

The Snooker Algorithm for Ultrasonic Imaging of Fatigue Cracks in order to use Parameter-Spaces to Aid Machine Learning

Dr Channa Nageswaran CEng FIMechE

TWI Ltd, Granta Park, Great Abington, Cambridge, CB21 6AL, UK

channa.nageswaran@twi.co.uk

Abstract

Fatigue cracks in a wide array of industrial components and structures pose a significant threat to their integrity. Detecting fatigue cracks using ultrasonic inspection techniques is a widespread activity for economic reasons but there are limitations to the techniques due to the morphology of fatigue cracks. In addition to detection there is a need to measure the size of the cracks which are often within the volume of the material. Ultrasonic techniques are well-suited to look inside the volume of the material but achieving sufficient sensitivity to the tip of the cracks in particular is practically difficult. Without an accurate knowledge of where the tip of the crack lies there can be significant uncertainty in sizing measurements. Machine Learning (ML) techniques are being developed to aid in the inspection and monitoring tasks but presenting the ultrasonic data in a suitable way for machine learning is very important. This paper presents a new approach to condition the ultrasonic data for machine learning settings so that they can be used effectively and confidently to detect and size fatigue cracks. The new approach, using images termed parameter-spaces, will also aid in conventional inspections as they are able to give information to human operators as to the existence or not of these very dangerous cracks.

Keywords

Fatigue cracks, ultrasonic inspection, machine learning, imaging, techniques

Introduction

Fatigue cracks can emerge in structures and components subject to cyclic loading conditions. The maximum stress levels are often below the yield stress of the materials and so detecting and managing any cracking is critical. The consequences of failure due to fatigue cracking can be catastrophic both to human life, the environment and the economy. Ultrasonic inspection is an established method for industrial use. Along with industrial radiography, ultrasonic techniques are the most commonly utilised options for interrogating the volume of materials to detect discontinuities such as cracks. However, only ultrasonic techniques have sufficient sensitivity to cracking and are able to be deployed for inspection during service in comparison to radiography, which usually requires the component to be brought to the facility for inspection. Hence ultrasonic techniques are typically used in a wide variety of configurations for inspecting components in service, which range from aircraft, ships, bridges, oil & gas installations and nuclear power plants.

Detecting fatigue cracks can often be straightforward because they commonly originate on the surface of the component, which can be done using both ultrasonic and surface techniques, such as dye penetrant inspection, magnetic particle inspection or eddy current array techniques. However, the morphology of fatigue cracks within the volume, sub-surface, makes detection and then sizing them problematic. The tips of fatigue cracks can be very sharp, often a few tens of microns in dimension, which is much smaller than the wavelength of typical ultrasonic inspection frequencies, often limited to 10MHz for practical use in industrial conditions. The interaction mechanism between the tip and the propagating sound wave can give rise to diffracted signals which can be detected by the ultrasonic instrumentation, allowing for sizing to be performed. However, very often, the diffracted signal is not generated or recorded by the instrumentation, making both detection of the tip and therefore sizing impossible. In addition, uncertainties in measured sizes can be significant when compared to actual sizes. These factors give rise to severe issues surrounding decisions on allowing structures to continue operation, repair, replacement and other management strategies.

Hence there is a need to develop and validate reliable techniques for detecting and measuring the size of fatigue cracks which may emerge in critical infrastructure. In recent times, the field of artificial intelligence has been rapidly developing tools and techniques to allow computing systems to make decisions and make measurements without human input. A key technology finding use in many arenas is machine learning (ML) where multi-level artificial neural networks, synonymous to the neural network of the human brain, is trained in an activity using examples. Once trained, the ML system is able to make decisions on input data without the need for a human interpreter. There are many industrial and economic benefits if such systems can be deployed.

However, there are several barriers to deploy such systems in the inspection industry at present. Primary amongst them is the availability of training datasets of sufficient quality to train ML systems effectively. Another is the nature of the data, which originates as a voltage signal over time, the so-called A-scan. From the A-scan, which is provided by the ultrasonic instrumentation, a variety of imaging schemes have been established in the inspection industry to aid the human operator to interpret the signals, to decide whether or not a discontinuity exists and if so, make measurements. The two key parameters that can be measured from the signals (and from the images) are the amplitude and the time-of-flight (TOF), which is often converted to distances when the velocity of the sound in the material is known. However, these images may not be well-conditioned for training ML systems as the information is rich but difficult for the deep neural networks to process effectively, requiring large training datasets which is difficult to acquire in the inspection industry.

In this paper, as the first step towards training ML systems, an approach to process the data into a new kind of image, termed parameter-space, is presented. The differences between the traditional images used by human operators and representation in the parameter-space is discussed. The parameter-space representation is considered generic, in that a number of different parameters could be used. For the example presented in this paper the approach used to generate the parameter-space image is described. In a subsequent paper, as the next step, the results of training an ML system using parameter-space images will be presented. See 'Supplementary Material' below for more information.

Ultrasonic Imaging Using Array Transducers

Nageswaran (2018) presented a critical review of ultrasonic techniques typically used for detection of a cracking mechanism termed high temperature hydrogen attack, but the analysis is equally valid when aiming to detect fatigue cracking as well as many other types of flaws. Nageswaran et al (2013) presented a project where a technique was developed to detect the onset of fatigue cracking in a large subsea threaded connector used in the offshore oil and gas industry. The integrity of this component

is very critical to ensure safe and pollution free operation, and hence the work described in the paper was aimed to ensure that the failure of the component under various different fatigue loading scenarios was well understood so that an effective monitoring programme could be put in place when the component went into service. This work illustrates the typical approach to manage fatigue cracks.

The reference image for this paper makes use of a technique termed Total Focusing Method (TFM). In order to create a TFM image, the data from the material needs to be collected using a technique called Full Matrix Capture (FMC). This approach has been recently standardised in ISO 23865 (2021) and an increasing number of ultrasonic instruments used in the inspection industry are able to implement data collection by FMC and then use that data to generate TFM images. This approach is considered the reference because it is able to achieve high imaging resolution. During the FMC process, each element of an array transducer (defined in Figure 1) is used as a transmitter to send a sound wave into the material. The sound wave interacts with the material and its contents, which includes the microstructure as well as gross discontinuities like porosity, and cracks. The echoes thus generated are then received and recorded on all elements of the array. Each transmit-receive pair of elements represents the basic A-scan signal of signal amplitude vs TOF. Hence a matrix of signals representing all transmit-receive paths is created. The data in this matrix is then processed by TFM. The algorithm for TFM is summarised below, with further information in Nageswaran (2018) and ISO 23865 (2021).

1. Select the region within the material to be imaged, which is termed the Region of Interest (ROI). The ROI is typically a 2D plane to minimise computational requirements but can be 3D.
2. Divide the ROI into a grid. This grid represents the resolution of the image and a grid point size is typically sub-millimetre. The number of grid points will also impact calculation times, and so in current instruments the total number of grid points is limited to between about 200,000 to 1,000,000. However, when the data processing is done on computers there can be no limit to how fine the grid spacing is set, except for the amount of time available for computation.
3. The ROI is typically a 2D matrix which will be the image. The elements of this matrix – which represent the grid points, and to be consistent with imaging terminology, will be referred henceforth as pixels - will contain the amplitude of echoes, represented on a colour scale. Start by setting the pixel values to 0, which represents the start of the summation process.
4. For each pixel, cycle through all transmit-receive pairs in the FMC data matrix. In the A-scan of each transmit-receive pair, calculate the TOF from the transmit event (typically at the start of the signal) to the position of the pixel, where it is assumed a wave interaction takes place, and then from the pixel to the receiver.
5. Take the amplitude value at the resultant TOF position in the A-scan and sum it to the amplitude value contained in the corresponding pixel of the ROI image matrix.
6. Once all transmit-receive pairs of the data matrix are processed as above for one pixel, repeat the process with all other pixels.
7. Once all pixels have been processed, visualise the resultant ROI matrix, which is the TFM.

The above process is described in Equation [1]:

$$I_{TFM} = \sum_{p=1}^P \sum_{1}^{n^2} A_{TOF_p}$$

In Equation [1]:

I_{TFM} is the TFM image of the ROI;

P is the total number of pixels, which is the same as the total number of grid points in the ROI;

n is the number of elements on the ultrasonic array transducer (see Figure 1);

A_{TOFp} is the signal amplitude at the calculated time-of-flight in the A-scan from transmitting element to a pixel p and then to receiving element.

There are two explicit loops for summation in Equation [1], as well as calculation of TOF in each case that involves simple geometry when sound propagates in an isotropic medium; which is the case in majority of engineering materials, but there are cases when this assumption is not true. As indicated earlier, the array transducer is composed of elements. Total number of elements, n , for industrial inspection with FMC/TFM is typically 32 or 64. Hence the inner loop could contain 4096 steps. A typical ROI size is 30mm by 60mm, at a grid size of 0.1mm, so that the total number of pixels, P , is 180,000. Hence, a lower estimate for the number of total steps is 180,000 times 4096, giving over 737 million total calculation steps. However, with optimisation and parallel calculation technologies such as field-programmable gate arrays and graphics processing units, this level of calculation is not daunting to modern instruments. Near real-time imaging is possible for typical industrial ROI sizes, but for complex techniques and configurations computational limitations still exist. Of more practical concern is management of FMC data, if it is elected to be stored, which can quickly become unmanageable. For this reason, FMC data is often discarded once it has been used to generate the required TFM image. TFM is a generic term used to cover many imaging configurations, as presented in ISO 23865 (2021). The key difference between different imaging configurations is the calculation of TOFs for each pixel.

Figure 1 illustrates the case considered in this paper, showing the array transducer and the position of fatigue cracks. In addition to fatigue cracks, two notches are used as representative flaws, as well as a drilled hole, termed side drilled hole (SDH) which is different in morphology to the fatigue crack and the notch. Figure 2 shows images of some of the actual cracks used for illustration in this paper. Whereas the notches and fatigue cracks are considered to be planar in morphology, the SDH is considered a volumetric flaw whose reflection characteristics are different to the planar flaw types. The fatigue cracks were generated in the Fatigue Laboratory of TWI, as shown in Nageswaran (2013b).

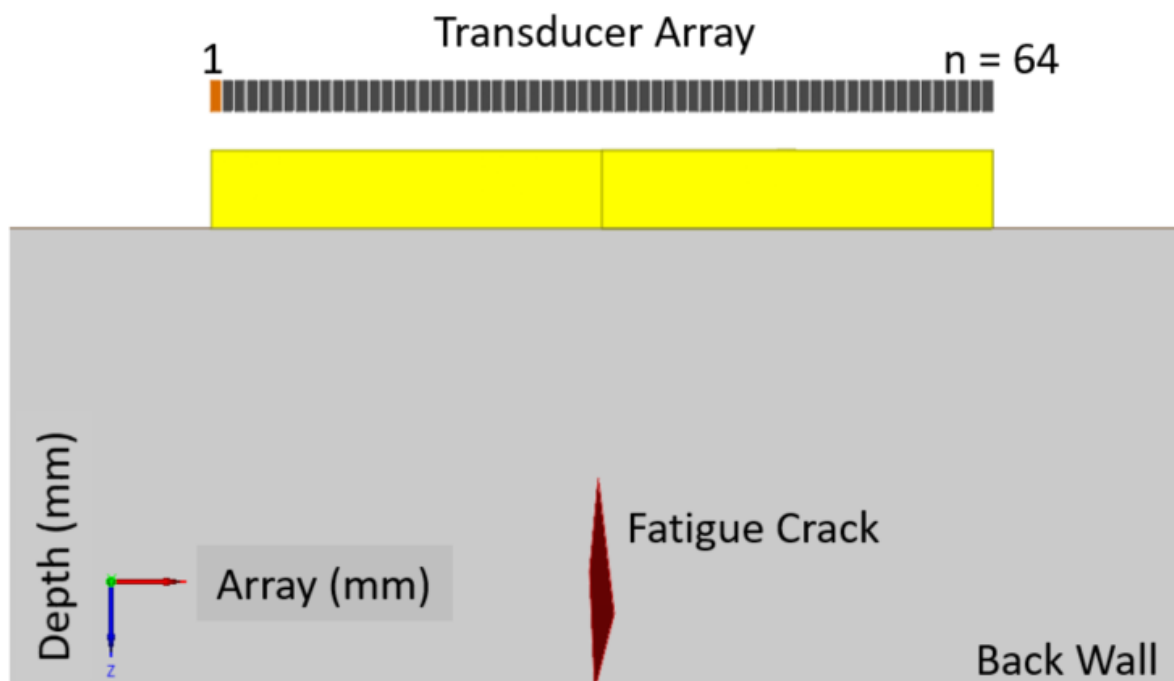


Figure 1 Illustration of the ultrasonic transducer array composed of 64 elements and the position of the fatigue cracks at the back wall surface opposite to the one where the transducer array is placed.

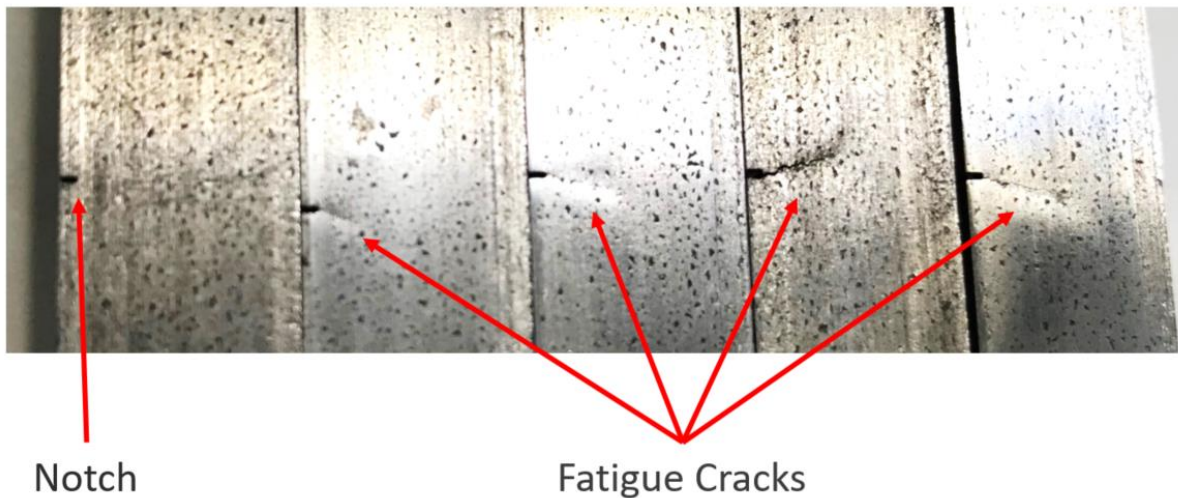


Figure 2 Image of a notch, which was machined, and several fatigue cracks generated for the project. Due to the nature of fatigue cracks they are often difficult to observe visually without the use of enhancements, such as dye penetrant inspections done on cleaned surfaces, or through microscopes. See 'Supplementary Material' below for more information.

Parameter-Space Images

There are several options available as parameters when undertaking imaging and utilising ML techniques. There are parameters related to the imaging scene, such as number of elements, frequency of the sound, pitch of the elements (the distance between adjacent elements) and ROI characteristics. There are parameters related to ML such as the number of layers and nodes in the network as well as manipulation of weighting functions (Google, 2021). Considerable care and thought needs to be taken in order to select parameters suitably for any given scenario, and ongoing work in TWI as well as in other research organisations are focused in this area. An area of difficulty for industrial use of ML systems is their perceived 'black box' nature – in that it is difficult to understand how they operate and the 'reasoning' behind their decisions. Therefore, at both a philosophical and a theoretical level, much remains to be resolved. However, well-established formal approaches for qualifying inspection system for industrial use, in particular within the highly regulated nuclear industry, have begun adopting ML systems and outlining a route for how they can be used safely in the near future (ENIQ, 2021). An important aspect will be traceability between how a machine is developed, configured, trained initially and progressively through its operational life. Capturing 'states' of the system at critical points of its evolution will be key as re-analysis of historical data is an important requirement for critical systems in industry, which regulation systems rely on for oversight.

In this paper a specific algorithm is presented for visualising fatigue cracks using FMC data that is not based on the TFM concept. The benefit of this particular algorithm is that it lives on the edge of where the human mind could also see aspects of the pattern in the data as well as the machine; other parameter spaces can be multi-dimensional and be incomprehensible for the human mind. Patterns are the key to an ML system 'making sense' of the training data. Attempts at training ML systems with TFM images used by human operators have failed or have been of low quality because there are too many exceptions in the training data, which in itself is not of sufficient volume. Strategies to supplement actual data with simulation results is being explored and show promise (Zou et al, 2021). Results from both supervised and unsupervised training will be presented in subsequent publications.

The algorithm for creating the parameter-space in this paper is termed the snooker algorithm. In this approach, each element is considered a transmitter and every element is considered a receiver. However, the ROI does not explicitly exist, so that the outer loop over P in Equation [1] does not exist, which minimises computation times considerably. Instead, from the FMC data matrix, the A-scan from the transmit element (hereafter referred to as the cue) to the receive element (hereafter considered the pocket) is selected. The TOF for propagation from the cue to the pocket which is incident on the back wall (BW) surface (hereafter termed the cushion) is calculated. An electronic gate is set over the expected arrival time of the wave reflected off the cushion, along with an amplitude threshold, and the peak amplitude above this threshold is recorded. The snooker and TFM algorithms are implemented using the Python programming language and make use of standard libraries, namely NumPy and Matplotlib, to process the data and visualise it. Python was selected for building the algorithms as it is an elegant high-level language where prototyping can be done efficiently. Furthermore, it is the language of choice for building ML systems using established open source frameworks, namely TensorFlow. For improved processing speeds the C programming language can be used to code the algorithms and used through Python as extensions, thereby allowing development of versatile and performant prototypes.

The steps for the snooker algorithm are outlined below.

1. Select the first element of the array as the cue. For each element in the array, including the transmitting cue, set it as the pocket.
2. Calculate the TOF from the cue to the pocket by reflection off the cushion. This calculation is based on the geometry, where the travel paths are straight lines when the material is isotropic, i.e. when the velocity of the sound wave remains constant in all directions.
3. In the FMC data matrix, using the A-scan for the cue-pocket (transmit-receive) pair, set a 'gate' over the expected arrival time of the wave. A gate is a concept where a signal within a defined period is captured or recorded for additional processing.
4. Using the gated signal, identify the maximum amplitude within the gate above a defined threshold. Selection of this threshold will depend on the noise level in the signal, but typically for the carbon steel material used for illustration in this paper the noise levels are low. Hence the threshold was set to 10% of the maximum possible signal level, so that all signal levels above this threshold will be considered valid for processing.
5. Record the maximum amplitude level of the signal in the gate, above the threshold, and store it in an $n \times n$ 2D matrix for each cue-pocket pair.
6. Repeat the process over all cue-pocket pairs. Once done, visualise the $n \times n$ 2D matrix, which is the parameter-space image for training the ML system.

This basic algorithm can be extended into more sophisticated cases, for instance building transmit-receive paths between the cues and pockets that use a postulated crack as a mirror reflector. Since the cue in theory has a wide angle of transmission (as it is nominally considered a point source with a divergent beam) it is possible to isolate particular paths for investigation. The only assumption in the snooker algorithm is that the law of reflection on a surface is assumed – i.e. the angle of incidence of a beam on a surface is equal to the angle at which the beam reflects from that surface (the BW and the crack). The assumption of a fatigue crack as a smooth mirror-like reflector is consistent with reality. A key difference between the TFM and snooker algorithms is that the amount of calculation required for the former to compute the image grows in proportion to the resolution of the ROI, whereas to calculate the parameter-space image only n^2 operations are required. The usefulness of the algorithm, in the version presented in this paper, lies in the hypothesis that when fatigue cracks exists, not all cue-pocket paths yield an amplitude – i.e. fatigue cracks stop 'balls' from being 'potted' in the 'pocket'.

Results and Discussion

The parameter-space images are all consistent in terms of profile – i.e. they are always an $n \times n$ matrix regardless of the size of the component and the position of the BW with respect to the array. This is an important factor in order to maintain a standard interface to train the ML system and to be able to train the system using data from components of different sizes. The only information contained in the parameter-space is a consequence of the presence, or not, of fatigue cracks; however, when the ML system is trained a variety of cases are being utilised, such as material with different types of reflectors other than fatigue cracks, and material with some degree of anisotropy, in order to study the impact of these other factors, which are considered parameters influencing performance of the ML systems.

Figure 3 shows the TFM image of a notch where the BW and the tip of the notch have been identified. The data was collected using a 64 element probe, with a pitch of 0.35mm and a frequency of 10MHz. The blocks with targets were all carbon steel, where the nominal sound velocity was taken as 5900m/s.

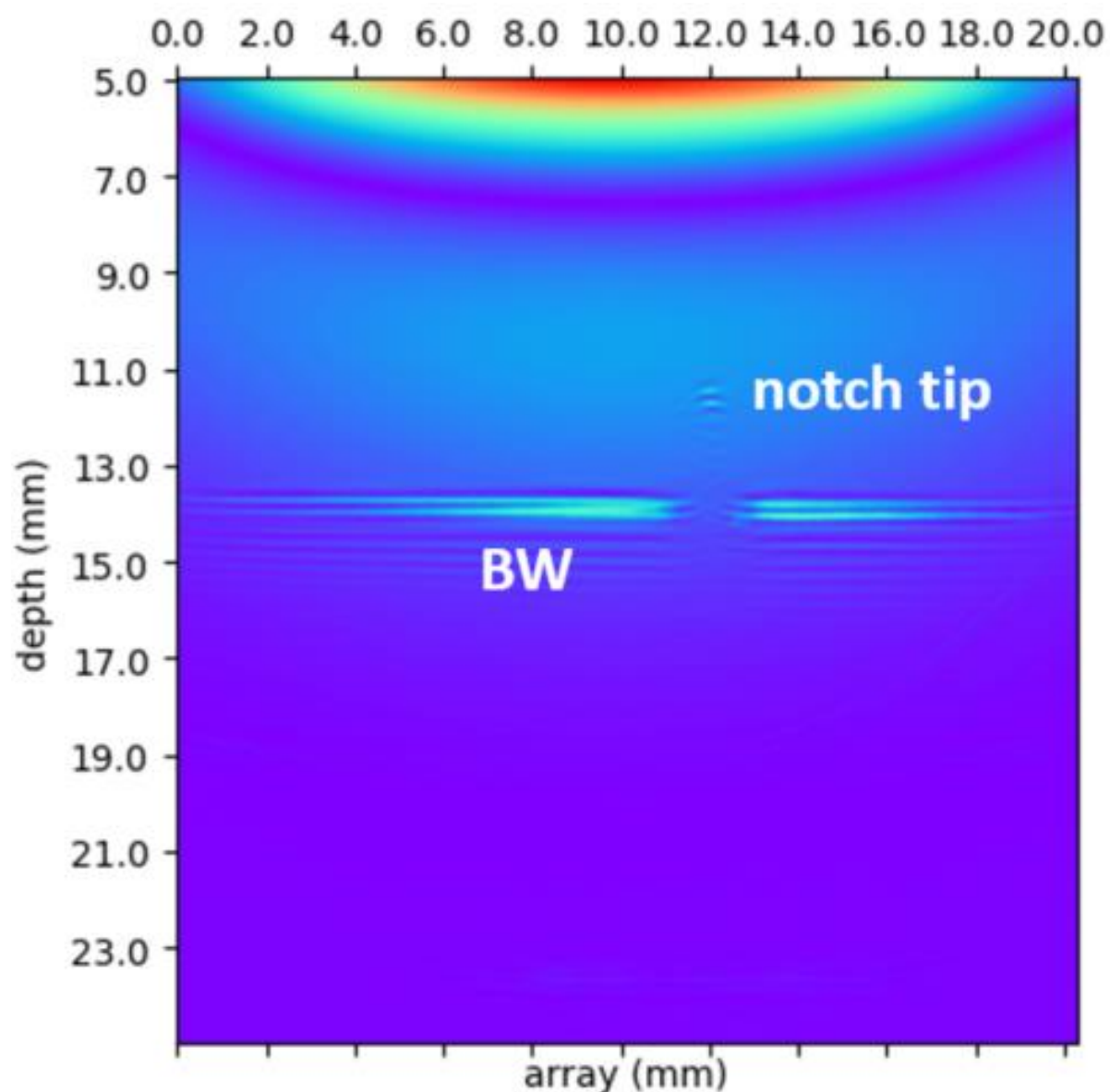


Figure 3 A TFM image of a block containing a machined notch, where the BW position and the notch tip have been identified. Refer to Figure 1 for the experimental setup.

Figure 4 shows the parameter-space image for the notch presented in Figure 3. A bicubic interpolation scheme was used to process the image and a rainbow colour map was used to represent the signals, which were all normalised to the maximum in the parameter-space image. The image can be interpreted thus: taking the x-axis as the cue and the y-axis as the pockets, red implies that a very high BW signal was received, whereas blue implies little or no signal was received. The overall 'pattern' is representative of the particular notch in this block. To compare, a notch that was much larger than that shown in Figure 3 was used, and Figure 5 shows the TFM image and the corresponding parameter-space image. It is clear when comparing the parameter-space image of the small notch (Figure 4) with that for the larger notch (Figure 5) that more of the transmit-receive paths were blocked by the notch. Qualitatively the data is consistent but it is not evident how a human operator could process the parameter-space images to establish any sizing information about the flaw – detection is clear but measurement of the through-wall extent of the notch is easier using the TFM images. Therefore the parameter-space images are not suited for use by human operators for the purpose of sizing any flaws.

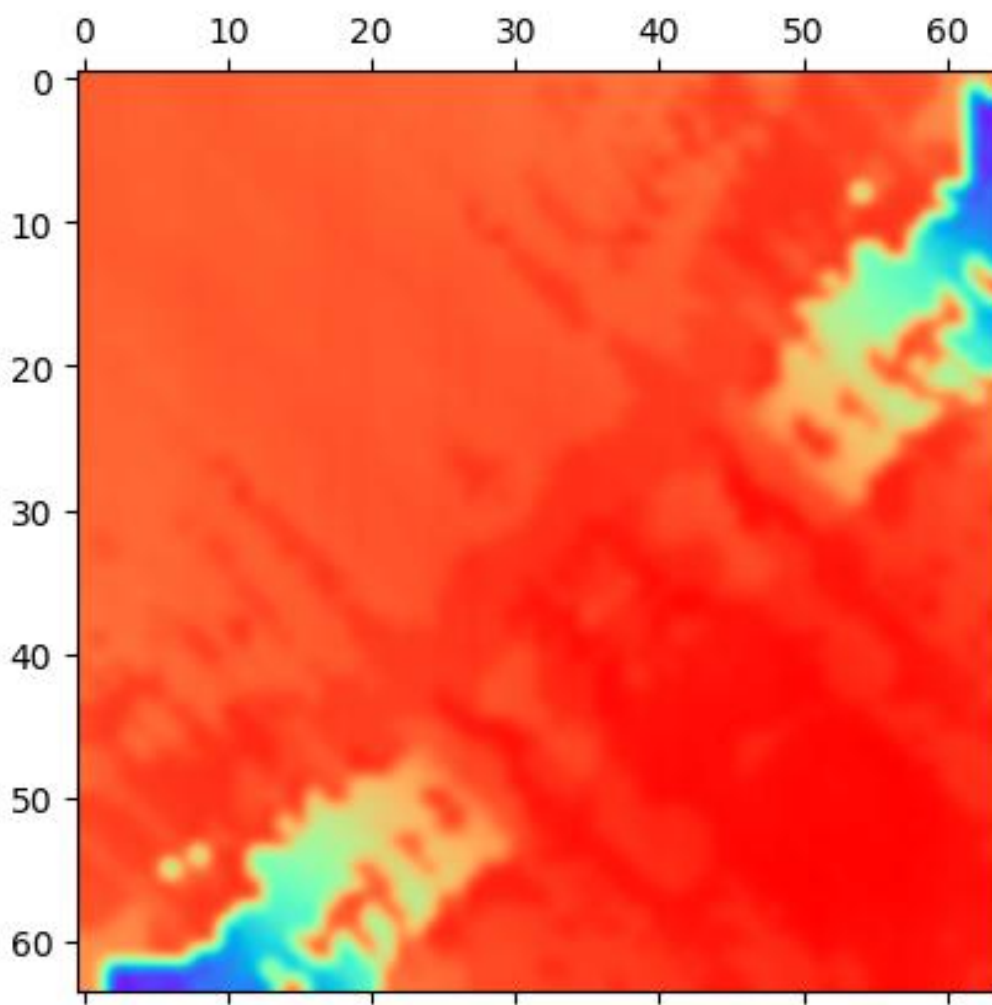


Figure 4 The parameter space image of a block containing a machined notch, created using the same FMC data used to create the TFM image in Figure 3. The scales are the number of elements, i.e. $n=64$.

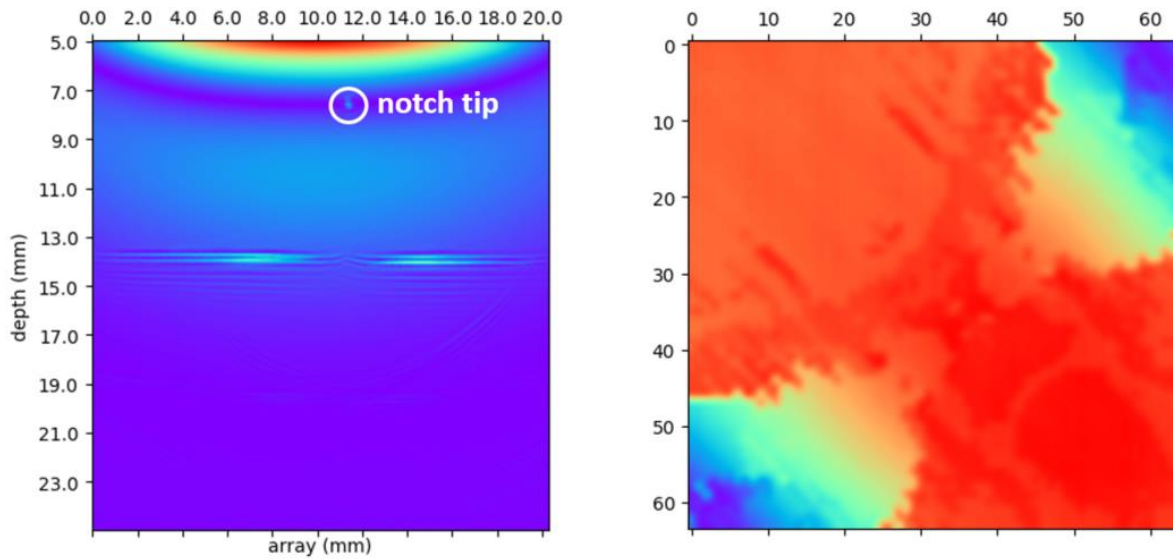


Figure 5 TFM and corresponding parameter space images from the same FMC data matrix for a large notch, where the tip of the notch approaching the surface is highlighted in the TFM image.

Figure 6 shows the TFM image of a volumetric SDH target in a block along with its parameter-space image. Note that the BW distance is much larger than for the blocks with notches but this information is not captured in the parameter-space image and does not affect it. On the other hand, the parameter-space image captures the information that there is no crack between the SDH and the BW surface, because many transmit-receive paths yield high amplitude signals. In this test scenario we know that the signal is from a SDH but in real inspection campaigns it may not be possible to be sure, and this is when the parameter-space view may help to make a more holistic assessment of the data.

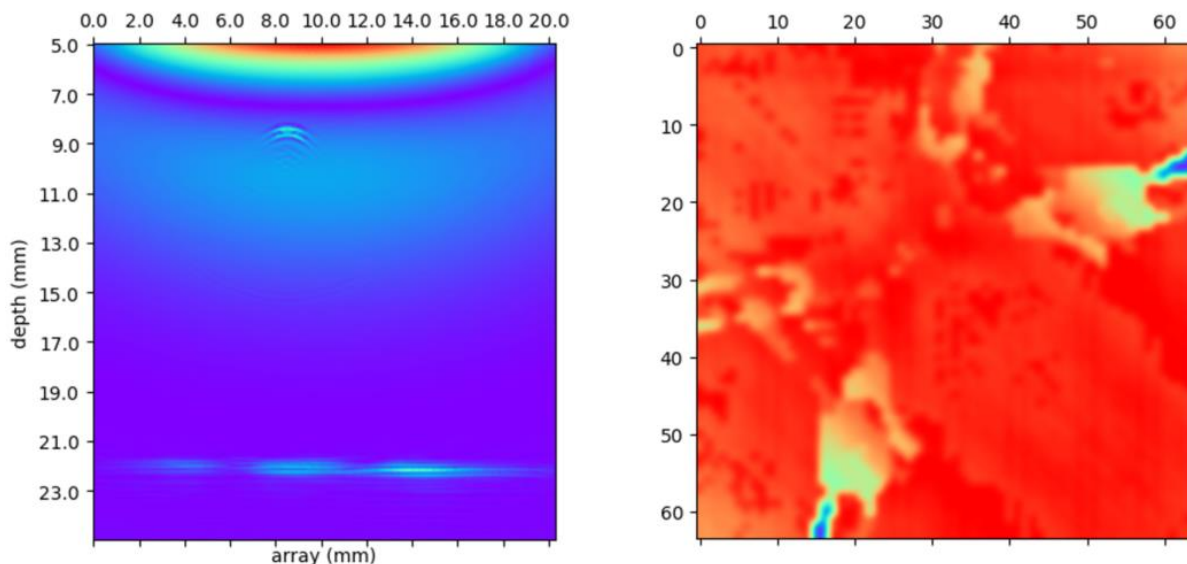


Figure 6 TFM and parameter-space images for a SDH, where there is no cracking.

Figure 7 shows the data from an actual fatigue crack in one of the blocks presented in Figure 2. Here the signals from the flaw are more difficult to ascertain using the TFM image: a faint rumbling is evident but there is no clear evidence of a fatigue crack tip. There is also a discrepancy in the second BW image which due to anisotropy in the steel becoming more evident over longer travel paths. This information is difficult to establish in the parameter-space image, but it does quite clearly indicate the

presence of a planar flaw blocking the shots. In addition there is evidence in the parameter-space image as to the through-wall extent of the crack, and that the profile of the crack is not uniform as when compare to the accurately machined notches of Figures 4 and 5. A further example of a fatigue crack is presented in Figure 8 and the asymmetry in the amplitude profile is clearer – note how the top right lobe of the image is subtly different to the bottom left. Normally, in array based imaging the rule of reciprocity is observed, which implies that the signal sent from the transmitter to the receiver should be identical to a signal sent from the receiver to the transmitter, such that a matrix such as the parameter-space should be symmetric about its main diagonal where the transmitter is the receiver. This level of asymmetry may be due to the use of a gate and a threshold to collect the amplitude data for the parameter-space and the impact of this variability – i.e. ‘noise’ – in the information used to train the ML system is being assessed in the ongoing work.

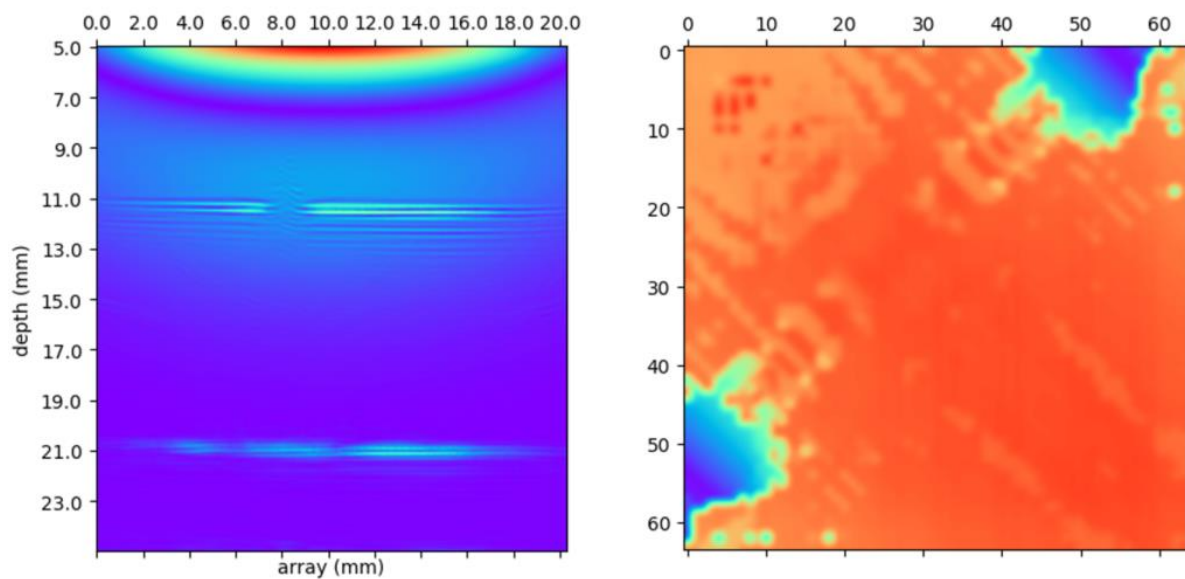


Figure 7 TFM and parameter-space images for a fatigue crack. Note that the echo at a depth position below the main BW echo in the TFM image is termed the second BW echo.

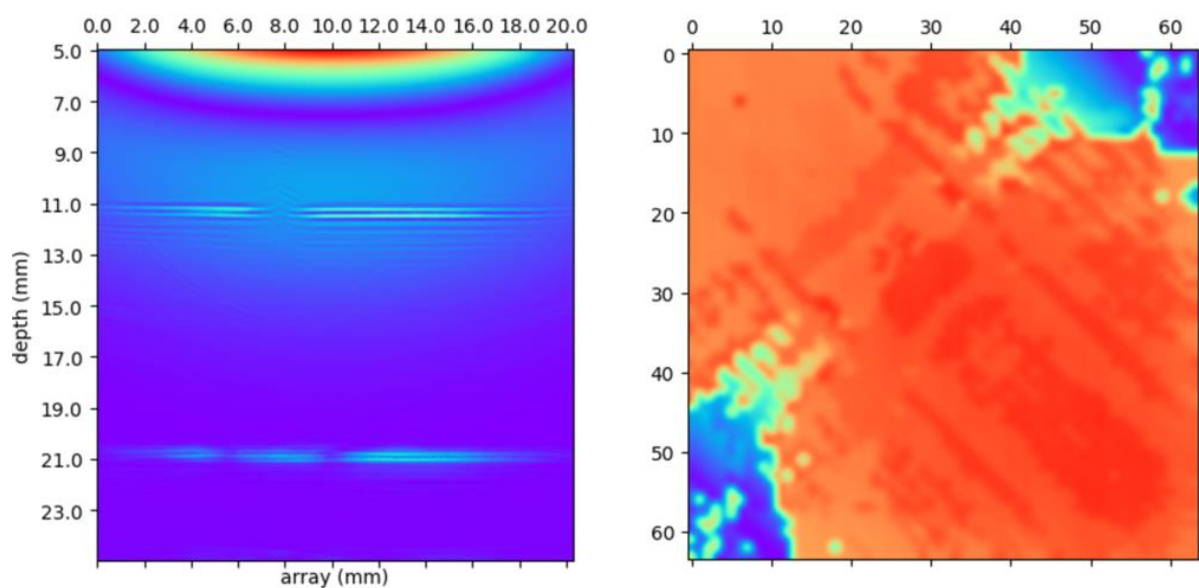


Figure 8 TFM and parameter-space images for a fatigue crack.

Figures 9 and 10 show further examples of TFM and parameter-space images for actual fatigue cracks presented in Figure 2. As part of the ongoing work these blocks will be sectioned to visualise the actual crack profiles in order to label the input data for supervised ML approaches. It is important to note that the size and profile of the cracking observed on the edge of the specimen is not necessarily the same one observed in the ultrasonic data, as the data is collected from within unobserved material. Therefore, in order to accurately train the ML system, the blocks will be sectioned and photographed. The aim of the wider programme of work is to investigate the use of parameter-space images – and other strategies – for training ML systems, and crucially to establish if it is possible to effectively train the system with a reduced amount of training datasets. In order for ML systems, in their current form, to be useable in the inspection industry the onerous training requirements must be reduced. In addition, the inspection industry requires stringent validation efforts, as inadequate or compromised systems can lead to catastrophic consequences, with loss of human life and other adverse outcomes.

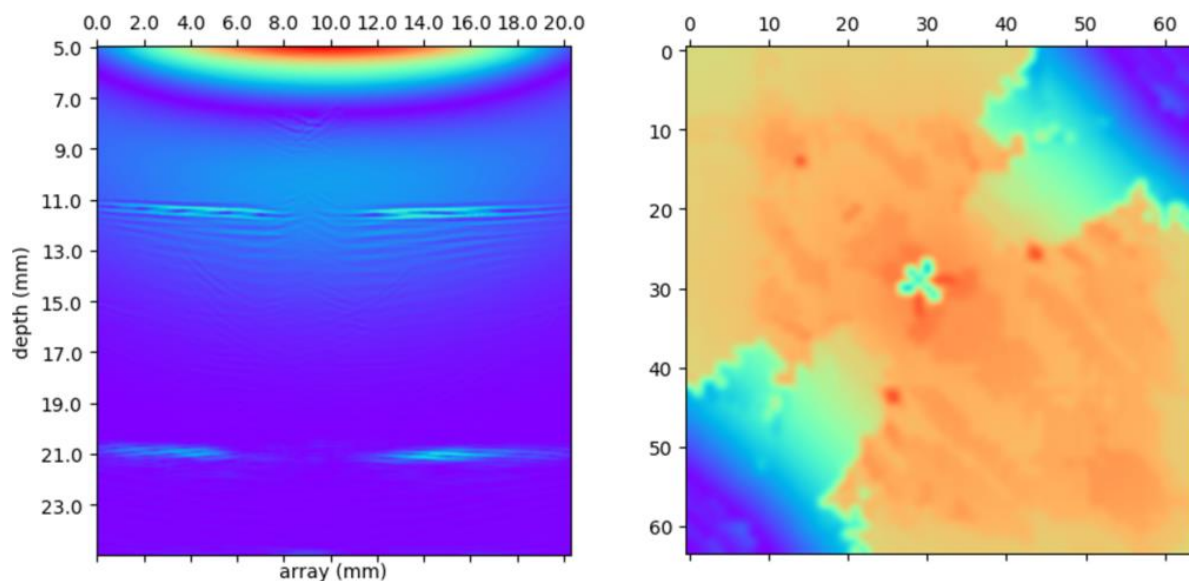


Figure 9 TFM and parameter-space images for a fatigue crack.

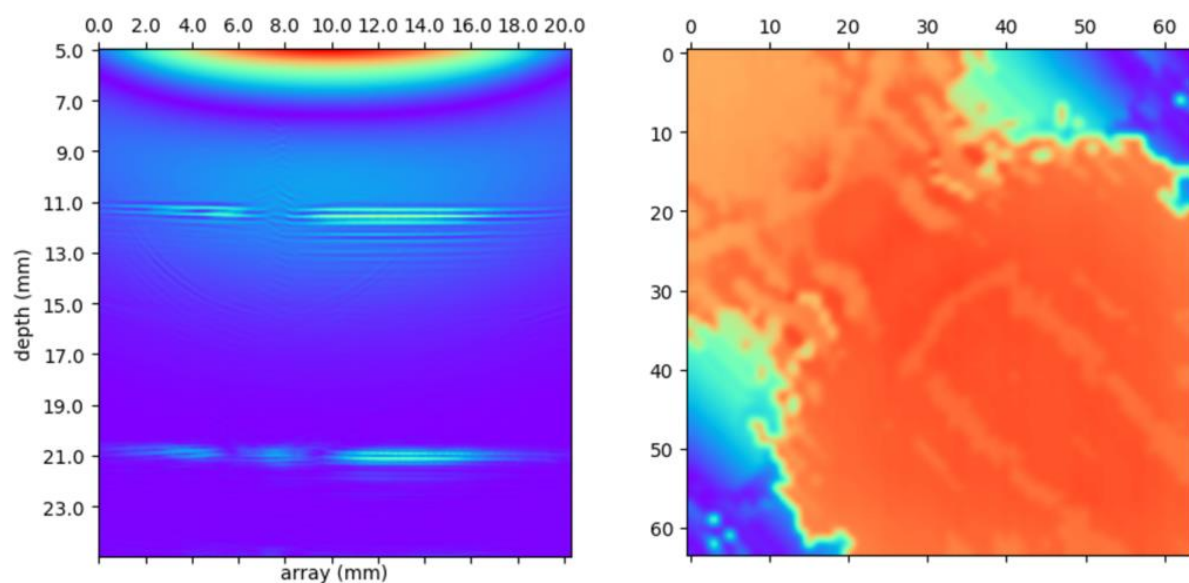


Figure 10 TFM and parameter-space images for a fatigue crack.

Conclusions

1. A new approach to presenting ultrasonic data for ML systems, termed parameter-space images, has been developed;
2. An example of a parameter-space image using an algorithm called ‘snooker’ which utilises the FMC data matrix but with reduced computational requirements in comparison to the TFM algorithm was developed;
3. Together, the snooker algorithm and the parameter-space images, can condition FMC data for input to ML systems for training, as they appear to be insensitive to extraneous information, instead focused on the fatigue cracking morphology;
4. The ongoing work, which will be reported in a subsequent publication, is aiming to establish whether a ML system can be effectively trained using smaller data sets when the training is done using parameter-space images;
5. In future, a methodology to interface real-world ultrasonic data to the ML framework will be developed as it will be crucial to realise the full benefits of ML based inspection systems.

References

1. Nageswaran C. Maintaining the integrity of process plant susceptible to high temperature hydrogen attack. Part 1: analysis of non-destructive testing techniques. *Health and Safety Executive Research Report RR1133*, <https://www.hse.gov.uk/research/rrpdf/rr1133.pdf> (2018, accessed 11 July 2021).
2. Nageswaran C, Day A and Hayward K. Monitoring Fatigue Crack Growth in Subsea Threaded Components using Ultrasonic Phased Array Techniques. *TWI-Global Web Portal*, <https://www.twi-global.com/technical-knowledge/published-papers/fatigue-crack-growth-ultrasonic-phased-array-techniques> (2013a, accessed 11 July 2021).
3. ISO 23865. Non-destructive testing – Ultrasonic testing – General use of full matrix capture/total focusing technique (FMC/TFM) and related technologies. *International Organisation for Standardization* (2021).
4. Nageswaran C. Measuring the Crack Growth Rate (da/dt) of a Fatigue Crack Using Phased Array Ultrasonics. *TWI-Global Web Portal*, [Measuring a Fatigue Crack Growth Rate Using Phased Array - TWI \(twi-global.com\)](https://www.twi-global.com/technical-knowledge/published-papers/measuring-a-fatigue-crack-growth-rate-using-phased-array-ultrasonics) (2013b, accessed 11 July 2021).
5. Google. Machine Learning Crash Course. [Machine Learning Crash Course | Google Developers](https://www.coursera.org/learn/machine-learning) (2021, accessed 11 July 2021).
6. ENIQ. Recommended Practice 13 – Qualification of Non-Destructive Testing Systems that Make Use of Machine Learning. *European Network for Inspection and Qualification* (2021).
7. Zou X, Nageswaran C, Tse YY and Jiang JJ. Convolutional neural network for classification of stress corrosion cracking using simulation-based training. *Submitted for publication* (2021).

Supplementary Material

Discussion of fatigue cracks, ultrasonic imaging and the background to this paper was presented by the author at a Technical Group meeting of The Welding Institute given on 6 July 2021, a video of which can be found here: <https://www.youtube.com/watch?v=7Jx0liwm5FY>

In addition to the talk by the author, other experts presented related information as background, including the use of surface inspection methods for fatigue cracks and a general discussion at the end on the use of ML in the inspection industry which is projected to grow significantly in the near future.