Ultrasonic Sensor Data Processing using Support Vector Machines

D.Isa, R.Rajkumar
School of Electrical and Electronic Engineering
University Of Nottingham, Malaysia Campus,Jalan Broga, 43500, Semenyih, Selangor, Malaysia
Dino.Isa@nottingham.edu.my.

Abstract— **Ultrasonic sensors are ideal for non-destructive testing due to its many advantages over conventional sensors. Oil and gas pipelines are an area which uses ultrasonic sensors for monitoring and detecting the presence corrosion and defects. The proposed techniques ultimately aims at providing a continuous monitoring system using an array of ultrasonic sensors strategically positioned on the surface of the pipeline to predict the occurrence of defects rather than just monitoring. The sensors used are piezoelectric ultrasonic sensors. The raw sensor signal will be first processed using the Discrete Wavelet Transform (DWT) as a feature extractor and then classified using the powerful learning machine called the Support Vector Machine (SVM). Preliminary tests show that the sensors can detect the presence of wall thinning in a steel pipe by classifying the attenuation and frequency changes of the propagating lamb waves. The SVM algorithm was able to classify the signals as abnormal in the presence of wall thinning.**

Key-Word- Ultrasonic Sensor, Pipeline, Support Vector Machines, Discrete wavelet transform

I. INTRODUCTION

Ultrasonic waves have been used in detecting defects in pipes, tubes and metal plates in many applications [1][2]. In the area of oil and gas pipelines, ultrasonic sensors are incorporated in many commercial products for monitoring corrosion and defects [3]. Ultrasonic waves propagate through the pipeline as Lamb waves thus picking up critical information on the condition of the pipe. Ultrasonic sensors enable detection without any contact with the object regardless of its material, nature, color and degree of transparency. The advantages of ultrasound detection include:

- No physical contact with the object to be detected, therefore, no wear and detection possible of fragile or freshly painted objects, etc.
- Detection of any material, irrespective of color, at the same distance, without adjustment or correction factor.
- Very good resistance to industrial environments (robust products entirely encapsulated in resin). Tough environments such as fumes, dust, noisy.
- Solid-state units: no moving parts in the sensor, therefore, service life independent of the number of operating cycles.

Currently, an established form of pipeline inspection uses smart pigs in a process called pigging [4]. These smart pigs travel within the pipeline recording critical information like corrosion levels, cracks and structural defects using its numerous sensors. Pigs can give pinpoint information on the location of defects using techniques like magnetic flux leakage and ultrasonic detection [5]. However, using smart pigs in pipeline inspection has a few disadvantages. The cost of implementing a pigging system can be expensive. More importantly, pigs measure the pipeline condition only at the instance it is deployed and does not provide continuous measurements over time. The proposed techniques aim at providing a continuous monitoring system using an array of different sensors strategically positioned on the external surface of the pipeline. The raw sensor signal will be first processed using the Discrete Wavelet Transform (DWT) and then classified using the powerful learning algorithm called the Support Vector Machines (SVM).

The DWT is used here as a feature extraction tool in order to single out any unique features in the sensor data. A useful property of the DWT is that it compresses signals and by doing so, it has the tendency to eliminate high frequency noise. The DWT is used here to eliminate noise in the sensor signals and also to compress large amounts of real-time sensor data for faster processing. The compressed data or the DWT coefficients are then used as inputs to the SVM classifier, which will fuse the different sensor data together and then perform the classification task. The SVM has been widely used lately for numerous applications due to its excellent generalization ability with small training samples. The SVM will be trained with normal and simulated defect conditions using an experimental pipeline rig in the laboratory.

II BACKGROUND

A. *Support Vector Machines*

SVM functions by creating a hyperplane that separates a set of data containing two classes. According to the SRM principle, there will just be one optimal hyperplane, which has the maximum distance (called maximum margin) to the closest data points of each class as shown in Fig. 1[6]. These points, closest to the optimal hyperplane, are called Support Vectors (SV). The hyperplane is defined by the equation

$\mathbf{w}.\mathbf{x} + b=0$, and therefore the maximal margin can be found by minimizing (1).

$$\tfrac{1}{2}\, ||\mathbf{w}||^2 \tag{1}$$

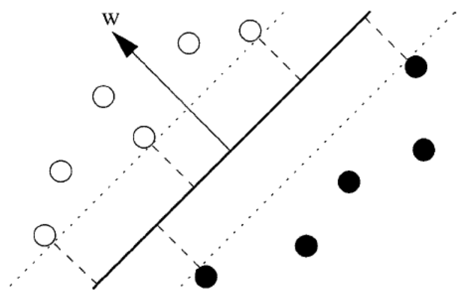

Fig. 1: Optimal Hyperplane and maximum margin for a two class data.

The Optimal Separating Hyperplane can thus be found by minimizing Eq. (1) under the constraint Eq.(2) that the training data is correctly separated [7].

$$y_i.(\mathbf{x}_i.\mathbf{w} + b) \geq 1 \text{ , } \forall \text{ i} \tag{2}$$

The concept of the Optimal Separating Hyperplane can be generalized for the non-separable case by introducing a cost for violating the separation constraints Eq.(2). This can be done by introducing positive slack variables ξ_i in constraints Eq.(2), which then become:

$$yi.(\mathbf{x}_i.\mathbf{w} + b) \geq 1 - \xi_i \text{ , } \forall \text{ i} \tag{3}$$

If an error occurs, the corresponding ξi must exceed unity, so $\Sigma_i\ \xi_i$ is an upper bound for the number of classification errors. Hence a logical way to assign an extra cost for errors is to change the objective function Eq.(1) to be minimized into:

$$\min \{\ \tfrac{1}{2}\, ||\mathbf{w}||^2 + C.\ (\Sigma_i\ \xi_i\ )\ \} \tag{4}$$

C is a tuning parameter which allows the user to control the trade off between maximizing the margin (first term in the objective) and classifying the training set without error. Minimizing Eq.(4) under constraint in Eq.(3) gives the *Generalized Optimal Separating Hyperplane*. This is a Quadratic Programming (QP) problem which can be solved here using the method of Lagrange multipliers [8].

After performing the required calculations [7], [9], the QP problem can be solved by finding the LaGrange multipliers, α_i, that maximizes the objective function in Eq.(5),

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\mathrm{x}_i^{\mathrm{T}} \mathrm{x}_j\right) \tag{5}$$

Since the input vectors enter the dual only in form of dot products the algorithm can be generalized to non-linear classification by mapping the input data into a high dimensional feature space via an a priori chosen non-linear mapping function Φ. Constructing a separating hyperplane in this feature space leads to a non-linear decision boundary in the input space. Expensive calculation of dot products in a high-dimensional space can be avoided by introducing a kernel function, K in Eq. (6).

$$K\left(x_i, x_j\right) = \Phi(x_i) \cdot \Phi(x_j) \tag{6}$$

By introducing the kernel function, it is not necessary to explicitly know Φ(.). So that the optimization problem in Eq.(5) can be translated directly to the more general kernel version in Eq. (7),

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1,\, j=1}^{n} \alpha_i \alpha_j y_i y_j K\left(x_i, x_j\right) \tag{7}$$

subject to .

$$C \geq \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i y_i = 0$$

The equation for the indicator function, used to classify new data (from sensors) is given in Eq.(11) where the new data **z** is classified as class 1 if i>0, and as class 2 if i <0 [10].

$$i_F(x) = sign\left[\sum_{i=1}^{l} y_i \alpha_i K(\mathrm{x}, \mathrm{x}_i) + b \right] \tag{8}$$

Note that the summation is not actually performed over all training data but rather over the support vectors, because only for them do the Lagrange multipliers differ from zero. Fig. 2 illustrates the SVM data flow, from input data point to the final decision value [11].

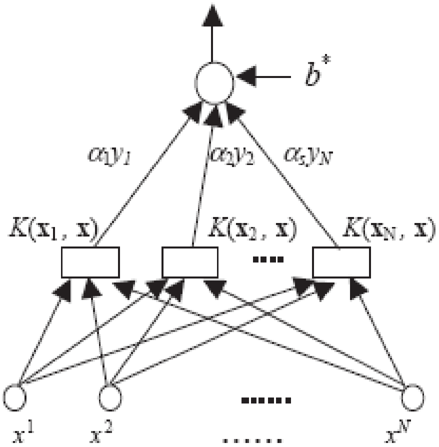

Fig 2: Diagram of SVM data flow

As such, using the support vector machine we will have good generalization and this will enable an efficient and accurate classification of the sensor signals. It is this excellent generalization that we look for when analyzing sensor signals due to the small samples of actual defect data obtainable from field studies. In this work, we simulate the abnormal condition and therefore introduce an artificial condition not found in real lie applications.

B. *Discrete Wavelet Transform*

A discrete wavelet transform (DWT) is basically a wavelet transform for which the wavelets are sampled in discrete time. The DWT of a signal *x* is calculated by passing it through a series of filters. First the samples are passed through a low pass filter with impulse response *g*, resulting in a convolution of the two (9). The signal is also decomposed simultaneously using a high-pass filter *h* (10).

$$y_{\mathrm{low}}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \tag{9}$$

$$y_{\mathrm{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \tag{10}$$

The output of the equations 9 and 10 gives the detail coefficients (from the high-pass filter) and approximation coefficients (from the low-pass). It is important that the two filters are related to each other for efficient computation and they are known as a quadrature mirror filter [12].

However, since half the frequencies of the signal have now been removed, half the samples can be discarded according to Nyquist's rule. The filter outputs are then down sampled by 2 as illustrated in Fig. 3. This decomposition has halved the time resolution since only half of each filter output characterizes the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled. The coefficients are used as inputs to the SVM [13].

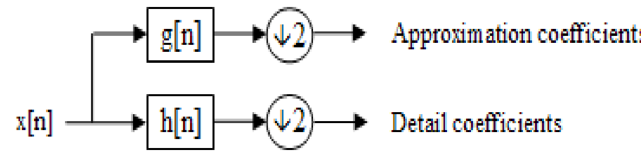

Fig. 3: DWT filter decomposition

C. *Corrosion Measurement*

A pipe failure and leakage of crude oil in Winchester, Kentucky on January 2000, was one of the biggest accidents that occurred and it incurred the owner Marathon Ashland Pipe Line LLC a clean up cost of $7.1 million. The crack was due to a small dent in the pipe that might have been caused by stone particles flowing along the path, in addition to the fluctuating pressure of the pipe wall [14]. An example of such a failure is shown in Fig. 4.

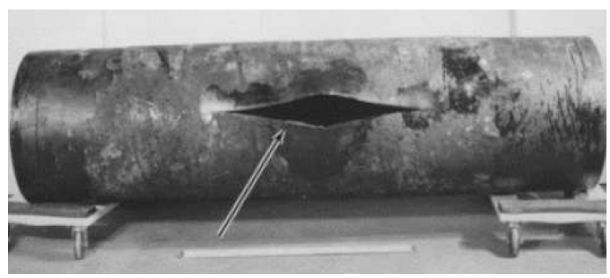

Fig. 4: The rapture pipe due to fatigue cracking [14].

Wall thinning, a common occurrence in the oil piping industry, is characterized by metal loss caused by surface erosion due to high temperature, high pressure and high flowing velocity of the flowing commodity [15]. The pipes are also subjected to combined loading by internal pressure, bending moment, and longitudinal forces. The internal wall thinning of a pipe cannot be observed from the outside of the pipe, hence a method of condition monitoring using ultrasonic waves as a non-destructive test of the metal loss can help to determine when the pipe may be at risk for leaks or failure. Ultrasonic sensor enables detection without any contact with the object regardless of its material, nature, color and degree of transparency.

The detection technology used here lies within the concepts of nonlinear acoustics. This basically states that when sound waves travels through a material, frequency and attenuation changes occur to the sound waves. The changes in the frequency and amplitude must be detected and analyzed to give precise information on the state of the material. Ultrasonic transmitters can be used to send ultrasonic waves and ultrasonic receivers can be used to detect the propagating waves. These sensors are very accurate as they can produce and detects high frequency sound wave based on Piezoelectricity [16]. Piezoelectric transducers have solid-state pressure sensitive elements that will expand and contract in step with input signals.

Demma [17] examined the effect of defect size with frequency on the reflection from notches and was able to show the relation between the value of reflection coefficient and the defect sizing. The cylindrical ultrasonic waves propagate along the pipe and are partially reflected when met with defects thus providing a fast screening technique to determine the presence of defects. Similar results and observation are recorded by Lin [18] by using guided waves and electromagnetic acoustic transducers (EMATs) to measure the wall thickness precisely. Wave propagation is performed for a specimen with thickness of 10mm, where different artificial defects are introduced to model local wall thinning. As shown in Fig.5, when transmitted waves impinge the wall thinning, they are reflected and the intensity of the reflected waves varies.

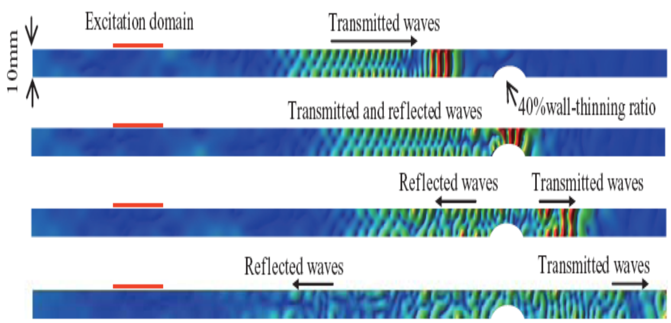

Fig. 5: The energy carried by the transmitted waves passes through the wall thinning and some reflected back as echoes. [18]

It is therefore a well known phenomenon, both theoretically and experimentally that defects in pipes can be detected by ultrasonic transducers [15].

III METHODOLOGY

This section details the experimental setup that will be used to simulate pipeline conditions and also defect conditions. The aim is to create a scaled downed version of an actual section of pipeline in the laboratory using commonly available materials. Fig. 6 shows the experimental setup. A motor pump is used to pump hydraulic oil in the reservoir through the pipeline section. A flow rate of around 5 m^3/h was achieved through a 1 m section of pipe (outer diameter of 48.30 mm and inner diameter of 42 mm). Two experiments will be carried out. The first experiment is to compare a defective pipe and a normal pipe.

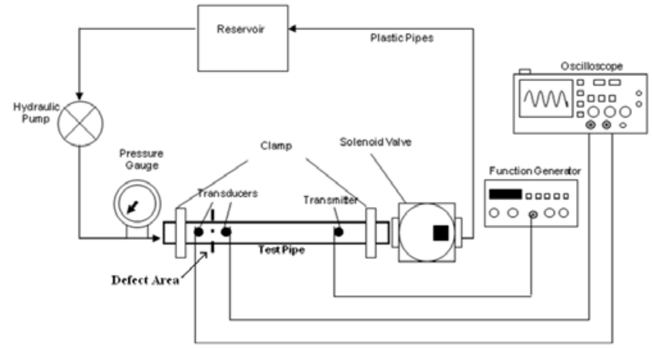

Fig. 6: Pipeline Experimental Setup.

A lathe is used to clear an area of 1mm wide and 1mm deep all around the circumference at the inner surface of the pipe. This is to simulate a crack or corrosion at the inner surface of the pipe. An ultrasonic transmitter is used to transmit a signal across the flowing pipe and through the defect area to see changes in the ultrasound signal. Ultrasonic receivers, placed at the other end of the pipe, will be able to pick-up the waveform that is vibrating in the pipe and can be used to monitor the condition of the pipe. The changes in the ultrasound signal will be used to determine the presence of any defects. The experiment was repeated on a pipe without any defects.

Ultrasonic sensors used are Murata analogue ultrasonic sensors which is an open structure type of sensor that has a range of up to 6 m and they will be attached to the outer surface of the pipe using epoxy. MA40B8S have the nominal frequency of 40 kHz with the maximum input voltage of 40V peak to peak. The stationary sensors can avoid any disturbance from the environment and will be able to transmit ultrasound along the length of the pipe by ringing the surface of galvanized steel pipe.

The second experiment is to simulate progressive defects or corrosion as seen in real pipelines. Therefore, a simulation was carried out to corrode the interior of the test pipe. In order to perform this, the test pipe is removed from the setup. Stones and rocks of different sizes are passed through the pipe for 5 hours a day to speed up the corrosion process. This is done by recycling stones and rocks by using a conveyor belt and clamping the pipe vertically. This process is repeated daily in order to cause random metal loss in the interior of the pipe.

IV RESULTS

The results of from the experimental rig will ultimately be used to ascertain whether SVM can detect the presence of cracks and whether DWT helps in the decision making. DWT is performed on raw time domain samples and the coefficients of the resulting DWT are inputted into the SVM for classification. Various wavelets can be tested including the Haar and Daubechies wavelets. A popular SVM algorithm called LIBSVM [19] is used to perform the SVM calculation. LIBSVM includes for kernel functions: linear, polynomial, radial basis function (RBF), and sigmoid. To train an SVM, the user must select the proper *C* value as well as any required kernel parameters.

Time domain samples before and after the defect area are first broken down into frames where the number of samples within the frame is a variable. Each frame will represent one instance or sample needed for the SVM and the frame size is the number of attributes or dimensions. Table 1 show the results of the first experiment where the SVM accuracy is shown as a percentage. The signals are decomposed into two frame sizes, 25 and 50 and inputted into the LIBSVM algorithm. 10,000 data points from the defective pipe and 10,000 data points from the normal pipe are used to obtain the results. This is therefore is a binary classification problem where the two classes are defective and non-defective.

Table 1: Classification accuracy (%) for pipeline data using LIBSVM for various kernel functions.

Wavelet	**Kernel Func.**	**Frame Size**	
		25	**50**
-	Poly	73.68	67.74
	RBF	75.44	70.97
	Sig	73.68	67.74
DB2	Poly	83.87	61.11
	RBF	80.64	72.22
	Sig	80.64	61.11
Haar	Poly	89.65	75.00
	RBF	89.65	75.00
	Sig	86.21	75.00

As can be seen from Table 1, the smaller frame size provides better classification accuracy than the bigger frame size. The radial basis function (RBF) kernel shows the highest classification rates among the kernel functions tested. The Haar wavelet also shows better classification accuracy as compared to the DB2 wavelet. Fig. 7 shows the results of the second experiment where the defects are progressively increased over time. The results shown are taken at weekly intervals with results for 5 weeks shown.

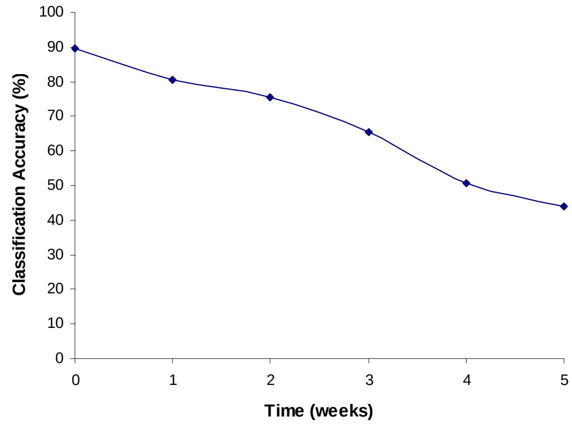

Figure 7: SVM lassification accuracy as defect size increased over time.

As can be seen from the data from Fig 7, as the defects are increased over time, the SVM classification accuracy decreases. This trend can be used to predict the occurrence of defects.

CONCLUSION

Monitoring hundreds of kilometers of pipelines is a difficult task due to the high number of unpredictable variables involved. Rapidly changing weather conditions, pressure changes and erosion due to gas or oil flow and ground movement are a few variables that can have direct impact on the pipelines. There variables can cause defects like corrosion, dents and cracks which will lead to loss of the valuable commodity and not to mention the series affects on the surrounding environment.

The use of an array of sensors with help of support vector machine processing intends to solve these problems in two ways. Firstly the array of sensors provides a continuous monitoring platform along the entire distance of the pipeline. Secondly the use of artificial intelligence tools like support vector machines, makes it possible to monitor and ultimately predict the occurrence of defects. Support vector machines are ideal for applications like these where there are high number of dimensions of data (sensors) and also the small number of samples for defect scenarios. SVM have been used widely in many such applications and has provided excellent generalization performance.

An experimental miniature pipeline rig provided the setting to examine the initial performance of the SVM on pipeline-like data. For now, no correlation study is made between the simulated and real situation. Ultrasonic sensors were used and the corrosion defect was simulated using human manipulation. The results of the first experiment showed good performance by the SVM using an RBF kernel function. Results with other kernels indicate that these other kernels do not accurately represent the inner product of feature vectors of that data set. The use of DWT further improved the performance of the SVM accuracy to 89.65%. This is due to the DWT compressing the data and filtering away unwanted noise from the high frequency acoustic signals. The second experiment show that as the defects are increased, the SVM classification accuracy decreases. The reason for this is that the signal attenuation and phase are changing as the defects are increased. So the ultrasonic signals are looking more and more dissimilar to that of the pipe at the beginning causing the SVM algorithm to classify more points to the defective class.

The conclusion is reached that a combination DWT and SVM algorithm can predict, to a high accuracy the presence of defects and also progressive defects in small pipes. The results of this paper will be used in future research where multiple sensors resulting in multiple time series signals need to be analyzed on bigger diameter pipes. Multiple defect scenarios will also be studied resulting in multiclass classification problems. The ultimate aim of the research will be to predict defects before they occurs thereby conserving the precious commodity and environment.

REFERENCES

[1] Tua P S, Quek S T and Wang Q, Detection of cracks in cylindrical pipes and plates using piezo-actuated Lamb waves, Smart Mater. Struct., Vol.14, 2005, 1325-1342

[2] Demma, A., Cawley, P., Low, M., Roosenbrand, A.G. and Pavlakovic, B., 2004, The reflection of guided waves from notches in pipes, *NDT&E International*, 37, 167-180.

[3] World Centre For Materials Joining Technolog, 'Long Range Ultrasonic Technologies' Part of TWI's Non-Destructive Testing Technology Group. URL: www.twi.co.uk

[4] Lebsack, S., 2002, Non-invasive inspection method for unpiggable pipeline sections, *Pipeline & Gas Journal*, 59.

[5] Bickerstaff, R., Vaughn, M., Stoker, G., Hassard, M. and Garrett M. 2002. Review of sensor technologies for in-line inspection of natural gas pipelines, Sandia National Laboratories.

[6] Müller K., Mika S., Rätsch G., Tsuda K., and Schölkopf B.," An Introduction To Kernel-Based Learning Algorithms, *IEEE Transactions On Neural Networks*, Vol. 12, No. 2, March 2001, 181

[7] Burges, Christopher J.C. "*A Tutorial on Support Vector Machines for Pattern Recognition*".Bell Laboratories, Lucent Technologies. Data Mining and Knowledge Discovery, 1998.

[8] Haykin, Simon. *"Neural Networks. A Comprehensive Foundation",* 2nd Edition, Prentice Hall, 1999

[9] V. Vapnik. The Nature of Statistical Learning Theory, 2nd edition, Springer, 1999

[10] Kecman, Vojislav, *"Support Vector Machines Basics"*, School Of Engineering Report 616, The University of Auckland, 2004. URL: http://genome.tugraz.at/MedicalInformatics2/Intro_to_SVM_Report_616_V_Kecman.pdf

[11] Ming Ge, R. Du, Guicai Zhang, Yangsheng Xu, 'Fault diagnosis using support vector machine with an application in sheet metal stamping operations, Mechanical Systems and Signal Processing 18 (2004) 143–159.

[12] Mallat, S. *A Wavelet Tour of Signal Processing*, 1999, Elsevier.

[13] Lee, K. and Estivill-Castro, V., Classification of ultrasonic shaft inspection data using discrete wavelet transform. *In: Proceedings of the Third IASTED International Conferences on Artificial Intelligence and application*, ACTA Press. pp. 673-678

[14] National Transportation Safety Board Washington, D.C. 20594, Pipeline Accident Brief. URL: www.ntsb.gov/publictn/2001/PAB0102.pdf

[15] Koji Takahashi, Kotoji Andoa, Masakazu Hisatsune, Kunio Hasegawa, "Failure behavior of carbon steel pipe with local wall thinning near orifice", *Nuclear Engineering and Design*, No. 237, April 2006, pg 335–341

[16] P S Tua, S T Quek and Q Wang, Detection of cracks in cylindrical pipes and plates using piezo-actuated Lamb waves, Smart Mater. Struct., Vol.14, 2005 1325-1342

[17] Demma, A., Cawley, P., Low, M., Roosenbrand, A.G. and Pavlakovic, B., 2004, The reflection of guided waves from notches in pipes, *NDT&E International*, 37, 167-180.

[18] S. Lin, et.al., 2006, "Development of Measuring Method of Pipe Wall Thinning Using Advanced Ultrasonic Techniques", CRIEPI Report Q05006

[19] Chang C , Lin C., 'LIBSVM : A Library for Support Vector Machines', 2001. Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

Dino Isa is an Associate Professor at the University of Nottingham, Malaysia Campus attached to the Faculty of Engineering and Computer Science. His current interest is in applying the support vector machine in various domains in order to further understand its mechanisms. Dr Isa obtained his BSEE (Hons) in Electrical Engineering from the University of Tennessee, Knoxville in 1986 and his PhD from the University of Nottingham, University Park, Nottingham, England in 1991. He was employed in Motorola Seremban as Engineering Section Head after he completed his PhD, and moved to Crystal Clear Technology in 1996 as Chief Technology Officer prior to joining the University of Nottingham in 2001.

R RajKumar is currently a PhD student in the School of Electrical and Electronic Engineering, Faculty of Engineering and Computer Science, University of Nottingham, Malaysia Campus. He obtained his Masters from the same University in 2005 and his main interest is in using the Support Vector Machine to predict failures in oil and gas pipelines.