.
\]

where; \( \mu_i \) is the mean of \( S_i \).

The Expectation-Maximization (EM) algorithm is a powerful probabilistic classifier method for finding the maximum likelihood solution parameters for unobserved latent variables \( S \). A set of \( n \) observations and a vector of unknown parameters \( \theta \) are used to calculate the maximum likelihood.

\[
L(\theta, n) = p(n|\theta) = \sum_{S} p(n, S|\theta).
\]

EM is an iterative method which alternates between performing an expectation E step and a maximization M step. The E step computes the expectation of the log-likelihood and evaluated responsibilities (posterior probabilities) for the latent variables.

\[
Q(\theta|\theta^{(t)}) = E_{S|\theta^{(t)}}[\log L(\theta; n, S)].
\]

The M step (Maximization) re-estimates parameters (means, covariance and coefficients). The expected log-likelihood is achieved in the E step.

\[
\theta^{(t+1)} = \arg_{\theta} \max Q(\theta|\theta^{(t)}).
\]

**RESULTS**

**Time Domain Signals**

A two-layer perfect (without defect) and imperfect (having some defects located at different depths and locations) plate has been used in experimental part of this study. A 1-kHz pulse was used to excite the probe and the response signals were digitized. Figure 1 shows the detected pulse response for defects ranging from 30 to 250µm respectively for each layer. The amplitude of the response signals increases as the thickness of the defects increase, i.e., the 30µm sample has a minimum amplitude whereas the 250µm defect sample has maximum
amplitude. At each location of the defects on the same layers, an increase in the thickness of the defect causes a decrease in eddy current field and therefore the pick-up coil (probe) receives more electromagnetic field. In Figure 1, acquired signals from three positions with different amounts are presented. These signals contain information from surfaces between layers and under layers. The signals are from defects placed on the bottom of the bottom layer (BOB), the top of the bottom layer (TOB) and the bottom of the top layer (BOT) at 30µm, 50µm, 70µm, 100µm, 150µm, 200µm and 250µm depths on each surface. Acquired signals are too closed. The signals are subtracted against reference signal which come from double layer without defects to show the metal losses variations clearly.

![Figure 1 - Pulsed Eddy Current results from seven different depths (metal-loss defects) placed in three positions](image)

Figure 2 shows the results of subtraction between the reference signal (double layer without defects) and a defected surface in a different layer ((a) bottom of top layer, (b) top of bottom layer and (c) bottom of bottom layer). It can be observed that the response signals for 3% to 10% are too close. These time-domain signals consist of data from interior layers, but the fact that the curves are so close causes difficulty in extracting reliable information from them. In this case, time-frequency methods are applied to create three dimensions for representation of data. A-Scan signals in the vertical direction are enhanced from each defected sample case (30µm to 250µm). Results of each sample case were acquired using 5-times detection (in total 105 signals).

**Time-Frequency Analysis**

The results of Time-Frequency Analysis are the Rihaczek Distribution of Pulsed Eddy Current signals. The output of Time-Frequency Analysis for each Pulsed Eddy Current signal is an n×n matrix and represent by an image. The Time-Frequency Analysis method provides a three-dimensional representation of a signal and allows detection of the size and location of defects in a double-layer system due to close-amount defects and noise resources such as interlayer gap and probe lift-off. Figure 3 shows the time-frequency representations of samples with defects located in the bottom-of-bottom layer. Deeper defects cause increased
penetration of eddy current magnetic fields and the probe receives more eddy current electromagnetic field. This effect causes a decrease in the amplitude of deeper defects. It also shows that the time-frequency analysis output for defected samples placed on the bottom-of-bottom layer (BOB). Defects were 3% to 25% of the thickness, respectively 30 micrometers to 250 micrometers. Increasing the depth of metal loss caused a shift in intensity from a narrow to wide state. In cases of more metal loss where depth of defects increase, the intensity magnifies, which shows that the strength of the eddy current magnetic field decreases at the top of the defected surfaces. Similarly, the same experiment has been carried out for the defects located at the top of the bottom layer and also the bottom of top layer, but not shown in this paper.

Fig. 2 - Subtraction result of defects with different metal loss. (a) Bottom of Top layer (BOT). (b) Top of Bottom layer (TOB). (c) Bottom of Bottom layer (BOB)

The comparison between same size metal losses that placed in different depth from surface show some critical issues that used in feature extraction and classification. Results of time-frequency analysis show when the defects are closer to surface, they have narrower shape against the same size defects that placed in deeper from surfaces. In other hand, maximum intensity of the defects that are closer to surface is more than the same size defects that placed deeper from surface.

Increasing the thickness of the defect in each surface layer causes an increase in the intensity and size of the response signal. In each defected layer, there is a narrow intensity in the range of 130µsec and this can be used as a visual specification for defects of less than 5% metal loss. Increasing the depth of defects causes a shift along the time axis. For example, for 5% defects the time of the peak is 130µsec, but in the case of a 25% defect, a wide peak is observed at 220µsec. The amplitude and shape (narrow or wide, extended in time axis or both, place of the peak) of the peaks defines the location and size of the defect. They are therefore
used as key features during feature extraction and classification. For each layer, feature extraction and classification were applied to 5 defect ranges: 0-5%, 5-10%, 10-15%, 15-20% and 20-25%.

**Feature extraction and classification**

The output of Time-Frequency Analysis for each Pulsed Eddy Current signal is an \( n \times n \) matrix which is represented by an image. The matrix is expressed as \( n^2 \) dimensional vectors. The value of each vector corresponds to intensity of each pixel of the Time-Frequency Analysis image. The image contains a large amount of redundant data. Principal Component Analysis (PCA) is applied for feature extraction to reduce the number of parameters and also to improve computational efficiency of the classifier.

PCA computes eigenvalues and eigenvectors of the covariance matrix. Eigenvectors are arranged in order of largest Eigenvalues. The first eigenvector has the direction of largest variance of data and it determines the direction of the most significant amount of energy. Principal Component Analysis retains the significant features (eigenvectors with largest eigenvalues) and leaves out some less expressive features. In this way, PCA reduces the number of parameters without loss of information.

![Fig. 3 - Time-Frequency results from defects placed on Bottom of Bottom (BOB) surface. (a) 3% metal loss. (b) 5% metal loss. (c) 7% metal loss. (d) 10% metal loss. (e) 15% metal loss. (f) 20% metal loss. (g) 25% metal loss](image-url)
These new features were used as input data for classification. In our research, a group of the 3 largest eigenvalues was selected for each class. Misclassification error is critical factor for each classification method. In this article, K-mean clustering and Expectation-Maximization have minimum misclassification error compared to other methods of classification. The misclassification errors of K-mean clustering and EM algorithms are calculated and the values of these errors are 3.11% and 1.78% respectively. The final results of classification using K-means clustering and Expectation-Maximization (EM) based on PCA output data for fewer than 10% and more than 10% are shown in Figures 4 and 5. In Figure 4, PCA output data are classified for samples with defect depths ranging from 30 to 100 micrometers where defects are placed on bottom of bottom (BOB), top of bottom (TOB) and bottom of top layers (BOT). Classification results are represented in 2-D where the first and second principal component have been chosen as the first and second dimensions of the data representation. In Figure 4, six separate clusters exist. These clusters belong to less than 5% metal loss and 5-10% metal loss in each layer. Results of 0-5% metal loss and 5-10% metal loss on the bottom of bottom layer are too close but can be separated by classifiers. Figure 10-(a) and 10-(b) show a map of classification for K-Means Clustering and EM algorithms respectively. Each circle (Fig. 5-(a)) belongs to each defected sample (0-5% metal loss or 5-10% metal loss) in three different layers. Defect distribution shows that low-size defects on the deep layer (BOB) are placed in the middle of the diagram. The defects which are placed on BOT and TOB are far from the center of the diagram. These effects depend on time frequency analysis results.

In Figure 5, PCA output data are classified for samples with defect depths ranging from over 100 to 250 micrometers where defects are placed on the bottom of bottom (BOB), top of bottom (TOB) and bottom of top layers (BOT). Figure 5-(a) and 5-(b) show maps of classification for K-Means Clustering and EM algorithms respectively. The results of the EM algorithm and K-means clustering are similar. For comparison, it is necessary to calculate the Euclidian distance of each unknown test point with each centroid to define minimum distance. The minimum distance of each unknown sample defines the cluster group in which the defected sample belongs.

![Fig. 4 - Classification of PCA output for defects less than 10% and 10% of thickness metal loss in three different positions (BOB, TOB and BOT). (a) k-mean clustering. (b) EM](image_url)
CONCLUSION

This paper reports the applicability of the time-frequency analysis method to treat pulsed eddy current data to detect corrosion defects at different depths in different positions in a multi-layer structure. Corrosion is a varied and complex phenomenon and metal loss on surfaces and between layers changes quickly. In many aspects, pulsed eddy current’s time domain signals are not able to show variation of metal losses on surfaces and cannot separate them correctly. In our approach, we convert PEC data from time domain to time-frequency domain and analyze the data in an enhanced form to detect defects of different sizes in each layer. Maximum variances of PCA were selected as features and a classifier was used to separate and classify these new features. The time-frequency method is a valuable tool to treat PEC data and is capable of detecting defects (such corrosion and cracks) in aluminum multi-layer systems when the size variation of the defect is approximately 30µm. This method can be used in aircraft industry applications to detect corrosion and metal loss in double-layer and multi-layer structures. In future studies, other time-frequency analysis methods will be used to compare the precision and reliability of the signal processing method when applied to overlap corrosion and defect detection in an aluminum multi-layer.

ACKNOWLEDGMENTS

This research was supported by the collaborative research and development grant (No. CRDPJ-335472-05) of Natural Sciences and Engineering Research Council of Canada (NSERC), Pratt & Whitney Canada, Bombardier Aeronautics and National Research Council of Canada (NRC)

REFERENCES


