NDT and E International 93 (2018) 86–97

Contents lists available at ScienceDirect

NDT and E International

journal homepage: www.elsevier.com/locate/ndteint

A confidence map based damage assessment approach using pulsed thermographic inspection

Yifan Zhao *, Sri Addepalli, Adisorn Sirikham, Rajkumar Roy

Through-life Engineering Services Centre, Cranfield University, Cranfield, MK43 0AL, UK

ARTICLE INFO

Keywords:
NDT
Reliability
Thermography
Impact damage assessment
Composite

ABSTRACT

In the context of non-destructive testing, quantification of uncertainty caused by various factors such as inspection technique, testing environment and the operator is important and challenge. This paper introduces a concept of contour-based confidence map and an application framework for pulsed thermography that offers enhanced flexibility and reliability of inspection. This approach has been successfully applied to detect three flat-bottom holes of diameter 32, 16 and 8 mm drilled onto a 5 mm thick aluminium plate with a high accuracy of damage detection ($R > 0.97$). Its suitability and effectiveness in assessing impact damage occurring in composites have also been demonstrated.

1. Introduction

Non-destructive testing (NDT) has been the front-runner in estimating the health of a component over the last few decades with specific emphasis on damage detection and quantification without causing further damage to the material. Pulsed thermography inspection has now been established as a reliable thermal NDT technique to detect near and sub-surface damage occurring in various materials. Pulsed thermography offers an effective alternative where damage detection and quantification is much faster and robust in comparison with traditional NDT methods such as ultrasonic testing and 3D X-radiography computed tomography methods [1,2]. The users of pulsed thermography are frequented with questions such as 'How do you estimate the accuracy of defect measurement? Or what is your confidence level of damage characterisation and how the confidence level affects the decision making?'. There is very limited reported research addressing these issues directly. Understanding the uncertainty of defect/damage characterisation is important because that is the only way to mitigate the uncertainty associated with the inspection and improve the accuracy of the measurement through identifying the source of errors followed by corresponding actions. Thermal data acquisition is a challenging process where the technique's dependence is heavily based on primarily the infrared detection system followed by an appropriate heat excitation source. Most of the current state-of-the-art systems still employ equipment such as flash lamps. These optical units are heavily dependent on capacitor bank systems where there is a level of uncertainty that exists in determining the flash initiation and end of flash and can only be monitored by a high frame rate infrared acquisition system. Further, the influence of environmental parameters such as the background temperature and humidity levels together with the inspected material, its type, surface finish and the data synchronisation all add uncertainty to the acquired measurement data, which adds disparity between inspection rendering repeatability as a challenging aspect [3]. Therefore, there is a strong demand to build the confidence level in results obtained from the thermographic inspection which becomes a driving factor to help establish and exploit the active thermal inspection method in the main stream inspection scenarios.

The use of Probability of Detection (POD) curves to quantify NDT reliability is common in the aeronautical industry [4]. There are studies that have been conducted to determine the POD for anomalies occurring in composite materials where traditional NDT techniques such as ultrasonic testing, radiography, and eddy current have been used [5–7]. Minkina and Dudzik [3] and Lane et al. [8] investigated the errors and uncertainties in infrared thermography in the passive mode, where it is mentioned that errors of temperature measurement with the infrared camera are typically classified into errors of the method, errors of calibration, and errors of the electronic path. However, manufacturing test pieces with representative flaws in sufficient numbers to draw statistical conclusions on the reliability of the NDT system being investigated is costly. The application of active thermography in detecting damage of metallic components and composites has been well established over the last few years but associated reliability research is limited. A few POD studies have been conducted to improve the applicability of pulsed and

* Corresponding author.
E-mail address: yifan.zhao@cranfield.ac.uk (Y. Zhao).

https://doi.org/10.1016/j.ndteint.2017.10.001
Received 12 May 2017; Received in revised form 29 September 2017; Accepted 2 October 2017
Available online 7 October 2017
0963-8695/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

lock-in into main stream inspection activities [9–11]. However, with the uncertainties and errors associated with the inspection process itself, it is important that a relative measure of confidence with limited trials needs to be addressed and established to make sure that pulsed thermography can be used to discriminate the health of the part being inspected. This paper is an effort to answer the very challenge described and thus develops a method to compute a unique confidence map that quantifies the confidence level of damage detection for each pixel statistically, and then introduces a confidence-map-based assessment routine to further exploit the applicability of pulsed thermography inspection technique to perform material degradation assessment.

2. Methods

In order to develop a contour-based confidence map and the associated assessment toolsets, this paper starts from the improvement of existing theory of defect characterisation, which is then integrated with the statistical theory to construct a new concept of representative of defect. This section presents the various concepts that support the development of the Adaptive Peak Temperature Contrast method (APTC), together with a proposed inspection framework that will enhance and highlight the merits of the pulsed active thermography system.

2.1. Defect detection using pulsed thermography

In pulsed thermographic inspection, the typical experimental setup of which is illustrated in Fig. 1(a), a short and high energy light pulse from the flash lamps is projected onto the sample surface. Heat conduction then takes place from the heated surface to the interior of the sample, leading to a continuous decrease of the surface temperature [12]. An infrared radiometer controlled by a PC captures the time-dependent response of the sample surface temperature. In areas of the sample surface above a defect (see point 2 in Fig. 1) the transient flow of heat from the surface into the sample bulk is wholly or partially obstructed, thus causing a temperature deviation from the sound areas (see point 1 in Fig. 1). Most of the defect detection methods are based on the classification of the temperature decay curve (see Fig. 1(b)). The time when the temperature deviation occurs can be used to estimate the defect depth. The surface temperature due to a defect at depth L for a plate is given by Ref. [13].

$$T(t) = \frac{Q}{\sqrt{\pi\rho ckt}}\left[1 + 2\sum_{n=1}^{\infty}\exp\left(-\frac{n^2L^2}{\alpha t}\right)\right] \quad (1)$$

where $T(t)$ is the temperature variation of the surface at time t, Q (unit: J) the pulse energy, ρ (unit: kg/m^3) the material density, c (unit: J/kgK) the heat capacity, k (unit: W/mK) the thermal conductivity of the material and α (unit: m^2/s) it's the thermal diffusivity.

The most widely used method to differentiate sound areas and defective areas is using the thermal/temperature contrast technique. Various temperature contrast definitions exist [14] but they share the need for specifying a sound area A_s as the reference. For instance, the absolute temperature contrast $\Delta T(t)$ is defined as

$$\Delta T(x, y, t) = T(x, y, t) - T_{A_s}(t) \quad (2)$$

where $T(x, y, t)$ denotes the temperature of a pixel at the location (x, y) at time t, and $T_{A_s}(t)$ denotes the temperature at time t for the pre-defined sound area A_s. Practically, the definition of A_s is important as issues such as non-uniform heat application and surface finish can cause considerable variations on the results and the same can be observed when changing the location of A_s [15]. A frame of absolute temperature contrast at a certain time is usually selected to represent the result of defect detection. For example, the Peak Temperature Contrast method (PTC) [14] calculated the thermal contrast between the defective/damaged region and an adjacent sound or non-defective region, and the frame where the maximum contrast between the sound and defective areas is chosen even though the defect peak occurs much later in time. Because of the 3D heat conduction effect, the temperature contrast first increases with time and then decreases. The time at which the temperature difference rises to its maximum value is approximately proportional to the square of the defect depth, and the proportionality coefficient depends on the size of the defect. Therefore, it should be noted that PTC is a defect detection method and only provides an approximation for defect depth measurement.

One limitation of PTC is that when the selected region of interest (ROI) includes multiple defects with a variety of sizes and depths, the selection of the optimal frame to visualise all defects in a single image is a challenge. The most common approach in defect characterisation in such cases is by considering defects occurring at similar depths and truncating the sampling time accordingly to achieve best contrast, and the same repeated for the remaining defects [16]. However, automating such a dynamic assessment approach to detect and quantify all defects at the same time is challenging. This paper proposes an Adaptive Peak Temperature Contrast (APTC) method to detect defects before estimating the confidence map. For each pixel on the image plane, the peak of temperature contrast is computed and a map of these peaks is constructed to represent the detection result, by which means, defects with different sizes or depths can be visualised with maximal contrast in a single image. To reduce the noise, the Thermographic Signal Reconstruction (TSR) algorithm [17] is employed to fit the raw data before the application of APTC. The estimation of APTC can then be written as

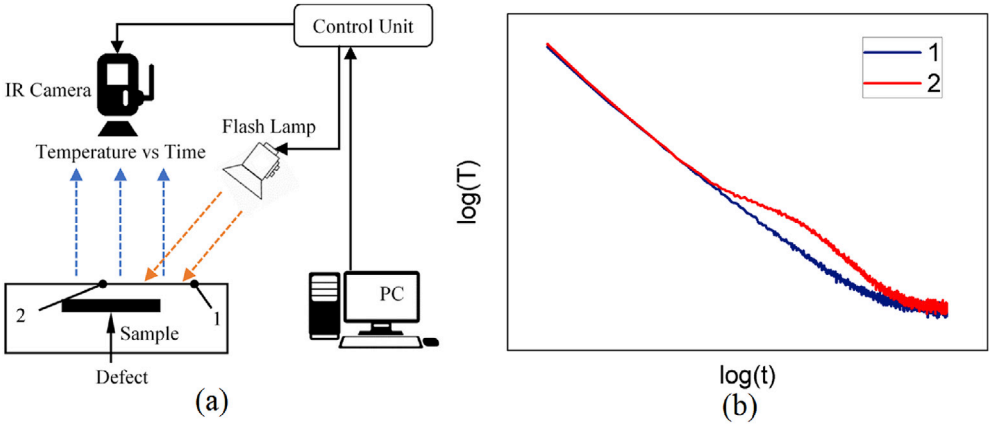

Fig. 1. (a) Experimental configuration of the pulsed thermographic inspection, where point 1 denotes a sound area on the sample surface and point 2 denotes a position with defects underneath; (b) Typical observed time-temperature decay curves in the logarithmic domain for the point 1 and 2, respectively.

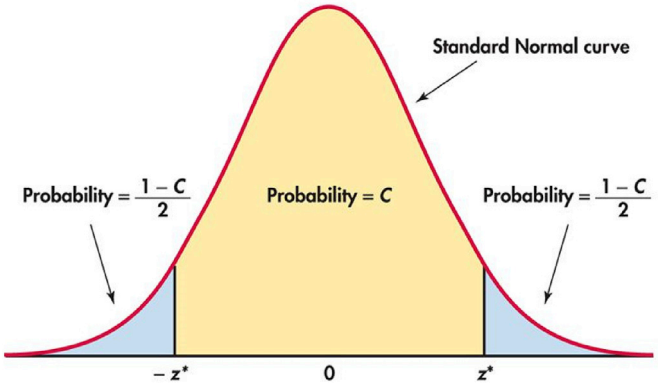

Fig. 2. Illustration of the confidence interval.

Table 1
The correspondence between the confidence level and the value of z^*.

Confidence Level	Value of z^*
50%	0.674
60%	0.842
70%	1.036
80%	1.282
90%	1.645
95%	1.960
98%	2.326
99%	2.576
99.8%	3.090
99.9%	3.291

$$APTC(x,y) = \max_t\left(\tilde{T}(x,y,t) - \tilde{T}_{A_s}(t)\right) \tag{3}$$

where $\tilde{T}$ denotes the TSR fitting of the raw temperature, which can be expressed as

$$\tilde{T}(x,y,t) = \exp\left(\sum_{i=0}^{N} a_i(\ln(t))^i\right) \tag{4}$$

where N is the model order and a_i the fitted coefficient of the data collected from the position (x,y). $\tilde{T}_{A_s}(t)$ denotes the averaged TSR fitting for the sound area. The selection of the model order is discussed by Zhao et al. [1], [18]. In all examples of this paper, the model order was chosen as 7 [18]. It has been reported that the first and second derivatives of TSR show improvements in detecting the defects [19]. The proposed APTC can be extended to an Adaptive Peak Temperature Contrast of the First Derivative (APTC1D) and Adaptive Peak Temperature Contrast of the Second Derivative (APTC2D) and is expressed as

$$APTC1D(x,y) = \max_t\left(\frac{d\left(\tilde{T}(x,y,t)\right)}{dt} - \frac{d\left(\tilde{T}_{A_s}(t)\right)}{dt}\right) \tag{5}$$

and

$$APTC2D(x,y) = \max_t\left(\frac{d^2\left(\tilde{T}(x,y,t)\right)}{dt} - \frac{d^2\left(\tilde{T}_{A_s}(t)\right)}{dt}\right) \tag{6}$$

respectively. It should be noted that the first and second derivative are computed using the fitted coefficients a_i to achieve better resolution [17]. It should be noted that APTC1D and APTC2D are based on the derivative of temperature contrast, so they are more sensitive for defects, as well as the associated noise in the data. Generally, if the captured data has high Signal-to-Noise-Ratio (SNR) similar to the ones acquired from composites, the APTC1D and APTC2D are recommended. Whereas metallic components tend to produce a low SNR as evidenced by data from aluminium or steel, the APTC is recommended.

Another reason why APTC is preferred over PTC in this paper is that we aim to evaluate the confidence level for different defects in a single map, therefore the way to produce the property of each pixel must follow a uniform rule. If multiple defects with different sizes or depths are considered, any selected frame using PTC will produce biased results whereas the APTC would produce a direct comparison based results.

2.2. Confidence map

While a normally-distributed (or Gaussian-distributed) random variable can have many potential outcomes, the shape of its distribution gives the confidence that the majority of these outcomes will fall relatively close to its mean. By assuming that this distribution is known or can be estimated, the distance between a new observed value and the mean can be used to quantify the confidence that this individual follows this distribution. Similarly, this distance can also be used to quantify the confidence that this individual does not follow this distribution. Let X be a random sample from a probability distribution with a statistical parameters θ, which is the quantity to be estimated, and φ, representing quantities that are not of immediate interest [17]. In statistical theory, a confidence interval for the parameter θ, with a confidence level C, is an interval with random endpoints $(u(X),\ v(X))$, determined by the pair of random variables $u(X)$ and $v(X)$, with the property

$$\Pr_{\theta,\varphi}(u(X) < \theta < v(X)) = C \text{ for all } (\theta,\varphi). \tag{7}$$

The number C, with typical values close to but not greater than 1, usually is given in the form of a percentage. As shown in Fig. 2, the z^* value measures the number of standard errors to be added and subtracted to achieve the desired confidence level (the percentage confidence you want). Table 1 shows a list of common confidence levels and their corresponding z^* values [20], which will be considered in this paper.

To relate the above theory with the studied application, we assume that $p(x,y,\ i)$ be the estimated property of the considered pixel at the position (x,y) in the i^{th} trial. It should be noted that the property p is not limited to APTC, APTC1D or APTC2D introduced above, and can be used as an unbiased feature that can differentiate pixels from sound and defective areas. It has been verified by Zhao et al. [18] that the thermal property (e.g. thermal diffusivity) of sound areas approximately follows the Gaussian distribution. To reduce the influence of experimental noise, a multi-trial process is proposed in this paper. In the classification process, there are three possible modes: supervised, semi-supervised and unsupervised. The supervised mode is defined when the ground truth of both defective areas and sound areas is pre-known, which applies to most of the numerical simulations and some experimental simulations. The semi-supervised mode is defined when only a limited number of pixels from sound areas are known, which applies to most of the experimental tests. The unsupervised model, also called 'blind test', is a task of organising data from 'unlabelled' data, which in this case is the most challenging scenario. In this paper, only the semi-supervised and unsupervised modes are considered because there is a lack of established inspection techniques that could provide the ground truth for real damage/defects occurring in the real-world, such as impact damage in composite materials. For the semi-supervised mode, the mean and standard deviation of the property, denoted by $\overline{\mu_p}(i)$ and $\overline{\sigma_p}(i)$ respectively, are estimated by randomly sampling the pixels from the defined sound areas with the number of sampled pixels, N, while for the unsupervised mode, the whole image is randomly sampled. The selection of N will be discussed through empirical tests in the next section. For each considered pixel of each trial, the z^* value can be estimated by,

To reduce the uncertainty caused by random sampling, the process of sampling and calculation of z^* is repeated for Q times and the z^* values for each trial are fused using the 'OR' or 'AND' operator. Therefore, Eq.

(8) can then be rewritten as

$$z^*(x,y,i) = \frac{|p(x,y,i) - \overline{\mu_p}(i)|}{\overline{\sigma_p}(i)} \tag{8}$$

$$z^*(x,y,i) = \bigcup_{j=1}^{Q} \frac{|p(x,y,i,j) - \overline{\mu_p}(i,j)|}{\overline{\sigma_p}(i,j)} \tag{9}$$

or

$$z^*(x,y,i) = \bigcap_{j=1}^{Q} \frac{|p(x,y,i,j) - \overline{\mu_p}(i,j)|}{\overline{\sigma_p}(i,j)} \tag{10}$$

Empirical tests show that the 'OR' operator usually produces a better result than 'AND', the details of which will be discussed in the next section.

Assuming the total number of trials is M, the averaged value of z^* can be calculated by

$$\overline{z^*}(x,y) = \frac{1}{M}\sum_{i=1}^{M} z^*(x,y,i) \tag{11}$$

The confidence level that the considered pixel is not from the sound areas, $\overline{C}(x,y)$, can be estimated by searching Table 1 using the averaged estimation of z^*. These steps are repeated for each pixel, and a confidence map can then be established based on multiple trials. It should be noted that this paper proposes the use of contour mapping to represent the confidence levels. The main reason is that the contour map is region-based rather than pixel-based, and is hence more suitable to represent defects or damage areas.

2.3. Inspection framework

To better utilise the proposed confidence map, this paper introduces a novel inspection framework based on pulsed thermography.

As illustrated by Fig. 3, this framework includes:

a. Data collection: Collect raw data based on pulsed thermography and select ROI. It should be noted that the selected ROIs must be same for different trials to avoid an extra task of image registration.
b. TSR fitting: Fit the raw digital intensity data using a polynomial model in the logarithmic domain to remove noise and improve temporal resolution.
c. Damage detection: Calculate the property of digital intensity decay curve of each pixel to distinguish the pixels from sound and defective areas using Eq. (3), Eq. (5) or Eq. (6).
d. Calculation of z^* value: For the semi-supervised mode, randomly select N pixels from the pre-defined sound areas while for the unsupervised mode, randomly select N pixels from the whole ROI. The mean and standard deviation of the sampled pixels are then computed. Now calculate the z^* value for each pixel based on Eq. (8). This step is repeated for Q times and the results are fused by Eq. (9) or (10).
e. Confidence map estimation: Repeat the step (a) to (d) for M times and compute the averaged z* value for each pixel. Then search Table 1 to construct the confidence map. The map is visualised by computing its contour.
f. Binarisation: The confidence map can be converted into a binary map by introducing a confidence threshold, for example, 95%.
g. Depth measurement: Base on the binary map, the depth of defects/damage can then be estimated using some reference-based methods, such as Peak Slope Time (PST) [21–23], or reference-free-based methods, such Log Second Derivative (LSD) [17], Absolute Peak Slope Time (APST) [24], Discrete Fourier Transform (DFT) method [25], Nonlinear System Identification (NSI) method [18] or Least-Square Fitting (LSF) method [12].

3. Results and discussion

This section presents the application of the proposed confidence map and corresponding inspection routine on two representative material samples; the inspection of flat-bottom holes in an aluminium plate, where the ground truth is known, and six carbon fibre reinforced polymer (CFRP) samples with different levels of impact damage.

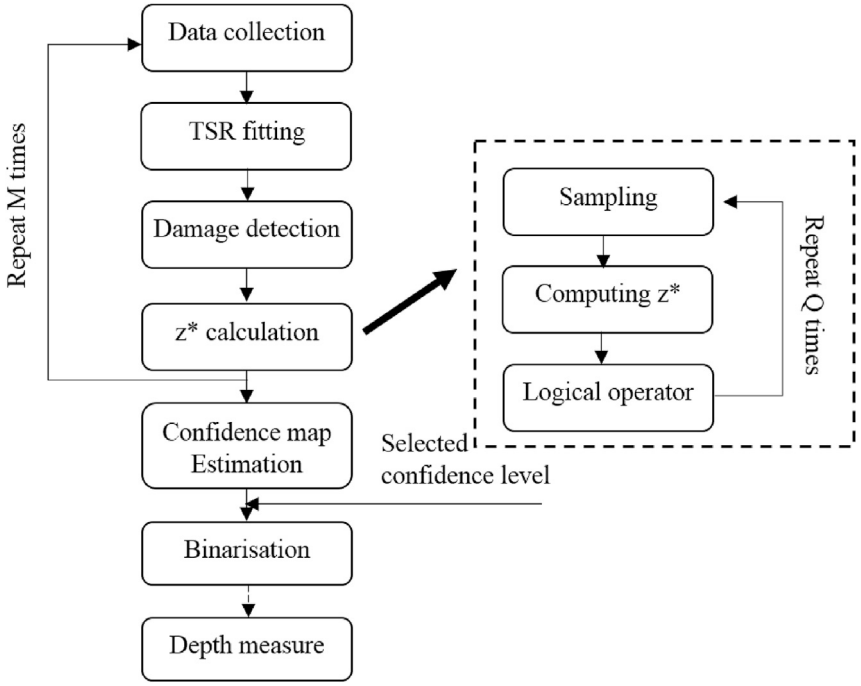

Fig. 3. Proposed framework to characterise defects based on pulsed thermography.

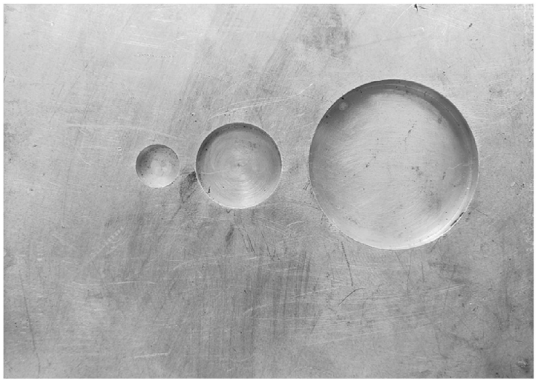

Fig. 4. The drilled surface of the flat-bottom hole samples.

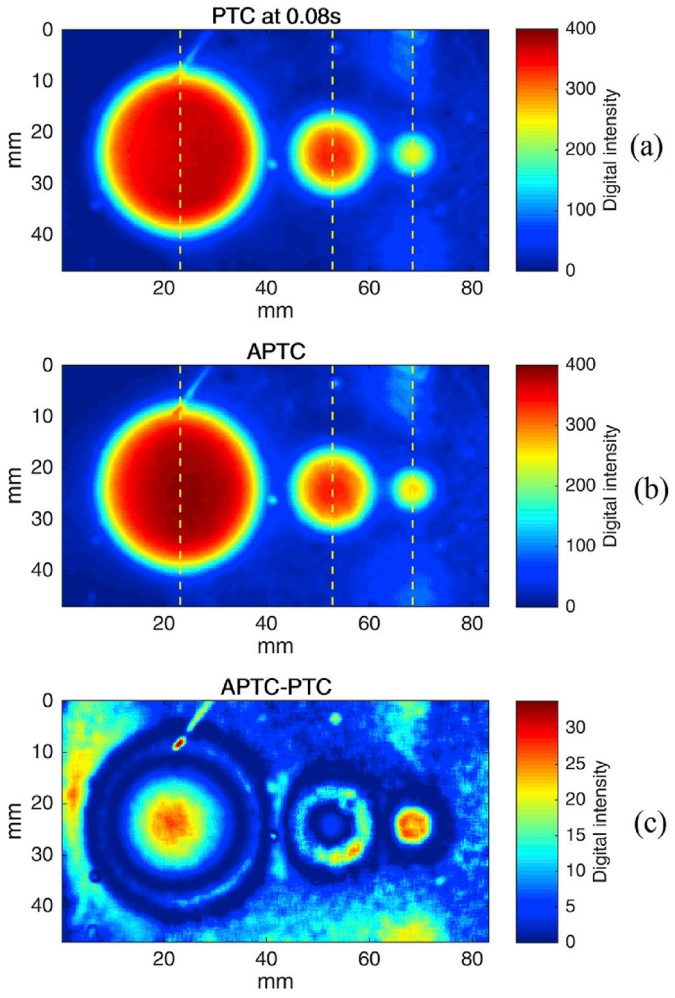

Fig. 5. Results of defect detection for the flat-bottom hole sample, where (a) shows the result of PTC at time 0.08 s, (b) shows the result of the proposed APTC method, and (c) shows the difference between these two images.

3.1. Experiment setup

The experiments were conducted using the Thermoscope® II pulsed-active thermography system which comprises of two capacitor bank powered Xenon flash lamps mounted in an internally reflective hood and a desktop PC to capture and store data. The scheme of the experimental set-up is illustrated by Fig. 1(a). A FLIR SC7000 series infrared radiometer (IR) operating between $3-5.1\mu m$ and a spatial resolution of 640×512 pixels was used to perform the inspection. The samples were placed with their surface perpendicular to the camera's line of sight at a distance of 250 mm from the lens (reflection mode configuration). It should be noted that the exported data of the used IR camera is in the unit of 'digital intensity', which was used for the analysis below instead of temperature.

3.2. Inspection of flat-bottom holes

Three flat-bottom holes were drilled at a depth of 1 mm from the top surface on a 5 mm thick aluminium plate, as shown in Fig. 4. The diameter of the holes are 32 mm, 16 mm and 8 mm from left to right respectively. The sample was inspected from the surface opposite to the drilled surface. The inspected surface of this sample was painted black to improve the surface emissivity of the sample. Considering the thickness of the sample and its high thermal diffusivity, a sampling rate of 50 Hz was used and a total 300 frames, equivalent to 6 s data length, were captured and analysed. After applying TSR with the model order of 7, a region of 10 by 10 pixels on the top left corner was defined as the sound area. Whilst there are multiple solutions to define the sound area, this simple reference system was adapted for ease of analysing the defect area.

To demonstrate the improvement of contrast between defective areas and sound areas, Fig. 5 shows the digital intensity contrast maps produced by PTC at the time of 0.08 s when the middle defect achieves maximum contrast and APTC respectively, as well as the difference between these two maps. It can be observed from the difference image that the overall contrast of defects has been improved and can be supported by the observation of high value of contrast increment on defective areas and very limited increment on sound areas adjacent to the defective area. This improvement is particularly prominent for the small defect represented by the 8 mm diameter hole. Such contrast improvement is important because the confidence level of the small defect will be lower when the proposed method is applied on Fig. 5(a) than on Fig. 5(b). The ground truth is that all three defects should have the same and maximal confidence level. To further inspect the contrast improvement for each defect, Fig. 6 plots the digital intensity contrast for three vertical lines as illustrated in Fig. 5. A close investigation of Fig. 6 suggests that the improvement of contrast for the large defects is limited. The improvement is significant for the small defect, where the ratio of the maximum and the minimum is increased from 8.01 to 8.53. The contrast improvement is not necessary for the small defect, which depends on the frame selected in the PTC. However, it is certain that APTC improves the overall contrast in comparison with PTC. As mentioned above, apart from the contrast improvement, another important reason to apply APTC rather than PTC in this paper is that the confidence map should be established on a property when all pixels arrive at the peak.

To demonstrate the advantage of APTC against PTC on confidence level quantification, Fig. 7 shows plots of the digital intensity contrast between three pixels in three defects respectively and the sound area. In the PTC method, an image of contrast at a certain time is selected to represent the result. The vertical dotted line shows an example of PTC, where the time when the large hole (black curve) arrives at the maximal digital intensity contrast is chosen. It indicates that P1, P2 and P3 are selected to represent the optimal contrast of these three pixels respectively. For the APTC method, the maximal digital intensity contrast for each pixel, denoted by P1, P4 and P5, is selected. Since the confidence level is directly related to the selected contrast value, the medium (red curve) and small defects (blue curve) have higher confidence levels using the APTC method than the PTC method. The selection of P2 and P3 is biased because they are not the optimal representative of the medium and small defects respectively.

The proposed method to calculate confidence map was then applied and the results for both semi-supervised and unsupervised mode are as illustrated in Fig. 8. For the semi-supervised mode, the boundary areas of

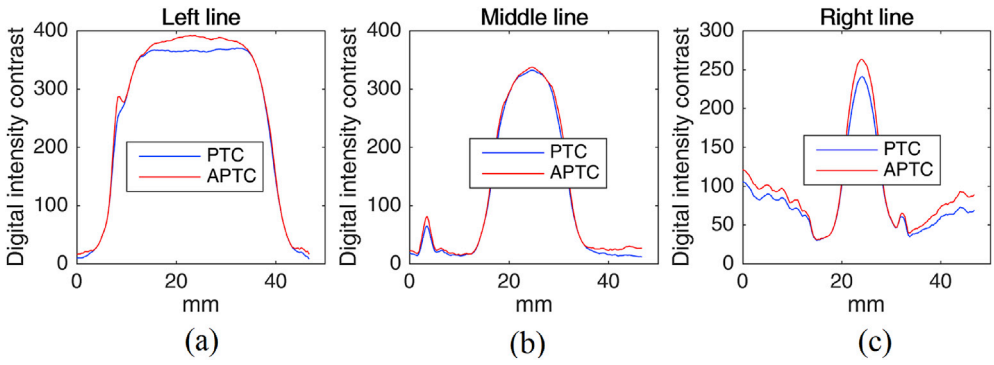

Fig. 6. Comparison of digital intensity contrast for the three vertical lines in Fig. 5 using the PTC and APTC methods respectively. (a) Left line, (b) middle line and (c) right line.

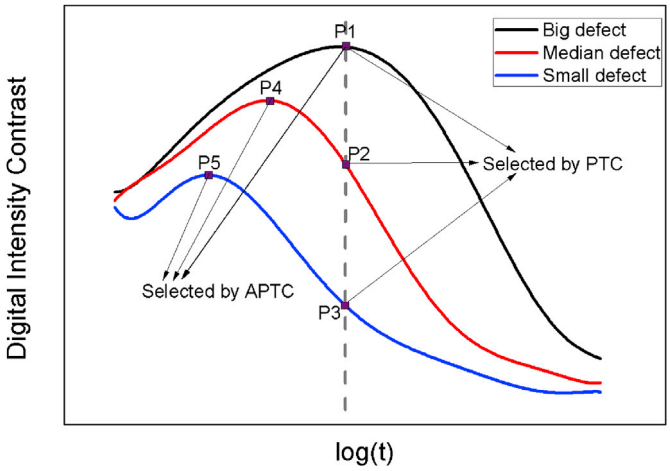

Fig. 7. Comparison of optimal property selection between PTC and APTC, where P1, P2 and P3 are selected property of three pixels selected from each defect using PTC, and P1, P4 and P5 are from APTC.

all four sides with a thickness of 5 pixels were defined as sound area, by which means, the effectiveness of non-uniform heat is considered. The number of sampled pixels to establish the reference, N, was chosen to be 100, the number of sampling iterations, Q, was set to 20 and the OR operator was used. Inspection of the left column of Fig. 8 suggests that the proposed method working under the semi-supervised mode produced a good result where three defective areas have been successfully detected with a very high confidence level (>99.8%). Some sound areas are also falsely detected as defective areas with relatively low confidence levels (<80%), which may be due to the surface finish of the sample itself. A confidence threshold of 95% was then chosen to produce the binary map, which was found to be in-line with the ground truth. If the unsupervised mode is chosen, as shown in the right column of Fig. 8, the three defective areas have been detected with relatively low confidence levels (>80%). It was further observed that for the small defect, the confidence levels were found to be less than 70%. From the numerical point of view, this could be due to the fact that the standard deviation of sampling is much larger in the unsupervised mode than that of the semi-supervised mode, which leads to a reduction of z^* value. From the confidence point of view, the higher confidence result from the semi-supervised mode is established assuming a high level of confidence that the sampling area is sound. In the case of the unsupervised mode, a reduced level of confidence was assumed in establishing the remits of the soundness of the sampling area. In Fig. 8, a confidence threshold of 70% was chosen to produce the binary map for the map using the unsupervised mode, where the defects with the large and medium sizes were detected successfully, the detection of the small defect failed.

To quantify the accuracy of the measurement, the dimensional cross-correlation between the ground truth and the produced binary map, denoted by R, is introduced to measure the accuracy of the defect detection. This value was calculated using the function *xcorr2* in Matlab. Values of R always range between 0 and 1. A higher value of R represents a high accuracy of detection. Table 2 shows the comparison of accuracy between the APTC and PTC (used the frame at 0.06 s, 0.08 s and 0.10 s) where the semi-supervised mode was used and the confidence threshold was chosen as 95%. Inspection of this table clearly showed the improvement of accuracy.

3.3. Sensitivity analysis

There are a few parameters introduced in the proposed method that may affect the performance.

The number of sampled pixels (N): Fig. 9 plots the influence of N to the R value in the semi-supervised mode, where the OR operator was used, Q was set to 20 and the confidence threshold was set to 95%. It can be observed from Fig. 9(a) that the R value increases significantly following the increment of N when $N < 50$ and when $N > 50$, the change of R value is very limited. Fig. 9(b) plots the ratio of change of R value against N, which indicates that the variation is less than 1% for $N > 100$. Thus N was chosen as 100 for all the remaining tests.

Logical operator and the iteration number Q: Fig. 10 plots the influence of Q to the R value in the semi-supervised mode for both 'AND' and 'OR' operators, where N was set to 100 with the confidence threshold of 95%. It can be inferred that the 'OR' operator produced improved results regardless of the selection of Q. The selection of Q can affect the results but the influence is very limited. Considering the computational speed, a small number of Q is suggested.

Confidence threshold: The selection of the confidence threshold is important to produce the defect map in a binary form. Fig. 11 plots the influence of this threshold in both semi-supervised and unsupervised mode, where N was set to 100, the OR operator was used and the Q value was defaulted to 20. For the semi-supervised mode, a higher confidence threshold produced better results and the performance was found to be consistent for all threshold values set more than 95%. It can be observed in the left middle plot of Fig. 8 that the confidence level of the left side of sound areas are between 70% and 80% potentially due to non-uniform heat application. If the threshold is selected to be less than 80%, these areas are classified as defect, which can be construed as a false positive. If the threshold is selected to be higher than 80%, these areas are correctly classified as sound areas. Thus, a sudden and a sharp change at 80% threshold has been observed. There is almost no change in performance if the threshold is higher than 95% because the confidence level of the true defects is about 99.9%. For the unsupervised mode, a relatively low confidence threshold should be chosen. For the considered example, the value of R peaks at a chosen threshold of 70% as indicated by Fig. 11. If the threshold is larger than 90%, there is no strong indication of the defect and therefore the R value is not applicable.

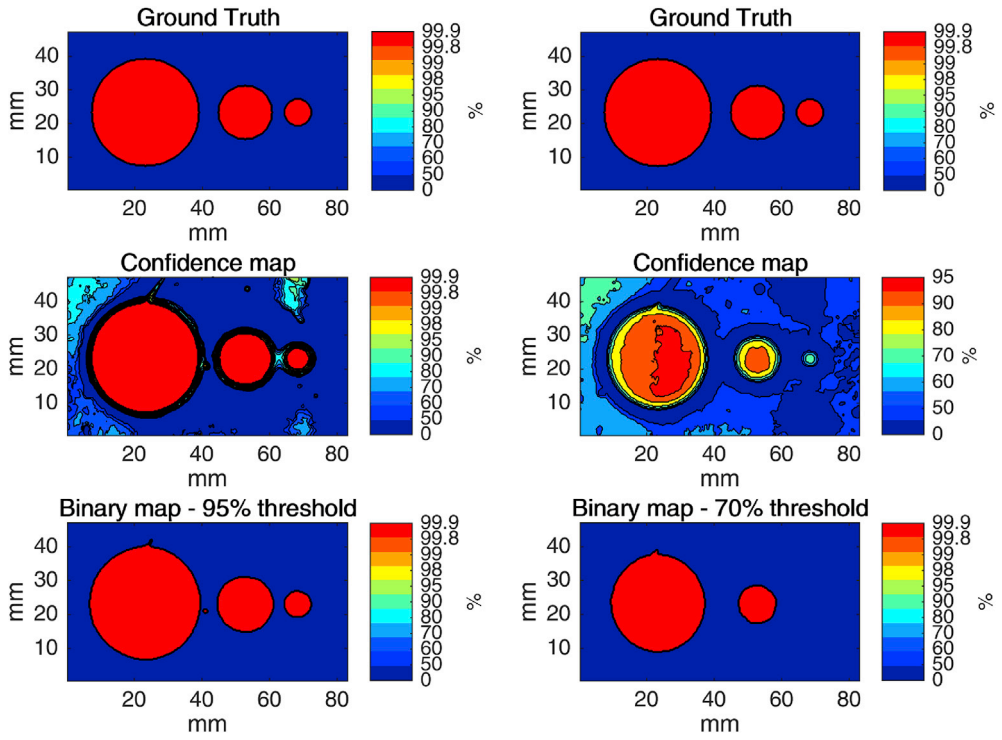

Fig. 8. Produced confidence map and binary map using a selected threshold for the semi-supervised classification mode (left column) and the un-supervised classification mode (right column). The top row shows the ground truth, the second row shows the confidence map, and the third row show the binary map with a selected threshold.

Table 2
Accuracy comparison between APTC and PTC for the flat-bottom hole sample, where the bold number indicates the highest R value.

Method		R
PTC	0.06 s	0.9653
	0.08 s	0.9613
	0.10 s	0.9486
APTC		**0.9743**

3.4. Inspection of impact damages in composite

Although impact damage in composites has been successfully detected with lock-in thermography [26], their evaluation using pulsed thermography still attracts more and more interests due to its advantage of speed and ease of deployment. Lock-in thermography allows better control of the energy deposited on a surface, which might be interesting if a low power source is to be used or if special care must be given to the inspected part.

Specimens were produced with the dimension of 150 × 100 × 4 mm, which were made of unidirectional Toray 800 carbon fibres pre-impregnated with Hexcel M21 epoxy resin. The laminates were subjected to a drop test machine with predefined energy levels using a 16 mm (diameter), 2.281 kg hemispherical indenter. The support used to hold the sample in place was designed as per BS ISO 18352. Impact

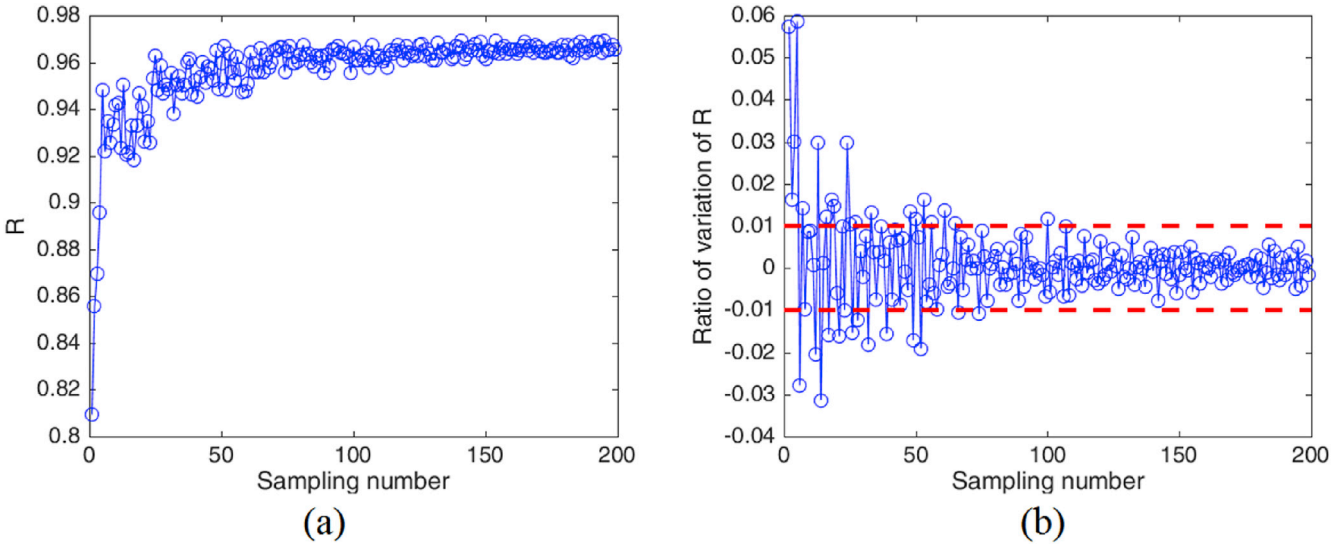

Fig. 9. Sensitivity analysis for the number of sampling N, where (a) shows the R value between the ground truth and the detected binary map and (b) show the ratio of variation of the R value.

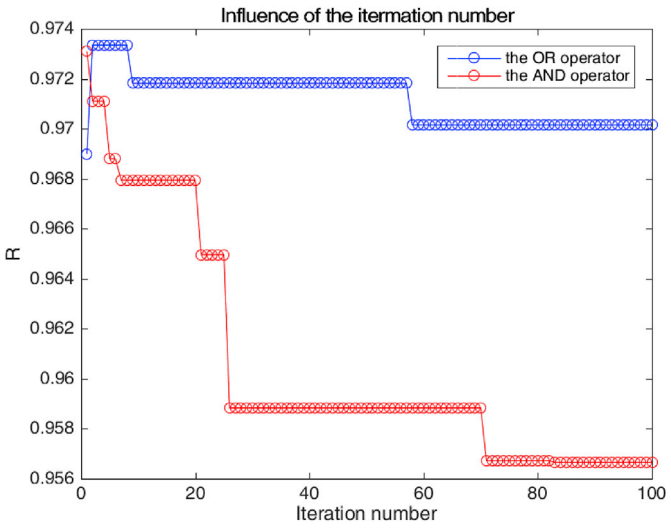

Fig. 10. Sensitivity analysis for the number of iteration Q and the logical operator.

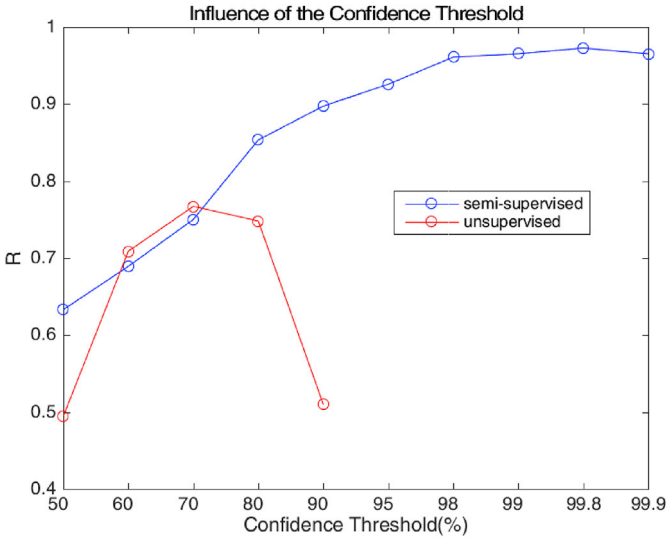

Fig. 11. Sensitivity analysis for the selection of confidence threshold for both modes.

energy was adjusted by changing the height of the drop-weight. The specimens were subjected to represent impact energies of 5, 10, 15, 20, 25 and 30 J (J) respectively. In all samples, each of the damage features are clearly visible from the impacted side, but they are hidden or less obvious from the rear surface, as shown in Fig. 12. The samples were inspected from the rear surface. Considering the thickness of the sample and its low thermal diffusivity, a sampling rate of 25 Hz was used, and totally 750 frames, equivalent to 30 s data length, were captured and analysed. After applying TSR with the model order of 7, a region of 10×10 pixels to the top left corner of the ROI was defined as the sound area. The ROI of 200×200 pixels was selected for the analysis and corresponds to an area of 63.8×63.8 mm.

The APTC1D was calculated based on Eq. (5) in this example and the mean and standard deviation of three trials ($M = 3$) for six samples are shown in Figs. 13 and 14 respectively. It can be observed from the mean maps that the overall value of APTC1D of the defective area increases following the increment of energy level, which suggests the increment of SNR. For the 5 J specimen, the visual inspection shows that the impact damage in the middle is so weak that it is not as prominent and was found to be close to the noise level of the sound area as seen in the top right corner of the image (see the top-left subplot of Fig. 13). With the increment of impact energy, the background noise significantly reduces visually. The maps of standard deviation measure the noise introduced due to multiple trials, where no direct indication of damage has been observed for low energy specimens, significant variation to the level of noise has been observed for the 25 J and 30 J specimen.

The proposed method working under the semi-supervised mode was then applied to acquire the APTC1D maps. The boundary areas of all four sides with a thickness of 5 pixels were defined as sound area. The value of N was set to 100, Q was set to 20 and the OR operator was used as deduced from previous trials above. Fig. 15 shows the estimated confidence maps based on the mean APTC1D. For the specimen with the impact energy larger than 5 J, the damage has been successively detected with high confidence levels (>99.9%), although some noise with relatively low confidence levels around the defects and boundaries have also been detected. For the 5 J specimen, although the damage is very weak, a defective area with confidence level around 80% has been detected. However, the noise in the top left corner is significant indicating that the damage and the sound area noise are at a similar level.

Fig. 16 shows the results using the unsupervised mode, where the overall confidence levels are less than those of the semi-supervised mode. It has been observed from the specimen with the impact energy larger than 5 J that there are consistently two areas (top right and bottom left) that have high values of confidence (>99.9%), while other areas have relatively low confidence. This phenomenon may be caused by the mechanism of the weight-drop machine itself. Fig. 17 plots the detected damage size after converting the confidence maps into binary maps with different confidence thresholds. For the threshold less than 90%, the damage size of the 10 J specimen is larger than that of the 15 J specimen, which in this case is not true because most of the background noise in both specimens is considered as the defective area. For the thresholds

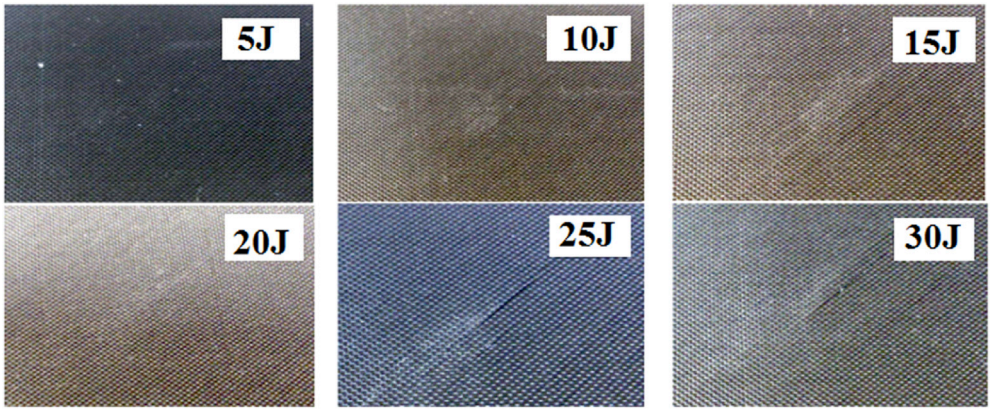

Fig. 12. A snapshot of the non-impact side of the samples showing limited indication of sub-surface damage.

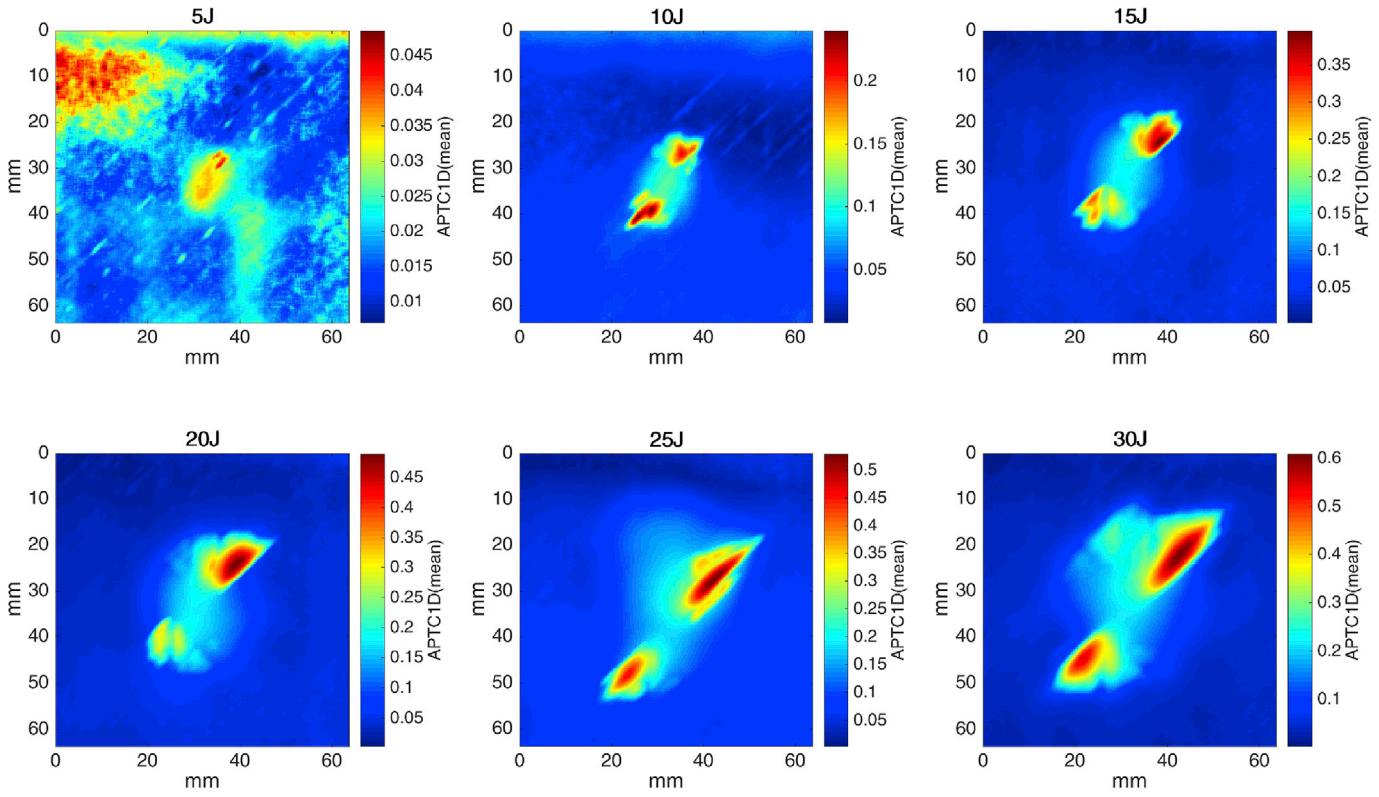

Fig. 13. Mean of defect detection based on 3 trials for all six samples.

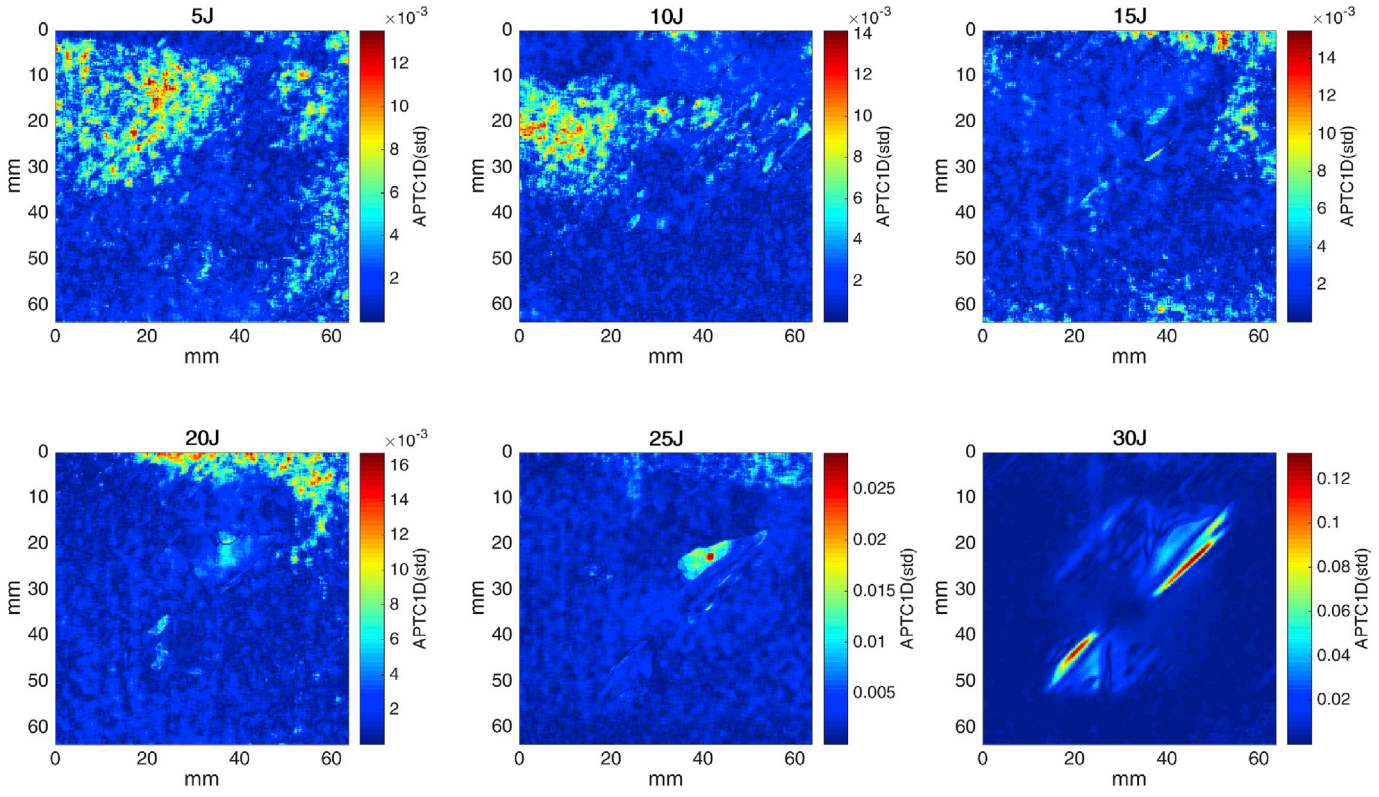

Fig. 14. Standard deviation of defect detection based on 3 trials for all six samples.

larger than 90%, the curves exhibit a linear trend against the increment of impact energy.

After the detection of the defect, some information can be extracted to further characterise the damage in more detail, such as depth measurement. A bonus of defect detection before depth measurement is that for the reference-based depth measurement methods, the selection of

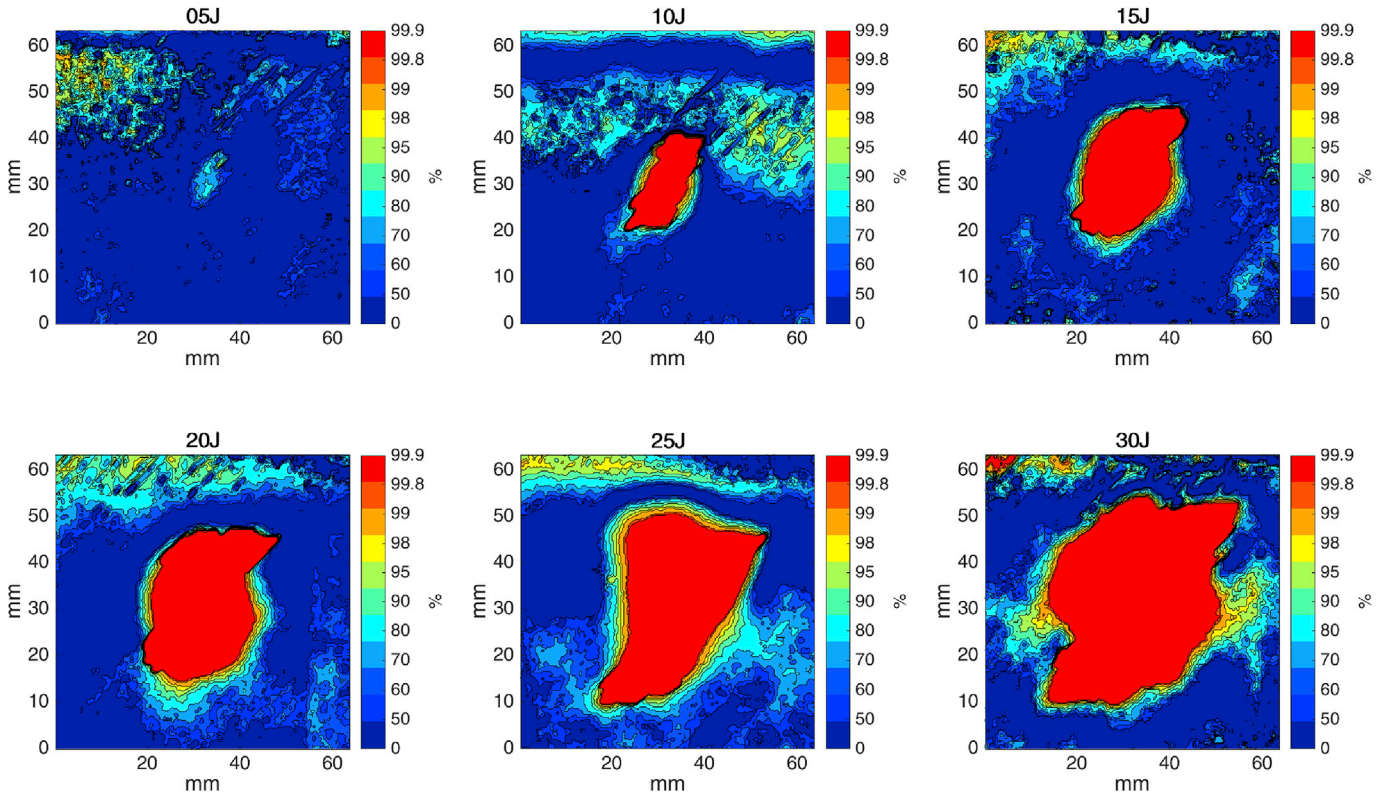

Fig. 15. Confidence maps produced from APTC1D using the semi-supervised mode for six samples.

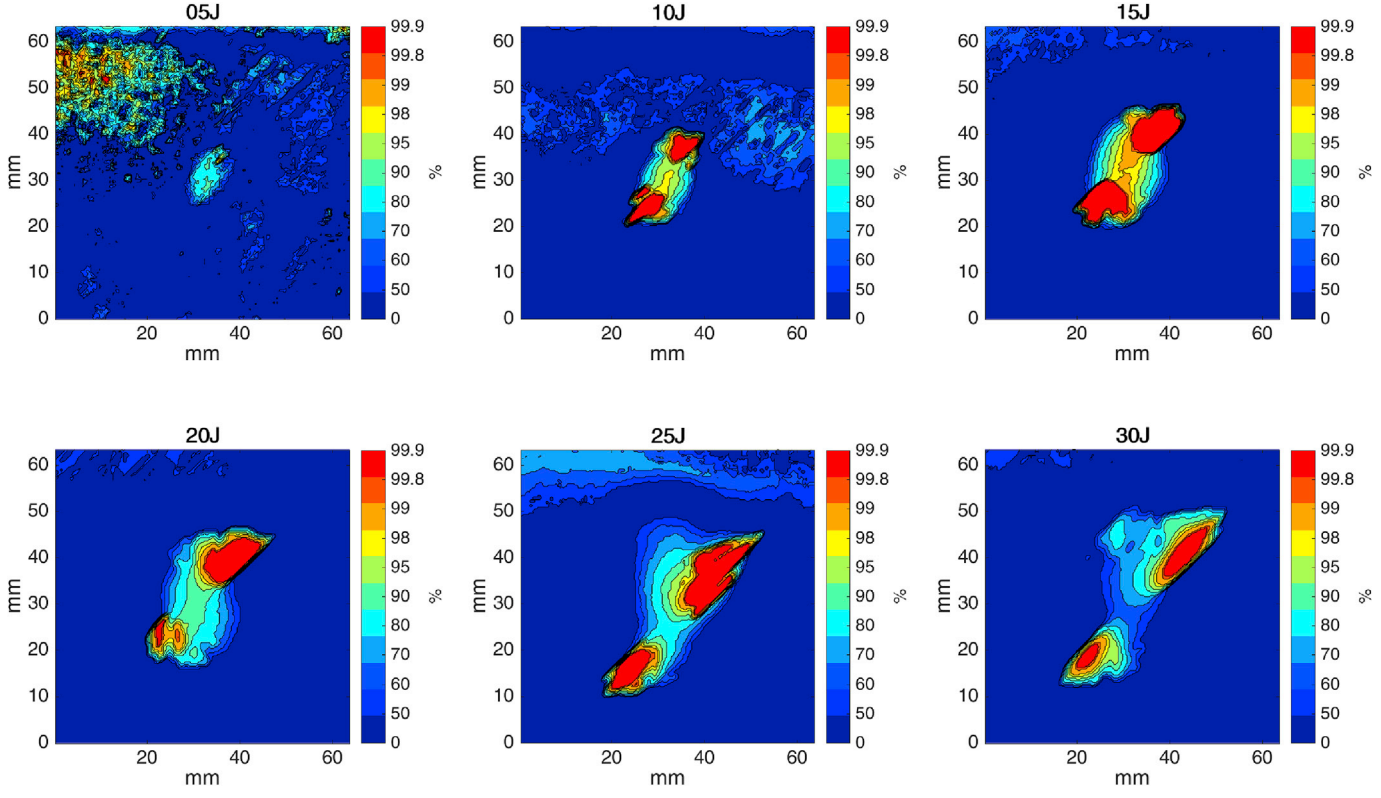

Fig. 16. Confidence maps produced from APTC1D using the unsupervised mode for six samples.

reference is straightforward now. Another benefit is that the depth measurement process is only applied on the defective pixels, which could significantly reduce the computational time. Fig. 18 shows an example of depth measurement using the PST method established on binary maps using the confidence threshold of 99.8%. The value of z-axis indicates the distance from the detected subsurface damage from the inspected surface

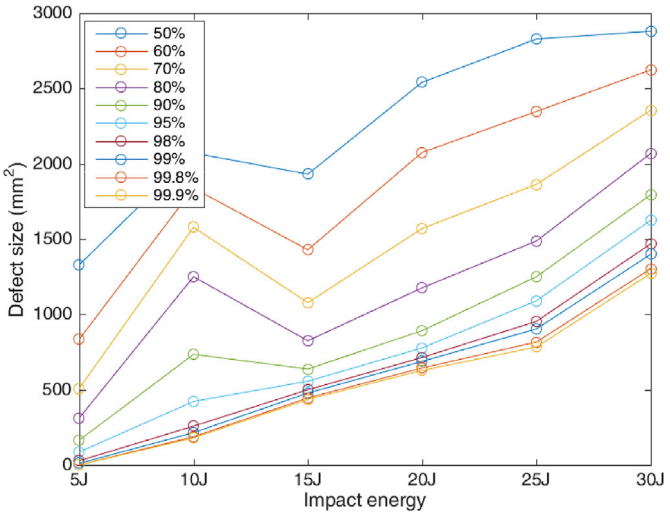

Fig. 17. Detected defect size for samples with different impact energy using different confidence threshold.

to the back surface. The 3D damage representation is based on reconstructing all the digital intensity maps acquired during the inspection. Based on the results obtained, it can be inferred that there is no strong indication of damage that can be detected on the 5 J specimen and hence the reconstruction fails to register any damage. The remaining reconstructions show specific pattern and are representative of the damage boundary as they occur from the surface of the laminate. Further, the 3D depth maps indicate the complex nature of impact damage occurring in composite materials. It has been observed that the delamination surfaces are not completely even as evidenced by the surface patterns which are found to be comparative for all specimen except the specimen with the 5 J impact. The main difference is that the delamination area continues to grow following the increment of impact energy. An interesting 'S' shape valley crossing the impacted circle area can also be observed, which is caused by the shock wave itself occurring at the time of the impact and the associated material response to the impact event. It should be noted that the depth maps presented are representative of the first boundary of damage from the surface of the laminate and any additional damage features occurring below this boundary have not been presented due to the nature and limitation of the pulsed thermographic inspection.

4. Conclusions

To enhance flexibility and reliability of NDT inspection, this paper introduced a concept of contour-based confidence map with its application framework for pulsed thermography. Traditional methods based on intensity of thermal images to interpret multiple defects in a single image can be misleading if one defect has significantly high temperature contrast than others. This approach transfers the intensity of thermal images into the confidence level of inspection to represent the defects, by which means multiple defects can be better visualised in a single image even when their temperature contrasts are significantly different.

Applications of the proposed technique on evaluations of flat-bottom holes and impact damage in composite laminates show that:

a. The APTC method can improve the overall contrast between defective areas and sound areas in comparison with PTC, particularly for an ROI with multiple defects.
b. The semi-supervised mode works better than the unsupervised mode due to the contribution of a pre-defined sound area.
c. The proposed method can successfully detect all three flat-bottom holes of diameter 32, 16, and 8 mm with a high accuracy ($R > 0.96$) when an appropriate confidence threshold is selected.
d. The proposed framework can effectively assess the impact damage in CFRP by offering the flexibility of decision making based on a selected confidence level and reducing the computation time for depth measurement.

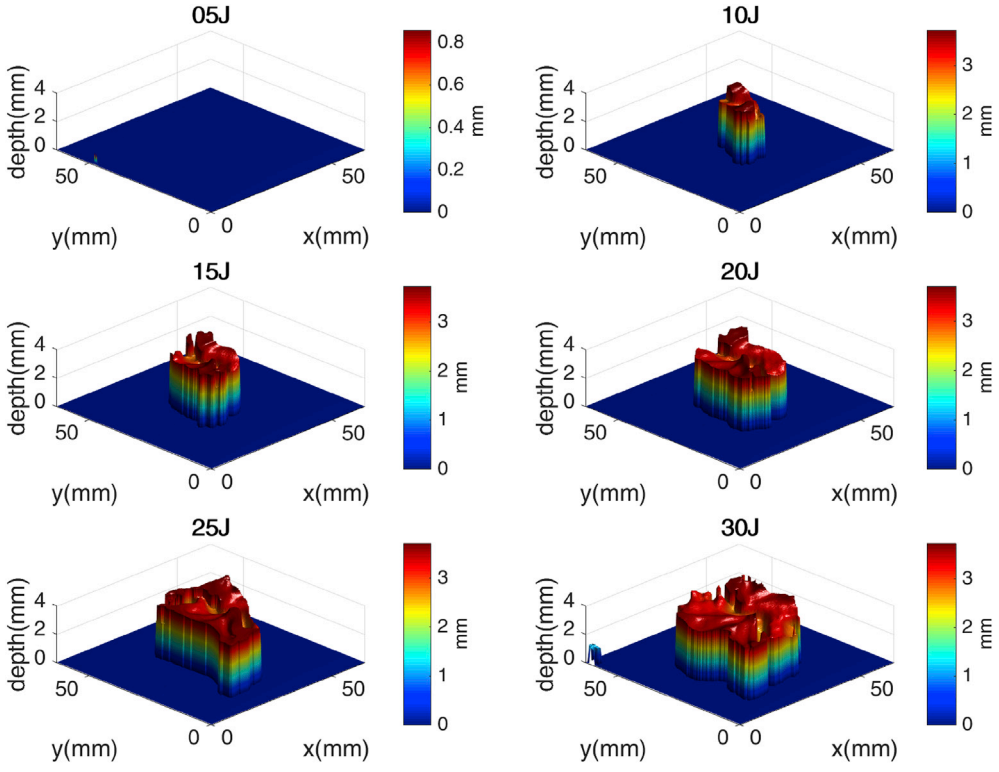

Fig. 18. Depth measurement using the PST methods for all specimen, which the depth indicates the distance from the defect surface to the back surface.

Sensitivity analysis shows that (i) the number of sampled pixels N is important and a larger number (N > 50) was suggested to produce more reliable results; (ii) the logical operator AND usually produces better results than OR; (iii) the influence of iteration number Q on results is very limited, so a small value of Q was suggested; (iv) A higher confidence threshold should be used to further analyse the defect where the approach works under the semi-supervised mode than that under the unsupervised mode.

To further improve the fidelity of the confidence map, different defect detection methods, different testing environmental parameters, different NDT techniques, and even different operators will be considered in future studies.

Acknowledgements

This work was supported by the UK EPSRC Platform Grant: Through-life performance: From science to instrumentation (Grant number EP/P027121/1). For access to the data underlying this paper, please see the Cranfield University repository, CORD, at: https://doi.org/10.17862/cranfield.rd.5480683.

References

[1] Zhao Y, Tinsley L, Addepalli S, Mehnen J, Roy R. A coefficient clustering analysis for damage assessment of composites based on pulsed thermographic inspection. NDT E Int 2016;44(0):59–67. Jun.
[2] Addepalli S, Zhao Y, Tinsley L. Thermographic NDT for through-life inspection of high value components. In: Redding L, Roy R, Shaw A, editors. Advances in through-life engineering services. Springer; 2017. p. 189–97.
[3] Minkina W, Dudzik S. Infrared thermography: errors and uncertainties. John Wiley & Sons Inc; 2009.
[4] Annis C, Gandossi L, Martin O. Optimal sample size for probability of detection curves. Nucl Eng Des 2013;262:98–105. Sep.
[5] Vavilov VP, Almond DP, Busse G, Grinzato E, Krapez J-C, Maldague X, et al. Infrared thermographic detection and characterisation of impact damage in carbon fibre composites: results of the round robin test. In: Proceedings of the 1998 international conference on quantitative InfraRed thermography; 1998.
[6] Kessler SS, Spearing SM, Soutis C. Damage detection in composite materials using Lamb wave methods. Smart Mater. Struct 2002;11(2):269.
[7] Gros XE. Characterisation of low energy impact damages in composites. J Reinf Plast Compos 1996;15(3):267–82. Mar.
[8] Lane B, Whitenton E, Madhavan V, Donmez A. Uncertainty of temperature measurements by infrared thermography for metal cutting applications. Metrologia 2013;50(6):637–53. Dec.
[9] Junyan L, Yang L, Fei W, Yang W. Study on probability of detection (POD) determination using lock-in thermography for nondestructive inspection (NDI) of CFRP composite materials. Infrared Phys Technol 2015;71:448–56. Jul.
[10] Duan Y, Servais P, Genest M, Ibarra-Castanedo C, Maldague XPV. ThermoPoD: a reliability study on active infrared thermography for the inspection of composite materials. J Mech Sci Technol 2012;26(7):1985–91. Jul.
[11] Duan Y, Huebner S, Hassler U, Osman A, Ibarra-Castanedo C, Maldague XPV. Quantitative evaluation of optical lock-in and pulsed thermography for aluminum foam material. Infrared Phys Technol 2013;60:275–80. Sep.
[12] Sun JG. Analysis of pulsed thermography methods for defect depth prediction. J Heat Transf 2006;128(4):329.
[13] Lau SK, Almond DP, Milne JM. A quantitative analysis of pulsed video thermography. NDT E Int 1991;24(4):195–202.
[14] Maldague XP. Theory and practice of infrared technology for nondestructive testing. New York: Wiley; 2001.
[15] Martin RE, Gyekenyesi AL, Shepard SM. Interpreting the results of pulsed thermography data. Mater. Eval 2003;61:611–6.
[16] Pilla M, Klein MT, Maldague XPV, Salerno A. New absolute contrast for pulsed thermography. In: QIRT 2002-6th int. Conf. Quant. Infrared thermogrvol. 1; 2002. p. 53–8. 1.
[17] Shepard SM, Lhota JR, Rubadeux B a, Wang D, Ahmed T. Reconstruction and enhancement of active thermographic image sequences. Opt Eng 2003;42(5):1337–42.
[18] Zhao Y, Mehnen J, Sirikham A, Roy R. A novel defect depth measurement method based on Nonlinear System Identification for pulsed thermographic inspection. Mech Syst Signal Process 2017;85:382–95. Feb.
[19] Vavilov VP, Burleigh DD. Review of pulsed thermal NDT: physical principles, theory and data processing. NDT E Int 2015;73:28–52.
[20] Altman DG, Bland JM. How to obtain the confidence interval from a P value. aug08 1 BMJ Jul. 2011;343. d2090–d2090.
[21] Favro LD, Han X, Kuo P-K, Thomas RL. Imaging the early time behavior of reflected thermal wave pulses. In: Proceedings of SPIE - the international society for optical engineering; 1995. p. 162–6.
[22] H. I. Ringermacher, R. J.A Jr, and W. A. Veronesi, Nondestructive testing: transient depth thermography US Patent No. 5711603, 1998.
[23] Krapez J-C, Lepoutre F, Balageas D. Early detection of thermal contrast in pulsed stimulated thermography. Le J Phys IV Jul. 1994;4(C7). C7-47-C7-50.
[24] Zeng Z, Zhou J, Tao N, Feng L, Zhang C. Absolute peak slope time based thickness measurement using pulsed thermography. Infrared Phys Technol 2012;55(2–3):200–4.
[25] Maldague X, Marinetti S. Pulse phase infrared thermography. J Appl Phys 1996;79(5):2694.
[26] Meola C, Carlomagno GM. Impact damage in GFRP: new insights with infrared thermography. Compos Part A Appl Sci Manuf Dec. 2010;41(12):1839–47.

https://dspace.org/

A confidence map based damage assessment approach using pulsed thermographic inspection

Zhao, Yifan; Addepalli, Sri; Sirikham, Adisorn; Roy, Rajkumar

2017-10-07
Elsevier
https://doi.org/10.1016/j.ndteint.2017.10.001
Zhao Y, Addepalli S, Sirikham A, Roy R. (2018) A confidence map based damage assessment approach using pulsed thermographic inspection. NDT & E International, Volume 93, January 2018, pp. 86-97

Downloaded from a DSpace repository. https://dspace.org/