Improving the Reliability of Automated Non-Destructive Inspection

Nick BRIERLEY, Trevor TIPPETTS, Peter CAWLEY
Mechanical Engineering Department, Imperial College London, South Kensington Campus, SW7 2AZ, United Kingdom; Phone: +44 207 5947227, Fax +44 207 5945709; e-mail: n.brierley09@imperial.ac.uk, t.tippetts09@imperial.ac.uk, p.cawley@imperial.ac.uk

Abstract
In industrial NDE it is increasingly common for data acquisition to be automated, driving a recent multiple increase in the availability of data. The data analysis remains a mostly manual task, performed by a skilled operator - a rather painstaking task given that much of data contains no indications. Partial automation, using software to prioritise regions of interest, could simultaneously increase inspection reliability and decrease data analysis time, by optimising the use of the operator’s time. The project output is a software system general enough to fit a wide array of NDE applications. Nonetheless, the work so far has focused on two specific examples of automatic NDE: the ultrasonic inspection of power station rotor bores and the ultrasonic immersion inspection of aerospace titanium turbine discs. The research incorporates elements of image-processing, machine vision, optimisation and information theory, as well as of course the theory of ultrasonic NDE. The paper will discuss the key issues in the development of the software and show initial results from the program.

Keywords: ultrasound, automation, reliability, optimisation, registration, data fusion

1. Introduction
Automated data acquisition is increasingly common in industrial NDE, driving a recent multiple increase in the availability of data. This is perhaps especially apparent for a single data stream in the particular case of ultrasonics, the technique central to much of this project, where the stored data now often consist of a full sequence of time-amplitude traces (A-scans), rather than several signal amplitudes at a particular gate. The problem may be further exacerbated by the acquisition of multiple channels of data, using different techniques, viewing angles or probe types, including many-element array probes.

The collected data need to be analysed and currently this is largely done manually by a skilled operator sifting through countless images displaying a subset of the data. Given the expense involved in doing NDT, there is a high probability of an inspected part being fit for continued service, otherwise it would in all likelihood be replaced without inspection. Consequently, the operator has to spend a large proportion of his/her time examining data where there are no defects to report. The resultant monotony is liable to increase operator fatigue and hence the probability of a defect indication being missed.

The current, typical output of a UT scanning package shows multiple views that allow detailed examination of data, but does not:
- Reduce the volume of data to be examined
- Prioritise regions of interest
- Fuse information in areas covered by multiple inspections or sensors
- Identify changes since the last inspection
There is therefore scope for the inspection reliability to be improved while reducing the time taken for the data analysis through partial automation, with significant potential cost savings. The project sets out to develop a system, ultimately to be implemented in commercial-grade software, for the analysis of automatically acquired data from a wide range of inspection types with the primary aim of improving the reliability of the overall inspection.

Despite the general applicability of the final project output, work so far has focused on two particular applications: firstly the Rolls-Royce mid-manufacture immersion ultrasonic inspection of titanium disks, using multiple passes of a single probe at different angles to search for defects, specifically hard-alpha inclusions, and secondly, RWE Npower’s regular, in-service inspection of turbine and generator rotor bores, using a bore scanner that carries four different ultrasonic probes and an eddy current array.

This paper will give a high-level description of the project goals in the Project Objectives section. The Current Development State section describes the data processing efforts to date, focusing on data registration and fusion. Next Steps outlines the remainder of the analysis software, including the metrics that quantitatively measure the reliability improvement.

2. Project Objectives

The overall project aim of improving inspection reliability is to be achieved by segregating inspected parts into regions that are defect free to a very high probability and regions that are of questionable integrity. This will be accomplished by maximising the use of information in new and historic data and applying an understanding of inspection physics. The prioritisation of regions for human examination will allow operators to focus where their skills are best applied, reducing fatigue and human errors. Statistical analysis will monitor the Probability of Detection (POD) and Probability of False Alarm (PFA) of the detection software. This self-evaluation capability will demonstrate the software's effectiveness and also make it possible to recommend improvements to inspection procedures.

The project does not aim to achieve automatic classification of defects as some researchers have attempted (see review [1]), and it does not use artificial intelligence (AI) [2]. This choice was made on the grounds that the performance of such systems is critically dependent on the quantity and quality of training data, but the availability of these data will not be adequate in many NDE applications.

3. Current Development State

3.1 Registration Introduction

To start the data processing, the raw data within each data channel are introduced into a geometry model. This builds up a volumetric field of data, incorporating knowledge of both the amplitude data as well as the spatial relationships of sample locations. Inevitably, the geometric model used is entirely application dependent, but the software’s modular structure allows for functionality given different inspection types.
The aligning of datasets to a common coordinate system, known as registration, is essential for a comparison or combination of data sets, be this between-channel, using past inspections, or with simulated inspections. There are a great number of papers concerning registration, particularly 2D image registration, and often in a medical context [3]. The publication [4] identifies four major steps in most image registration procedures: feature detection, feature matching, transform model estimation, and image re-sampling and transformation. These steps also apply to the registration of volumetric NDE datasets as considered here:

- **Feature extraction:** given the huge size of many of the datasets, individually, and especially when processing several together, it is essential for computational efficiency to select a subset for registration calculations. The features extracted from the data should be associated with component geometry and therefore detectable across the different datasets, despite the possible distortions that the registration seeks to compensate for.

- **Feature pairing:** features must be matched across datasets so meaningful comparisons between datasets are possible - see illustration in Fig. 1. For this the extracted features are summarised in terms of suitable descriptors [5], and these are used to pair up similar features across two datasets, but bearing in mind the possibility of outliers [6].

- **Evaluation of a registration error metric:** the quality of fit for a given registration transformation parameter combination is quantified by evaluating a suitable metric over all the identified feature pairs. Mutual information is an example of a possible registration metric [7].

- **Optimisation:** the registration error metric must be minimised as a function of the parameters governing the dataset alignment, to find the best possible combination. The optimisation must be completed carefully to reach the global optimum and avoid any local minima of the error metric space. A wide range of generic algorithms may be used for tackling such problems, including multi-start [8], pattern search [9], and particle swarm optimisation [10].

- **Transformation:** once the best combination of parameters has been determined, this is applied to the dataset to achieve registration.

![Figure 1](image-url)  
*Figure 1. Illustration of feature matching across datasets - example from the UT rotor bore inspection. Arrows point from each feature in Dataset 1 to the matching feature from Dataset 2.*

Each of these stages involves a number of challenges that differentiate them from their typical 2D image-processing equivalents [11].
3.2 Registration Optimisation

The registration optimisation to find the global minimum of the error metric is challenging due to complex features in optimisation space. The problem is illustrated in Fig. 2, a cross-section through the optimisation space in the plane where the rotational and axial offsets of the scan start position in the rotor bore inspection are varied whilst other registration parameters are held at their correct values. The figure shows a very sharp minimum that is surrounded by large areas of zero gradient. This registration metric compares features from the two data sets that overlap each other. Only a small region in the geometry parameter space produces overlapping features, so the metric is constant outside of this basin of convergence. Traditional gradient-descent operations would fail unless they initialise within the small convergent region.

![Figure 2. Registration metric 1. A cross-section through the multi-dimensional optimisation space, one dimension for every registration parameter, plotting variations with axial and circumferential offsets in scan start positions for an example input from the UT rotor bore inspection. The colour scale gives the value of the metric. This metric is based on the root-mean-square residual between paired features.](image)

The authors have found that combining different metrics in a multi-objective optimisation allows the advantages of each metric to overcome the disadvantages of others [12]. The principle is illustrated by Figure 3 – it shows an equivalent surface to Fig. 2 for a different error metric. This second surface is noticeably smoother and does not have any zero-gradient regions, so locating the global minimum is easier, even though this minimum is less clearly defined. Considering optimising over a simple combination of these optimisation spaces, such as a weighted sum, it becomes apparent that the resultant surface will be much more suitable for an optimisation algorithm's search than either of the component spaces in isolation.
As there is no limit on the number of metrics included in the optimisation, and hence dimensions in the objective function space, further metrics can be introduced to allow the incorporation of prior knowledge about the most likely correct parameter combination. For example, if the optimiser is unbounded, a metric can be used to down-weight parameter combinations that lie outside of the anticipated domain.

Of particular significance in the case of inspection path geometries, such as the helical rotor bore scan, where data is recorded with no discontinuities unequivocally demarcating adjacent probe passes, is the parameter that describes the number of samples (A-scans) recorded per “scan line”. If this parameter is far from its correct value, the volumetric representation of the data is grossly distorted, more so than for other parameters as this one affects spatial adjacency. Hence the features identified for registration may be unrepresentative, undermining any optimisation attempt. To overcome this problem, a metric based on the parameter difference to a pre-computed “best guess” is introduced to the optimisation: the authors established that such an estimate could be found using autocorrelation to interrogate the amplitude periodicity in the scan-path direction. Figure 4 illustrates the effectiveness of this autocorrelation calculation.
The workings and outputs of multi-objective optimisation are further explained under the heading of Detector Evaluation (section 2.1), in the context of optimising the Receiver Operating Characteristic (ROC) of an inspection.

3.3 Data Fusion

Once registration is complete, data from multiple sensors and/or scans of the same part can be fused in regions of overlapping coverage. Other researchers have proposed a range of techniques for such data fusion, and have demonstrated the possible reliability gains [13, 14]. The processing could also entail data subtraction to detect changes over time or compared to a standard reference component as in [15]; baseline subtraction used in structural health monitoring (SHM) is an analogous problem, e.g. [16]. The authors are currently exploring using a variant of the Total Focussing Method (TFM), familiar from the literature on ultrasonic arrays [17, 18], as a means of combining different A-scans within and across ultrasonic channels.

4. Next Steps

4.1 Inspector’s Software Output

A detection algorithm identifies indications in these fused or subtracted data and assigns to each indication an estimated probability that it represents a reportable defect. The final output to a human inspector is a sequential rendering of indications, ordered by severity. Two probabilities are reported for each indication: (1) that the indication is actionable and (2) that there are actionable indications yet to be seen. These probabilities help the inspector decide when to terminate the viewing of the inspection data. Figure 5 illustrates this inspection sequence.

Figure 5. Illustration of the inspection procedure with the project software. After reviewing each indication, the inspector decides either to continue to the next or to stop. Previously viewed indications and the probe scan path (helical – for the case of the rotor bore inspection) provide references for orientation in the 3D visualisations.
4.2 Detector Evaluation

The Receiver Operating Characteristic (ROC) plot is perhaps the best way to characterise detector performance [19]. A detector can make two types of errors: false positives and false negatives. NDE inspectors obviously want to reduce both the probability of a false positive (PFA) and probability of false negative (1-POD), but these are generally competing objectives. The ROC plot shows the trade-off between PFA and POD. Figure 6 shows a simulated ROC curve as an example.

A perfect detector would have zero probability of a false alarm and would also detect a real flaw with a probability of one. The (0,1) point in the upper left corner of the ROC plot therefore represents the best possible detector performance. Ideally, the ROC curve should come as close to that point as possible. In the past, NDT practitioners have sometimes plotted a curve in the ROC space by sweeping through values of a single, relatively simple parameter such as amplitude threshold. An alternative approach is to hold PFA constant while optimising POD as a single objective, or *vice versa*.

![Simulated ROC plot](image)

Figure 6. Simulated ROC plot. Each point represents a detection algorithm with a specific set of detector parameter values. The point locations show their POD and PFA for a data set with known flaws. The dominant, optimal points are shown as circles, while dominated detectors are plotted as crosses.

The parameter-sweeping and single objective paradigms are insufficient for data-fused detector evaluation. Data fusion and indication detection algorithms can be much too complicated to sweep through parameters or to optimise one objective at a time. Detection algorithms often have many more free parameters and optional settings than a detector operating on a single data set; there is not necessarily a monotonic relationship between a detector parameter and PFA or POD. The various parameters can interact in complex ways, and there is typically no way to hold either PFA or POD constant.
Following [20], we instead treat the ROC plot as the output of a multi-objective optimisation. Detector parameters define a search space, such that each point in the space defines a unique detector. The optimiser samples points in this space and calls simulation code that solves the forward problem: estimating PFA and POD for each sampled detector. The optimisation solves the inverse problem by seeking sample points that improve the PFA and POD objectives simultaneously. The software records a Pareto frontier, or optimal points in the parameter space, which collectively represent the best achievable detectors.

The resulting dominant points have an optimal trade-off between the competing objectives. Inspectors can choose detector parameter values from the dominant set. The dominated points are detectors that are worse than at least one other detector in POD, PFA, or both. Therefore, inspectors should never use dominated detectors. The area of the dominated region, equal to the integral of the ROC curve, is a performance metric of the detector class as a whole. A good detector class would approach the maximum possible value of 1.

This approach can work with arbitrarily complex data-fused detection algorithms. A multi-objective optimisation can also include other objectives, in addition to PFA and POD, very easily. For example, a quantity measuring accuracy in predicted location, size, or characterisation readily extends the two-dimensional PFA vs. POD. Black-box multi-objective optimisers can solve the higher dimensional problem without modification. The authors’ latest optimiser is able to efficiently explore the whole Pareto front, providing a significant range of optimal parameter combinations.

5. Conclusions

Advances in NDE technology have vastly increased the amount of available data through automated acquisition. However, many applications still rely on slow, human-intensive analysis and do not use potential correlation information from overlapping data sets. A methodology for automated processing is being developed that screens and prioritises data for evaluation by human inspectors. The implementation software searches for optimal detection parameters by improving PFA and POD simultaneously in a multi-objective optimisation algorithm.

Acknowledgements

The Engineering and Physical Sciences Research Council (EPSRC) funding for an Engineering Doctorate studentship for Nick Brierley is gratefully acknowledged, as is the Royal Commission for the Exhibition of 1851 Industrial Fellowship for Nick Brierley. This research is also financially, technically, and practically supported by RWE Npower, Rolls-Royce and the wider UK Research Centre in Non-Destructive Evaluation (RCNDE).

References