TC
separata
01

Feature Extraction Methods for Classification of Steam Generator Tube Defects Using Self-Organizing Maps

R.N. de Mesquita[1,2], D.K.S. Ting[1], E.L.L. Cabral[2] and B.R. Upadhyaya[3]
[1]IPEN/CNEN-SP, São Paulo, Brazil
[2]University of São Paulo, São Paulo, Brazil
[3]The University of Tennessee, Knoxville
E-mail: rnavarro@net.ipen.br
Phone: (55) 11-3816-9399

Abstract

This work presents a new approach on feature extraction methods selection and classification techniques for steam generator tube defects using Eddy Current Test (ECT) data. A set of feature extraction methods is applied to different tube defect type ECT signals and each obtained feature vector is projected onto a bi-dimensional map obtained by a Self-Organizing Map (SOM) neural network. These maps show different clustering distributions that depend on the importance of each of the feature characteristics for a given defect type. An estimate of the characterization importance of each feature is made through the comparison between clusters on those maps. This method allows an optimal feature extraction selection for the defect type classification and posterior use in a fuzzy inference system.

Introduction

Steam generators are important components of pressurized water reactor (PWR) power plants. Their basic principle is heat transfer. The thousands of tubes that constitute the steam generator are interfaces between the coolant liquid of the reactor primary system, which contains the thermal energy generated by fission reactions, and the fluid on the secondary reactor system, which moves the power turbines. The steam generator efficiency is directly related to the power it can generate. This efficiency is affected significantly when a fraction of the steam generator tubes have to be plugged due to defects. The tube degradation occurs by different mechanisms such chemical corrosion, stress, vibration or a combination of these.

The Eddy Current Test has been used for many years as a non-destructive inspection technique on a periodic schedule for Steam Generator tubes (SGT) maintenance. The acquisition and processing of ECT data in nuclear power plants uses an automated and well-developed technology that control the ECT probe displacement and positioning in the tubes. The processing and visualization include commercial software that allows a database storage that can be analyzed off-line. Multiple teams of specialists perform this off-line analysis to make decisions about the plugging or repair of each of the SGT. This analysis takes into account economic and safety considerations and compares the actual data with a database of previous defects in similar conditions and similar power plants. Besides the large size of these databases, there is very little time for this analysis to be done as it generally occurs during a power plant shutdown for refueling and other scheduled maintenance.

The ECT analysis automation has been part of an effort to extend the Steam Generator mean life and raise the reliability of nuclear power plants in a predictive maintenance program. The imprecision of this diagnostic system implies a usual conservative tendency for making decisions about plugging the examined tubes. This conservative posture in making decisions necessarily leads to an extra loss on these equipment life and capacity, which may represent the premature need to replace the Steam Generator. The total maintenance cost, including power plant stoppage and steam generator replacement may run cost hundreds of millions of dollars.

Since the 1990's several methods have been developed for automated diagnostics systems to be used as auxiliary tools to ECT data analysis [1-4]. The automated diagnostics system depends on a reliable classifier algorithm. The classifier itself depends strongly on a well-organized knowledge database. A common difficulty to the development of any specialized system is the knowledge representation. In ECT signal classification, knowledge representation is particularly important due to the few and poor characteristic descriptions that can be found related to each tube defect type and its associated signals. The problem becomes worse if we consider that we have different design and characteristics depending on the nuclear power plant in which the ECT is performed.

There were few tries to classify ECT signals in their different defect types. Most of the contributions to ECT classification were directed to the defect depth estimation. Fuzzy inference systems using different signal characteristics were developed by Upadhyaya et al [5] with important results. Most of these results were obtained using pattern calibration signals. Other classification efforts [1,6] use different feature extraction methods.

This paper proposes a selection of feature extraction methods based on a new method that can estimate the contribution of each of this feature extraction methods to classification task. This selection is done based on SOM two-dimensional maps obtained for each of these feature extraction methods. This algorithm permits both visualizing the multi-dimensional configuration of this feature space and an output that can be used by a fuzzy inference system for SGT defects classification.

This procedure makes possible a classifier algorithm that does not take into consideration other related knowledge base information as defect localization in tube, excitation frequency channel, or which nuclear power plant the tube pertains to. This classifier algorithm will only use information that comes from the signal itself. This study can contribute to ECT signal modeling problem as it reveals the signal intrinsic features that are useful on their characterization.

Eddy Current Testing (ECT)

The basic principle of ECT is the relationship between the electromagnetic field generated by a probe and an electrically conductive material sample (Figure 1). The probe is composed of a coil, which is subject to an altemated current. The coil, thus, produces an alternating electromagnetic field (primary field) that induces a set of circular currents in the sample. The circular trajectories formed by these currents on the conductive material occur in planes perpendicular to the primary field. These are the so-called "eddy currents" and they generate a secondary electromagnetic field, which opposes the primary field. The secondary field results in a reduction on the overall magnetic flux inside the bobbin. The presence of defect or material non-homogeneities generates an eddy current redistribution, changing the complex impedance of the bobbin coil [7].

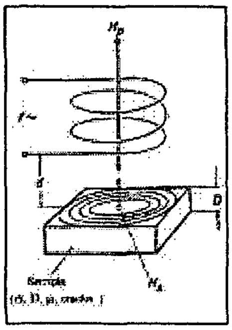


Figure 1. Electromagnetic interaction between bobbin coil and conductive sample (McMaster, 1959).

The electromagnetic phenomenon involved in ECT technique is almost static based on experimental apparatus and characteristic frequencies used. The equation that describes this phenomena, if we don't take the displacement current into consideration, is:

$$\nabla \times \left(\frac{1}{\mu}\nabla \times \vec{A}\right) = \vec{J}_s - j\omega\sigma\vec{A} \qquad (1)$$

where J_s is the source current density; A is the magnetic potential vector, μ is the magnetic permeability, σ is the material conductivity and ω is the excitation frequency. For a homogeneous medium Equation (1) can be rewritten as [7]:

$$\frac{1}{\mu}\nabla^2\vec{A} = j\omega\sigma\vec{A} - \vec{J}_s \qquad (2)$$

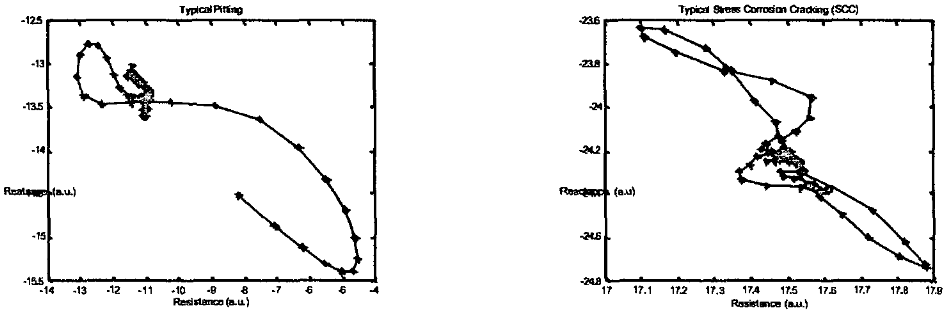
The bobbin impedance changes are translated into electric potential variations. The ECT equipment usually represents these complex impedance variations by two signals: resistance and inductive reactance. Traditionally the impedance plane trajectories are used to characterize the defect.

The eddy currents depend on three different material characteristics of the sample. These are geometry, electrical conductivity, and magnetic permeability. An important electromagnetic property of this technique is described as skin depth, which can be obtained by the solution of the simplified diffusion equation and represented by:

$$d = \frac{1}{2\pi\,(\mu_{rel}\,\sigma\,f\,10^{-7})^{1/2}} \qquad (3)$$

where d is the electromagnetic penetration depth and f is the ECT test frequency. This ECT property implies in the phase rotation phenomena where the impedance plane phase increases as the penetration depth is increased. Through this the defect depth estimation can be done after appropriate phase calibration. A common technique in ECT inspection of steam generator tubes is the multi-frequency analysis in which a single probe stimulates the sample with different excitation frequencies in order to detect different depth defects.

Typical steam generator tube defects are pitting (PIT) (Figure 2a), stress corrosion cracking (SCC) (Figure 2b), inter-granular attack (IGA), anti-vibration bar damage (AVB), mechanical fretting caused by tube-support plate interaction, and others. A set of different power plant operational conditions over many years cause these tube degradation. The usual operational conditions include high pressure, temperature, and flow with severe water chemistry reactions.



(a) (b)

Figure 2. Typical impedance plane plots (Lissajous like figures) of ECT data from steam generator tube defects: (a) PIT and (b) SCC.

Pitting (PIT) defect is due to galvanic differences in the tubing and assumes volumetric characteristics. Most of these defects are associated with acid chemical reactions. A typical PIT signal is shown in Figure 2a. Stress corrosion cracking (SCC) is caused by simultaneous actions of chemical corrosion and mechanical stress. The main physical characteristics of SCC are multiple major cracks with branching and two-dimensionality. A typical signal is shown in Figure 2b. Anti vibration bar (AVB) damage is due to mechanical interaction between the tubes and their supports. It consists of material removal characterizing a volumetric defect. Finally, the inter-granular attack (IGA) is a defect mainly caused by chemical corrosion. Mechanical stress has a less significant contribution and its physical characteristics may be volumetric or two-dimensional.

Methodology

The method used in this work consists of feature extraction techniques that are applied to a set of ECT steam generator tube defect signals. Two-dimensional SOM maps are obtained for each feature extraction in order to estimate the contribution of each feature for further defect classification. This algorithm will also permit multi-dimensional feature space visualization.

Self-Organizing Map (SOM)

The Kohonen map or Self-Organizing Map (SOM) is an unsupervised pattern recognition method whose objective is to represent every point in the source space by points in a target space keeping proximity and distance relations as much as possible. The algorithm is less complex than multidimensional scaling and generally can be applied to almost any source or target space [8]. The SOM is a nonlinear smoothed mapping of high-dimensional input data manifolds onto the elements of a low-dimensional array [9].

Suppose that the set of input variables $\{?_i\}$ is definable as a real vector $x = \{?_1, ?_{2,\dots,} ?_n\}^T \; ? \; R^n$, each element in the SOM array is also associated with a parametric real vector (model):

$$m_i = \{\mu_1, \mu_{2,\dots,} \mu_n\}^T \qquad (4)$$

This mapping implementation is done by an algorithm that attributes new values for each node (and its neighbors) in each iteration comparing each input sample x(t) with all the m_i, optimizing the distance measure. If the general distance between x and m_i is $d(x, m_i)$ the input vector SOM image is defined as the matrix element m_c that better fits the input vector (x), where c is the index:

$$c = \arg\min_i \{d(x,m_i)\}. \qquad (5)$$

It has been proved that if the x(t) and m_i are Euclidean vectors and a locally smoothed distance measure is used, then the process converges [10]. The nodes are arranged initially at positions that obey a topology function and are updated using the Kohonen rule [11] and some updating algorithms [12]. An example of this topological arrangement is shown in Figure 3 with natural speech spectrum in a two-dimensional map.

Algorithm implementation

During the training phase the neuron winner and its nearest neighbors (in an Euclidean sense) are adjusted to the input vector. The sequence is as follows:

1) Initialization of weights with small values.

2) Finding the winner through $c = \arg\min_i \{d(x,m_i)\}$.

3) Weight update using Kohonen rule: $\Delta w(t_{k+1}) = \alpha(t_k)\,(x(t_k) - o(t_k))$, where x is the input vector and o is the output vector.

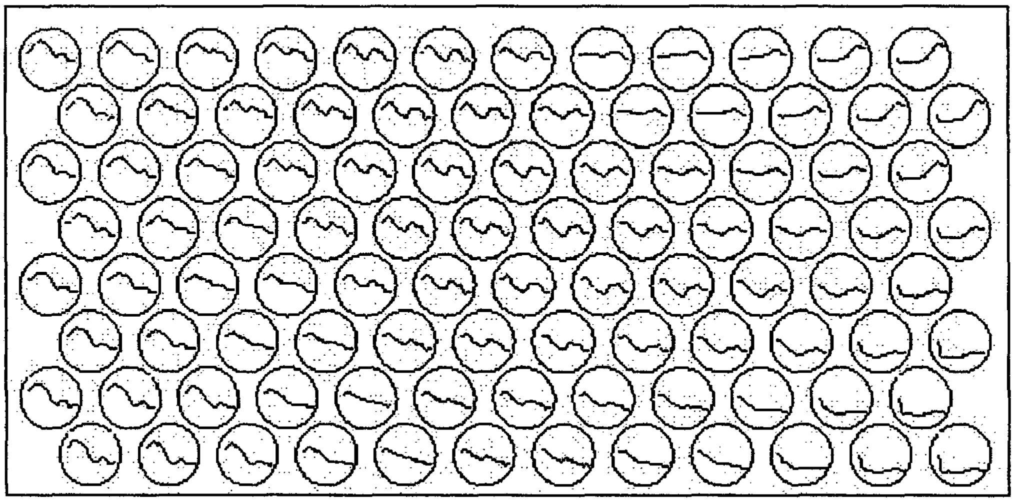


Figure 3. Hexagonal grid with natural speech spectrum [9].

One of the first signal processing and classification using SOM was image classification and character clustering by Kennedy and Morasso [13]. A vector quantization in images is done using SOM by Chiueh et al [14]. Debeljak et al [15] applied SOM to classify chromatographic systems.

Feature Extraction Methods

The following feature extraction methods are used in this work.

Signal Segmentation

Sets of 64-point vectors are obtained by sampling around the defect center point. This point is the median sample of the degradation area signal. The affected area generally contains 200 to 500 samples and the central sample is also used for phase estimation.

Phase

This is the main feature used during ECT inspection. The vectors are obtained from the ECT Lissajous figure formed in the impedance plane. It takes into account that the ASME calibration has been performed with 40-degree angle for a 100% through hole. The phase is defined as

$$\phi = \tan^{-1}\left(\frac{X_l}{R}\right) \tag{6}$$

Linear Predictive Coding (LPC)

LPC is a signal analysis technique mostly applied in speech recognition tasks [16]. The basic principle of this analysis is that the signal sample can be approximated as a linear combination of past samples. By minimizing the sum of squared prediction errors (over a finite interval) between the actual samples and the linearly predicted ones, a unique set of predictor coefficients can be determined. It has been used as feature extraction for signal recognition [17] and image compression [18]. An autoregressive one-step predictor model of a discrete signal x(k) has the form

$$x(k) = \sum_{i=1}^{n} a_i x(k-i) + w(k) \tag{7}$$

where $\{a_i\}$ are the model parameters. The model parameters and the model order, n, are estimated using least squares techniques. The sequence $\{w(k)\}$ represents uncorrelated noise.

Wavelet Zero-Crossing (WZC)

In the last decade the wavelet transformation has become a popular tool for signal processing. This technique was used by Upadhyaya et al [19] to establish a fuzzy inference system to classify ECT defect types based on the discontinuities detected on a set of signals.

By definition the function $\psi(x)$ is said to be a wavelet if:

$$\int_{0}^{+\infty}\frac{|\Psi(\omega)|^2}{\omega}d\omega = \int_{-\infty}^{0}\frac{|\Psi(\omega)|^2}{\omega}d\omega < +\infty . \tag{8}$$

Using $\psi(x)$ mother wavelet, the wavelet transformation of a function $f(x)$ at the scale s and position x is defined by the following convolution:

$$W_s^{\psi} f(x) = f * \psi_s(x) . \tag{7}$$

The wavelet transformation decomposes a signal into components using a dyadic scale where $s = 2^j$ is used. It is rarely computed for continuous scale. The wavelet transformation of a signal is proportional to the first derivative of the signal smoothed by θ(x), if the wavelet is the first derivative of a smoothing function, and it is proportional to the second derivative of the signal smoothed by θ(x) and if the wavelet is the second derivative of a smoothing function [19]:

$$W_s^{\psi} f(x) = f * \left(s\frac{d}{dx}\theta_s \right)(x) = s\frac{d}{dx}(f * \theta_s)(x) \tag{8}$$

If a discontinuity has a larger localized singularity than the background noise it is possible to separate the noise from the useful part of the signal. In this way, the number of discontinuities can characterize the signal regularity. Zero-crossings localization in wavelet representation can be considered as the inflection points of the original signal. A more stable representation by integrating the wavelet representation between two consecutive singular points:

$$c_n = \int_{z_{n-1}}^{z_n} W_{2^j}^{\psi} f(x)dx \tag{9}$$

where z_n and z_{n-1} are two consequent zero-crossing points, respectively.

Depending on the number of scales of wavelet transformation, a zero-crossing representation can be obtained through four basic steps [19]:

1) Symmetry operation on both sides of the signal for avoiding border distortion,
2) Multi-scale wavelet transformation,
3) Determination of zero-crossing points,
4) Calculation of integrals between two adjacent zero-crossing points.

Results

The data were taken from EPRI's Performance Demonstration Database (PDD) [20] and include typical signals from actual ECT measurements in operating steam generators. Typical signals taken from a specific frequency channel were grouped by defect type and used to generate the training data set. Typically, the excitation frequency of the used channel was 400 kHz. There were a few slightly different excitation frequency channels used in this training set. This methodology looks for characteristic features that do not depend on the excitation frequency channel. This training data set consists of 246 signal segments of 64 points each. The total number of signal segments, which includes all the frequency channels, is 1920. Each defect type was represented by two different nuclear power plant signals. These training signal segments were labeled in order to be visualized in the SOM hexagonal plane as "ws" and "gs" (for SCC), "wa" and "ga" (for AVB), "wp" and "gp" (for PIT) and "wi' (for IGA). This data set was properly calibrated and normalized.

The SOM matrix size used is (10 x 12) that is approximately equal to the total number of signal segments to be clustered. The map lattice geometry was hexagonal. The initialization was random and a batch-training algorithm was used. The SOM Toolbox [21] for Matlab was used for this implementation.

Sets of intrinsic features from these signal segments were taken. The features were divided in: Real Segment (Figure 4a), Imaginary Segment (Figure 4b), WZC of Real Segment (Figure 4c), WZC of Imaginary Segment (Figure 4d), Signal Phase (Figure 4e), and LPC of Imaginary Segment (Figure 4f). The LPC of the Imaginary Segment used 10 LPC coefficients taking into account the predictor error level and computational effectiveness. The colored boundaries in the figures were manually drawn to emphasize the different clusters extracted. These maps can be used to select the best feature extraction methods for classification.

It is clear from these initial results that the SOM clustering applied directly to the signal segments or their extracted features produces clearly separated clusters, especially for the Real Segment (Figure 3a) and for the WZC maps (Figures 4c and 4d).

A systematic set of experiments using a mixture of these feature extraction methods in order to obtain two or three maps that can be used to create fuzzy inference rules for ECT defect type classification will be performed. This method uses the unique information that comes directly from the signal.

The classification will be made using membership functions created over the SOM clusters obtained on the different maps. As an example, the rule would be: "If Real_Segment is wa AND WZC_Imaginary is wa THEN the signal is AVB type".

Concluding Remarks

The SOM algorithm proves to be applicable to ECT defect classification. The established cluster show that the differences among these feature vectors are sufficient to separate the data into different classes. A supervised classification and testing is necessary to validate an optimal methodology.

The continuing work includes a systematic study with additional feature extraction methods and more eddy current test data.

Acknowledgments

This research is sponsored by the Nuclear Engineering Center, IPEN, Brazil. The research at The University of Tennessee is supported in part by the U.S. National Science Foundation under Grant No. 9901201.

References

[1]. L. Udpa and S.S. Udpa, "Eddy current defect characterization using neural networks." Materials Evaluation, v.48, n.3, pp. 342-347, 1990.

[2]. B.R. Upadhyaya and W. Yan, "Hybrid digital signal processing and neural networks for automated diagnostics using NDE method." Annual Report of Nuclear Engineering Department of the University of Tennessee, Prepared for the U.S. Nuclear Regulatory Commission Office Of Nuclear Regulatory Research, Nov. 1993.

[3]. R.J. Kurtz, P.J. Heasler, and C.M. Anderson, "Performance demonstration tests for eddy current inspection of steam generator tubing." Pacific Northwest National Laboratory, prepared for U.S. Nuclear Regulatory Commission, NUREG/CR-6227, PNLL-9433, USA, May 1996.

[4]. S.J. Kang, J.C. Moon, and D.H. Choi, "A distributed and intelligent system approach for the automatic inspection of steam-generator tubes in nuclear power plants." IEEE Transactions on Nuclear Science, v.45, n.3, pp.1713-1722, 1998.

[5]. B. R. Upadhyaya, W. Yan, M. M. Behravesh, and G. Henry, "Development of a Diagnostic Expert System for Eddy Current Data Analysis Using Applied Artificial Intelligence Methods," Nuclear Engineering and Design, Vol. 193, pp.1-11, 1999.

[6]. W. Yan and B.R. Upadhyaya, "Development of an automated diagnostics system for eddy current analysis using applied artificial intelligence techniques." Annual Report, The University of Tennessee, Prepared for the Electric Power Research Institute, February 1996.

[7]. S.N. Rajesh, L. Udpa, and S.S. Udpa, "Numerical model based approach for estimating probability of detection in NDE applications." IEEE Transactions on Magnetics, v.29, n.2, pp.1857-1860, March 1993.

[8]. R. O. Duda, P.E. Hart, and D.G. Stork, "Pattern Classification," John Wiley, New York, 2001.

[9]. T. Kohonen, "Self-Organizing Maps," p.59, Springer-Verlag, 2001.

[10]. Y. Cheng, "Convergence and ordering of Kohonen's batch map", Neural Computation, Vol. 9 (8), pp. 1667-1676, 1997.

[11]. Z.-P. Lo and B. Bavarian, "On the rate of convergence in topology preserving neural networks," Biological Cybernetics, Vol. 65, pp.55-63, 1991.

[12]. Z.-P. Lo, Y. Yu and B. Bavarian, "Analysis of the Convergence Properties of Topology Preserving Neural Networks," IEEE Transactions on Neural Network, Vol. 4, pp.207-220, 1993.

[13]. J. Kennedy and P. Morasso, "Application of self-organizing networks to signal processing." Lecture Notes on Computer Science, n.412, pp.225-232, 1990.

[14]. T.D.Chiueh, T.T. Tang, and L.G.Chen, "Vector quantization using tree-structured self-organizing feature maps." IEEE Journals on Selected Areas in Communications, v.12, n.9, pp.1594-1599, December 1994.

[15]. Z. Debeljak, M. Strapac, and M. Medic-Saric, "Application of self-organization maps for the classification of chromatographic systems and prediction of values of chromatographic quantities." Journal of Chromatography A, v.925, pp.31-40, 2001.

[16]. L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signals," Prentice-Hall, 1978.

[17]. M. R. Azimi-Sadjadi, D. Yao, Q. Huang and G. J. Dobeck, "Underwater Target Classification Using Wavelet Packets and Neural Networks," IEEE Transactions on Neural Networks, Vol. 11, No. 3, pp.784-794, May 2000.

[18]. M. A. Wade, "Numerical Study of Linear Predictive Coding for the Compression of Gray Scale Images," MS thesis, The University of Tennessee, Knoxville, Dec 1993.

[19]. B.R Upadhyaya, A.S. Erbay, and G. Házi, "Eddy current test data analysis for steam generator tubing diagnosis using artificial intelligence methods." Annual Report, The University of Tennessee, Prepared for the Electric Research Institute (EPRI), February 1999.

[20]. Electric Power Research Institute (EPRI), "Steam generator eddy current data analysis performance demonstration: review material," Palo Alto, California, Oct. 1995.

[21]. SOM Toolbox team, Laboratory of Computer and Information Science, P.O.Box 5400, Finland.

MARCON
2002

“Moving Forward Toward Excellence”

May 5-8, 2002
Radisson Summit Hill

MARCON 2002 returns to Knoxville and will be held at the Radisson Hotel, located just a short distance east of the University of Tennessee campus. We invite you to bring your family and friends and arrive early, stay late, or both, to really enjoy the warmth and hospitality of this beautiful area. Knoxville has much to offer visitors, both in town and in the nearby area. From the Old City area to the new Women's Basketball Hall of Fame to the new Riverfront area to the University area, conferees will find a myriad of shops, restaurants, and other points of interest. The surrounding area includes The Great Smoky Mountains National Park, Gatlinburg, Pigeon Forge, Oak Ridge, and many other attractions to appeal to your interests. You will find plenty to do for yourselves or for family members in addition to attending this outstanding conference.

Three Three days filled with the latest thinking from around the world
Set in the heart of East Tennessee

MARCON 2002 Program **Index of Authors**

If you need to load Acrobat Reader, Click Here!

Maintenance and Reliability Conference.

9288