NDT.net - The e-Journal of Nondestructive Testing (May 2008)
For more papers of this publication click: www.ndt.net/search/docs.php3?MainSource=25

Robust defect detection algorithm based classification on features extracted from MFL signals

ALI SADR [1] , SAEEDREZA EHTERAM [1, 2]

[1] Department of Electrical Engineering
Iran University of Science & Technology (IUST) , Narmak,Tehran,Iran
[2] Researcher of Young Researchers Club of Iran.
E mail: Sadr@iust.ac.ir , sehteram@ee.iust.ac.ir

Abstract

ANALYTICAL models at recent years are developed to determine surface-breaking defects along the applied field when using the magnetic flux leakage (MFL) non-destructive technique. The theoretical model fits the experimental MFL results from simulated defects. For MFL sensors, the normal magnetic leakage field is subsequently used for evaluation of defects. Permeability variations were neglected by employing a flux density close to sample saturation. Three different defect geometries were experimentally investigated and the validity of the analytical model was verified. Different Feature extractor functions are applied in this paper to yield fast decision and more accurate. Indeed more accuracy is because of decision on different features that yields by employing two kinds of feature extractors, PCA and DCT. Both feature extraction and Multilayer perceptron (MLP) methods for identifying erosion defects are described and employed in this paper. Great accuracy rate in compare between results of related approaches suggests that this Method can be used as an algorithm of MFL data interpretation technique.

Keywords: Magnetic flux leakage, PCA, DCT, Multilayer perceptron, erosion defects

1. Introduction

Pipeline transportation is one of the fundamental modes in petroleum and natural pipeline long distance transportation. Ferromagnetism pipeline forms various defect because of corrosion, attrition and mechanical damage. It is necessary for pipeline's security evaluation and maintenance to detect the pipeline regularly using pipeline detector and obtain the precise information of the defect [1,2]. Among various pipeline inspection technology MFL inspection is the most widespread and perfect one. It has well Effect in ordinary defect detection, such as loss of metal. Applying MFL inspection technology, the defect recognition is mainly completed by man at present. And in this way, the defect only can be recognized. With the improvement of MFL device's precision and the extension of inspection distance, the quantity of the data grows sharply. It need long time for man to analyze a long pipeline data [3]. So finding the intellective technology to recognize pipeline defect quantitatively is urgent. During such an inspection a large amount of data is acquired and the detection of flaw-shaped features in the signal has to be performed manually. This is not only a time consuming and tiring task, moreover, the result depends on human elements of uncertainty. so for this reason we applied a mathematical relation between the magnetic field applied on the surface and the defect properties. in this way an approach is to find many exactly samples from a defect which is sorted in the surface by its various radial and depth. In this paper, the pipeline MFL image is recognized in an artificial algorithm that is trained BP neural network [4]. also In this work, an approach for the automatic detection of a defect is presented, where the NDE data are preprocessed using an analytical model of the magnetic flux and the extracted information is passed on to a panel of neural networks.

2. Database of defects from MFL testing

The database of the experimental MFL signals that is employed in this project is from The department of physics from Queens in Canada which is the Applied Magnetics Group (AMG). this database conclude signals of MFL that measured from Outside and Inside of a pipeline. Details of this database will lead to both un annealed and annealed data plots of increasing dent depths from 3mm to 7mm, resulting in a total of 10 plots for each one**.** For an instance Figure 1 illustrates a measurement from an annealed and not annealed MFL measurement from inside and out side of a pipeline.

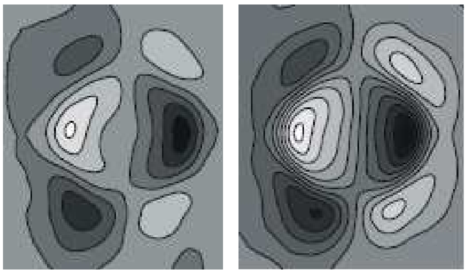

Figure1. after anneal(left) and before anneal (right) from outside of a 3 mm pipeline.

3. Preprocessing of the NDE data

A major problem in the design of a neural network for any application is to find the significant input parameters as well as the topology of the network. Although a neural network is able to learn and to generalize nonlinear relations between input and output, it is not possible to make a decision from input information which is not relevant [5]. In this approach we would like to feed our neural networks by mathematical signal producer that could specify a real defect signal with a defined shape than we test our system by some real MFL measuring that are mentioned above. The expected signal shape can be calculated analytically by means of a dipole approximation model. Such a physical model uses a number of unknown material and system parameters. Due to many function variables it becomes difficult to use the analytical model as a fit function for measured data. Thus, the variables have to be reduced by summarizing them into one single parameter. The task of the neural network in the defect detection algorithm presented in this work is to compare the defect free and defected signals achieved by mathematical and physical simulation. Thereby many sort of defects could be generated with this procedure but we focus on some of general sort of defect that are described by types of q and p parameters that are described follows.

3-1. MFL signal processing

The MFL signals are processed in a series of Algorithms which include defects of MFL signal interference eliminating and compensating. The defect MFL signals or maps are extracted from background magnetization signals.

The calculated signals covers almost all sorts of defects caused by the benign pipeline artifacts (such as valves, welds, tees, flanges, etc), corrosion, metal loss, mechanical damage and dents in the pipeline. Because most of defects could be sorted in two features and those are various radial and various depths. The MFL signal is normally contaminated by interference which is caused by geomagnetism, residual magnetic flux, electromagnetic interference coupling, probe sensor vibration, oil-gas pipeline pressure, liquid or gas noise and additive white Gaussian noise and etc [10]. The MFL signal patterns due to benign pipeline artifacts have obviously different characteristics and can be separated directly. MFL signals are actually also altered by the stress inside the oil-gas pipeline, the variations of magnetic conductivity due to inhomogeneous pipeline material and inspection tool motions.

The algorithm to achieve this goal can be described by (1).

$$g(x_A) = m + \frac{1}{x_A} + n\exp(x_A) \qquad (1)$$

Where *A x* is the amplitude of MFL signals, *g* is the function of the grade of pipe material , m , l , n are design Parameters.

Compensation for the effects of sensor velocity is achieved using an adaptive filter as shown in Figure 2. The Adaptive algorithm of the recursive least squares technique is used to adapt the filter weights to improve the defected MFL signals.

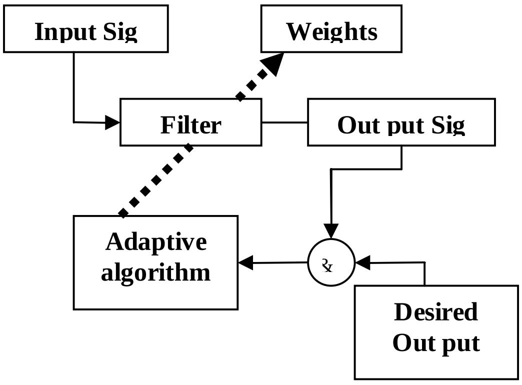

Figure2.Adaptive filter Scheme diagram

3-2. Formulation of an analytical model from MFL defect measurements

If a material is magnetized near saturation, the MFL field generated by a subsurface flaw can be described as follows:

$$H_y(x,y)=\frac{2xy\left(m-2H_a a^2\right)}{\left(x^2+y^2\right)^2} \qquad (2)$$

where m is the dipole moment per unit length this is measured as follows

$$\text{m}=\frac{\sqrt{3}}{2}\text{h} \qquad \text{h}=1.05\times10^{-34} \qquad (3)$$

Where h bar is the plank coefficient ,Ha is the applied magnetic field that is 1 tesla[6] and a is the radius of the defect [7,8]. If the MFL on the surface of a sample is calculated, the variable y is constant and is equal to the depth h of the defect(see fig.3). so the magnitude of h could specify the depth of defect. As mentioned above, it is not necessary to get physical information, like size or position of the defect. If the unknown system and material properties are defined in parameters we obtain $q=h^2$ and $p=2h(m\text{-}2H_a a^2)$, so the following simple fit function for the MFL on the surface of a sample could be illustrated as below:

$$f(x)=\frac{px}{\left(q+x^2\right)^2} \qquad (4)$$

In the developed device the signal is measured by induction coils and for this reasons the measured signal is the derivative in x direction times the velocity of f(x) of measuring device. With regards to the previous equation, the MFL signal becomes as below. in this relation we try to calculate the rate of measured signal in time. So with acknowledge of velocity, that is rate of measuring device distance in time, and by timing this term to deviation of *f(x)* we could reach to rate of delta f to delta t that is rate of depth in time.

$$F(x)=v\cdot f'(x)=v\left(\frac{p}{\left(q+x^2\right)^2}-\frac{4px^2}{\left(q+x^2\right)^3}\right) \qquad (5)$$

On the assumption that the velocity is constant, a new parameter P can be defined as:

$$P=v\cdot p=2hv(m\text{-}2H_a a^2) \qquad (6)$$

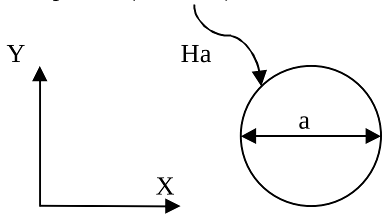

Figure 3. System of coordinates for the calculation of the magnetic flux leakage of a subsurface flaw.

4. Feature extraction for recognition

PCA is a well-known statistical technique for feature extraction. Each $M\times N$ MFL image in the training set was row concatenated to form $MN\times1$ vector x_k. Given a set of training images $\{x_k\}_{k=0,1,\ldots,N_T}$ the mean vector of the training set was obtained as [30].

$$\bar{x}=\frac{1}{N_T}\sum_{k=1}^{N_T}x_k$$

A $N_T\times MN$ training set matrix $X=[x_k-\bar{x}]$ can now be built. The basis vectors are obtained by solving the Eigen value problem:

$$\lambda=V^T\sum\nolimits_X V$$

Where $\sum_X=XX^T$ is the covariance matrix, V is the eigenvector matrix of $\sum_x$ and λ is the corresponding diagonal matrix of Eigen values. As the PCA has the property of packing the greatest energy into the least number of principal components, eigenvectors corresponding to the m largest Eigen values in the PCA are selected to form a lower-dimensional subspace. It is proven that the residual reconstruction error generated by discarding the $N_T - m$ components is low even for small m [31].

As has been said, PCA computes the basis of a space which is represented by its training vectors. The basis vectors computed by PCA are in the direction of the largest variance of the training vectors. These basis vectors are computed by solution of an Eigen problem, and as such the basis vectors are eigenvectors. These eigenvectors are defined in the image space. They can be viewed as images and indeed look like its inherent shape. Hence they are usually referred to as Eigens.

A discrete cosine transform (DCT) is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even. The DCT, is often used in signal and image processing, especially for lossy data compression, because it has a strong "energy compaction" property: most of the signal information tends to be concentrated in a few low-frequency components of the DCT [32]. This function is mathematically explained below:

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \qquad (7)$$

$$k = 0, \ldots, N-1$$

For our reason this function is employed because if its strong behaviour on collecting the important information in low frequencies at the top left of the DCT matrix. the squared low dimensional matrixes could lead to best and rapid decisions in some cases this later could lead to the accuracy rate of 100% in some cases.

4.1 Recognition of defects

The recognition of pipeline corrosion defects in this paper includes preprocessing and classification analysis. The former can be accomplished by recognizing and classifying typical features of signals from magnetic flux images in types of real images or mathematical forms. An approach is to classifying and performs a true decision. For this reason, these are some different kinds of neural networks such as Multilayer Perceptron (MLP), Learning Vector Quantization (LVQ) [9], Self Organized Machine (SOM) [10] and so on. In this work multilayer perceptrons are applied with BP structure.

4-2 Classification for recognition

According to construction of combiners, they are all made of learning process. Therefore to have different combiners different ways of training is essential. The process of learning is based on many ways such as: different ways to show inputs, samples for learning, training process, differ consulting technologies although in this task many theories are offered but each of them should due to some results:

1. The first requirement is that each expert has high level of performance and independently in deciding feature
2. expert has an arithmetic mathematics table to refer this point as strong point of each expert.

Classifying is done by many ways such as: multilayer perceptron, (MLP), radial basis function (RBF), k-mean etc.

This paper presents MLP for classifying. MLP means multi layer perceptron. Classifying is done by neural networks such as MLP. Fundamental work of MLP is to changing weights between layers and each layer has (m) nodes. Number of input nodes is depended on dimension the database. Amount of nods located in hidden layer are subject to change by complicated rate of the expert. In this paper an approach is shown in follows that specifies the number of each layer this equations for this reason is earned exrimentally but the result of this employment is satisfied. In training situation the weights are subject to change until reaching the best weights .The number of training situations is determined by the number of epochs it is kept done until less mistakes appears in out put.

in this algorithm three Networks with the names of +1 0 -1 are employed. all of these three experts are learned by a same set of database and the result of classification is achieved by voting the triple output.

5. Employed algorithm

We have applied similar algorithm to SSCE[11]to database of MFL signals. In this map we apply preprocessing to the crude data this section is discussed and as a brief it contains extracting different kinds of defects from physical formulation and normalization then three classes perform a decision on the their inputs, the rate of each of which is composed by a voter to achieve a well decision.

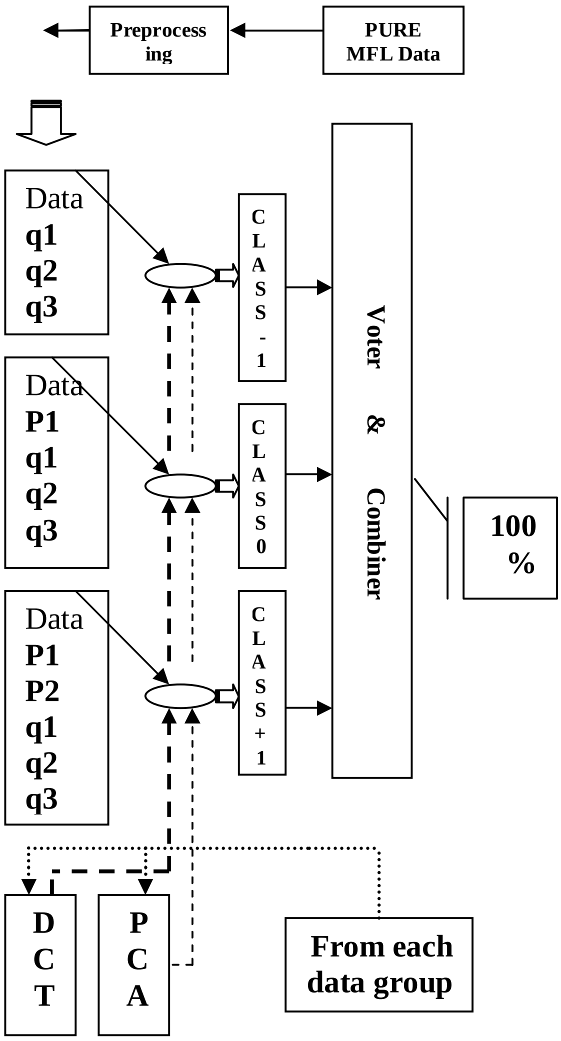

Figure4.Devised algorithm

6. Results and discussion

In order to investigate the statistical distribution of the error rate, three neural networks with the same structure and transfer functions(but with different number of neurons that are referred to initial state) were trained with the not same data set [12,13,14].In this approach each expert is trained to recognize o sort of defect so that each of which experts in final are tried to find three common sort of defects. Then the accuracy rate of each network is calculated. To calculate different numbers of input parameters were trained and compared to the network described in the above sections. The following experiential rule was used to define the structure of the networks:

$$
\begin{aligned}
&N_{input} = 2\times P \\
&N_{hidden} = approx\left(N_{in} + N_{out}\right) \qquad (8)\\
&N_{out} = 2\times P
\end{aligned}
$$

Where N is the number of neurons in the corresponding layer and P is the number of input parameters that could be even or odd.

In this project first we try to test a simple network by different characteristics and then we design three experts.

In some information about the set of trained networks is given by accuracy rate as well as the worst and the best network, respectively. Furthermore, maximum or minimum of the average of output of each network in ten times training is mentioned. Summary of the network performance for different input parameters is as follows:

P1, p2, q1, q2, q3

As is demonstrated in the table below there is q1, q2, q3, P1, P2 parameters. These parameters are described as follows:

$P = v\cdot p = 2hv\,(m-2H_a a^2)$

$q = h^2$

P1= P for h=0.002 & a=0.001 [m]

P2= P for h=0.003 & a=0.0015 [m]

q1=q for h=0.002 [m]

q2=q for h=0.003 [m]

q3=q for h=0.004 [m]

Comparison And details discussion	Class1 / Class2 / Class3												**CON CLUS ION**
	Clas1												
F.E.U. / Data/Res	PCA								DCT				**DCT**
Types of F.E.U.	10				15				4 sq				**6 sq**
Best result with best structure	63.6				83.3				90.0				**100.00**
	Expert details				Expert details				Expert details				***
Class structure (hidden neurons)	8	10	16	20	8	10	16	20	10	16	10	16	**16**
Avg result (10 times)	37.7	40.3	50.0	42.5	50.7	55.0	61.3	53.3	65.0	66.6	73.3	83.3	**83.3**
Worst result	10.0	17.0	30.0	22.0	20	23.3	40.0	30.0	50.0	55.0	60.0	53.3	**53.3**
	Clas2												
F.E.U. / Data/Res	PCA								DCT				**PCA**
Types of F.E.U.	15				20				4 sq		6 sq		**20**
Best result with best structure	73.6				100.00				93.3		95.0		**100.00**
	Expert details				Expert details				Expert details				***
Class structure (hidden neurons)	8	10	16	20	8	10	16	20	10	16	10	16	**16**
Avg result (10 times)	12.0	32.6	45.5	40.0	62.5	37.6	60.3	25.0	77.0	70.0	95.0	87.7	**60.0**
Worst result	7.0	19.3	23.0	11.2	21.0	12.5	21.0	17.6	58.7	57.3	63.5	60.0	**21.0**
	Clas3												
F.E.U. / Data/Res	PCA								DCT				**PCA**
Types of F.E.U.	22				25				4 sq		6 sq		**25**
Best result with best structure	86.6				100.00				83.0		90.0		**100.00**
	Expert details				Expert details				Expert details				***
Class structure (hidden neurons)	8	10	16	20	8	10	16	20	10	16	10	16	**20**
Avg result (10 times)	53.3	47.7	50.6	37.0	55.5	44.2	56.6	43.3	66.6	67.9	73.5	90.0	**43.3**
Worst result	32.0	27.0	14.0	17.6	19.5	10.7	17.0	22.0	41.0	30.0	52.2	45.5	**22.0**

Table 1.Three Networks performance by different parameters.

So by training classes with these triple clusters of data, class one could discriminate features of depth, better than radius and momentum.

6.1 Historical discussion

To date, all published research based on the analytical model of dipolar magnetic charge, [17, 18-20,21], this later is discussed before and defined as *m* parameter.but for an exception, reference [22] is presented.this reference is just discussed a single defect. , The often encountered practical situation of two adjacent defects is also discussed only by Uetake and Saito[22], but their study is limited to slots with parallel walls, of a maximum of 4mm in length. In this study we consider a multiple defect case. That is consist of triple recognition we claim that this algorithm could satisfy almost all of defects. With increase in computational capabilities, finite element analysis can now compete with analytical methods. Since the proceeding numerical modeling of MFL phenomena is exposed by Lord and co-workers [23, 24, 25], the finite element analysis of defect-induced magnetic signals has become increasingly popular. In oppose of the significant progress made in this area to include non-linear material properties [26,27,28], a quantitative relationship between magnetic leakage field and defect length has not been clearly specified. Furthermore, numerical modeling involves a direct MFL approach, since it includes predefined defect geometries and material characteristics. Calibration of the MFL signals in terms of defect depth has been studied both through finite element modeling [24, 25, 26, 28] and through analytical methods based on dipolar magnetic charge [25, 29]. Two of the numerical analysis studies [24, 28] correctly predicted that the amplitude of the normal MFL signal Component increases with defect depth, and that the separation between the extreme MFL values is directly proportional to the Defect length In this paper, with regards to previous works, a new simple algorithm is applied that could exactly determine defects with various shapes. For problem of encountering different kinds of defects we initializes deferent defects with three classes which each of them tries to learn a defect with determined characteristics. These features are an estimate of three large groups of defects.

7. Conclusion

In this study, we have discussed intelligent defect recognition directly from MFL signals. An analytical model is employed to account defects in order to correlate the normal component MFL profile with the defect dimension along the Impregnating magnetic field. The efficiency of the model was confirmed through experimental results in MFL defect detection. A clear advantage of the method presented here is the low number of parameters that have to be considered. for a satisfactory estimation we classify all the defects in three groups with different shapes in this case all the defects ranged to depth of 2 till 4 millimeter and radius of 1 up to 1.5 millimeters. These later are subject to recognize. For this reason three expert systems were learned to recognize the request. And at the end. voter starts to vote between the results of three experts. the accuracy rate of 100 percent shows the efficiency of the mentioned devised algorithm.

8. References

[1] A. Bergamini, Nondestructive testing of stay cables, IABSE conference on cable-supported bridges(2001), pp. 312–313.
[2] A. Bergamini, Nondestructive testing of stay cables—field experience in South East Asia, Third World conference on structural control vol. 2 (2002), pp. 1057–1064.
[3] M. Afzal and S. Upda, Advanced signal processing of magnetic flux leakage data obtained from seamless steel pipeline, NDT&E Int 35 (2002) (7), pp. 449–457. SummaryPlus | Full Text + Links | PDF (428 K) | View Record in Scopus | Cited By in Scopus (15)
[4] P. Ramuhalli, L. Udpa and S.S. Udpa, Electromagnetic NDE signal inversion by function-approximation neural networks, IEEE Trans Magnetics 38 (2002) (6), pp. 3633–3642. Full Text via CrossRef | View Record in Scopus | Cited By in Scopus (20)
[5] R.R. da Silva, S.D. Soares, L.P. Caloba, M.H.S. Siqueira and J.M.A. Rebello, Detection of the propagation of defects in ressurized pipes by means of the acoustic emission technique using artificial neural networks, Insight 48 (2006) (1), pp. 45–51. Full Text via CrossRef View Record in Scopus Cited By in Scopus (3)

[6] C Mandache, B Shiari and L Clapham"Defect separation considerations in magnetic flux leakage inspection" Insight Vol 47 No 5 May 2005 pp.271
[7] D.E. Bray, Nondestructive evaluation (revised ed.), CRC Press, Boca Raton, FL (1997).
[8] R. Christen, A. Bergamini and M. Motavalli, Three-dimensional localization of defects in stay cables using magnetic flux leakage methods, J Non Destructive Eval 22 (2004) (3), pp. 93–101.
[9] Martin Golz, David Sommer," The Performance of LVQ Based Automatic Relevance Determination Applied to Spontaneous Biosignals, KES 2006, Bournemouth, UK, October 9-11, 2006. Proceedings, Part III, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2006. pp. 1256-1263
[10]Hiroshi Wakuya, Hiroyuki Harada, Katsunori Shida "An architecture of self-organizing map for temporal signal processing and its application to a Braille recognition task" Wiley Periodicals, Inc. Syst Comp Jpn, 38(3):

62- 71, 2007
[11]Saeedreza Ehteram, Seyed Z. Moussavi "Semantic Supervised clustering to land Extraction on satimage database" journal of Global engineering , science and technology (GESTS),Seul korea March 2007, pp.117-125
[12]Saeedreza Ehteram, Ali Sadr , Seyed zeinolabedin Mousavi "Rapid face recognition by regional feature extraction ", INISTA 2007 conference Istanbul Turkey20 – 23 June2007, pp. 262-269, 2007.
[13] 14. R. Ebrahimpour, S. R. Ehteram, E. Kabir, "Face Recognition by Multiple Classifiers, a Divide-and-Conquer Approach ", Lecture Note in Computer Science (LNCS), vol. 3686, pp. 225-232,September 2005.
[14]R. Ebrahimpour, Seyed Zeinolabedin Moussavi, and Saeedreza. Ehteram (Iran)" Multiple Binary Classifier Fusion (MBCF) in Application of Satimage Database" IASTED from proceeding (522) Applied Simulation and Modeling 2006-Greece

[15] Y. Chung Bang, L. Jong Won, K. Jae Dong and M. Kyung Won, Damage estimation method using committee of neural networks, Proceedings of the SPIE—the international society for optical engineering vol. 5047 (2003), pp. 263–274.
[16] Qi Jiang, Qingmei Sui, Nan Lu, Paschalis Zachariades, Jihong Wang." Detection and estimation of oil gas pipeline Corrosion defects "http://corporate.coventry.ac.uk/conten
[17]Dobmann G and H¨oller P 1980 Research Techniques in Nondestrucrive testing R. S. Sharp (New York: Academic) vol IV, pp.39–69
[18]Shcherbinin V E and Pashagin A I 1972 Defektoskopyia pp.874–82
[19]F¨orster F 1986 NDT Int. 19 3–13
[20] Edwards C and Palmer S B 1986 J. Phys. D: Appl. Phys. 196pp.57–73
[21] Mandal K and Atherton D L 1998 J. Phys. D: Appl. Phys. 31 pp.3211–17
[22] Uetake I and Saito T 1997 NDT & E Int. 30 pp.371–7
[23] Hwang J H and Lord W 1975 J. Testing Eval. 3 pp.21–5

[24] LordWand Hwang J H 1977 Br. J. Non-dest. Testing 19 pp.14–18
[25] Lord W, Bridges J M, Yen W and Palanisamy R 1978 Mater.
Eval. 36 pp.46–54
[26] Atherton D L and Daly M G 1987 NDT Int. 20 pp.235–8
[27] Patel U and Rodger D 1995 IEEE Trans. Magn. 31 pp.2170–3
[28] Altschuler E and Pignotti A 1995 NDT & E Int. 28 pp.35–40
[29]Philip J, Rao C B, Jayakumar T and Raj B 2000 NDT & E Int.
33 pp.289–95.
[30] M. Turk, A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, *vol. 3, pp. 71-86, 1991.*
[31] Duda, R.O. and Hart, P.E., Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.
[32] Ken cabeen, peter gent."Image compression and the discrete cosine transform" math45 collage of Redwoods, pp1,2.