Metrol. Meas. Syst., Vol. 27 (2020) No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

AUTOMATIC DEFECT DETECTION AND CHARACTERIZATION BY THERMOGRAPHIC IMAGES BASED ON DAMAGE CLASSIFIERS EVALUATION

Giuseppe Dinardo, Laura Fabbiano, Rosanna Tamborrino, Gaetano Vacca

Politecnico di Bari University, Department of Mechanics, Mathematics and Management, Via E. Orabona 4, 70125 Bari, Italy (✉ *giuseppe.dinardo@poliba.it*, +39 080 5963225, *laura.fabbiano@poliba.it*, *rosanna.tamborrino@poliba.it*, *gaetano.vacca@poliba.it*)

Abstract

In the framework of non-destructive evaluation (NDE), an accurate and precise characterization of defects is fundamental. This paper proposes a novel method for characterization of partial detachment of thermal barrier coatings from metallic surfaces, using the long pulsed thermography (LPT). There exist many applications, in which the LPT technique provides clear and intelligible thermograms. The introduced method comprises a series of post-processing operations of the thermal images. The purpose is to improve the linear fit of the cooling stage of the surface under investigation in the logarithmic scale. To this end, additional fit parameters are introduced. Such parameters, defined as damage classifiers, are represented as image maps, allowing for a straightforward localization of the defects. The defect size information provided by each classifier is, then, obtained by means of an automatic segmentation of the images. The main advantages of the proposed technique are the automaticity (due to the image segmentation procedures) and relatively limited uncertainties in the estimation of the defect size.

Keywords: thermography, non-destructive testing, thermal barrier coatings, image segmentation, uncertainty analysis.

1. Introduction

The *infrared thermography* (IR) is among the most attractive *non-destructive evaluation* (NDE) techniques allowing for the identification of material defects affecting the structures and specimens under test. The IR technique has the advantages of low cost, easy operation, high speed, and wide area coverage, from defect detection and characterization of engineering materials to the assessment of the state of historical and archaeological artefacts for cultural heritage preservation, [1–4].

The crucial aspect of such kind of NDE technology is the combination of infrared thermography technology and the specimen heating excitation which can be performed by means of different techniques, [1–4].

One of the most often employed NDE testing is the pulsed thermography. In this technique, the surface of the analysed component is heated with a brief light pulse by means of a high-power source.

An infrared camera, then, collects the temperature response of the specimen's surface, that can be obtained by means of a series of post-processing operations to be implemented to the acquired raw thermographic images. The possible presence of superficial and sub-superficial defects provokes an alteration of the heating up and cooling rates of the flawed regions, [2]. The thermographic images so obtained, exhibit the presence of thermal contrasts in those pixels corresponding to the defected regions of the specimen under analysis.

Such a contrast highlights possible defects. However, the acquired raw thermograms are characterized by non-neglectable noise levels. These interfering effects are caused by the presence of the background radiation and, possibly, not uniform heating of the target surface.

Therefore, a direct analysis of the raw thermal images does not allow for an accurate defect detection [5]. In this regard, the scientific literature reports several effective and efficient methodologies, aimed at signal enhancement, *e.g. thermographic signal reconstruction* (TSR) [6, 7], *pulsed phase thermography* (PPT) [8, 9], *principal component thermography* (PCT) [10], the slope (m) and the determination coefficient (R^2) of the linear fit of the properly treated thermal data on the logarithmic scale [11].

Recently, the *long pulse thermography* (LPT) technique has been profitably developed. In LPT technique, the surface of the component under investigation, is heated up for few milliseconds [12] (typically hundreds of milliseconds up to a few seconds). There are so many applications mentioned in the scientific literature according to which the long pulse excitation technique can be more effective for some applications [13–16].

The authors of references [12] and [17] have shown that the processing techniques, commonly used for the Dirac delta heating pulse, can be applied to finite duration heating pulse.

The aim is to refine and improve the least square fit of the thermal data (during the surface cooling) obtained by a series of novel processing steps to be implemented in the thermograms, acquired by means of the LPT technique. The effectiveness of defect detection of the fit slope (m) and the determination coefficient (R^2) will be significantly enhanced by adopting an optimization algorithm to better utilize the spatial information derived from the thermographic data. Additionally, the authors introduce the fit intercept (q) and fit standard deviation (S_ρ) as defect detectors, [18].

Within the framework of the proposed technique, the thermal image segmentation is fundamental for the detection of defects. The segmentation distinguishes the image pixels related to flawed regions from the sound pixels and it is mainly based on grouping pixels (sharing common colour features) into connected regions, [19–21]. Such approach, allows the determination of defect size directly from thermographic images, as already proven in [22].

The presented approach not only enhances the resolution of the detection but also facilitates automated selection of the suspected regions without any prior knowledge about features. The experimental results strongly indicate that the algorithm can reliably extract the inner defect information and improves the contrast between the defect area and the sound area, allowing also for an accurate and precise sizing of the detected defects.

The study will be conducted on a series of test specimens exhibiting *thermal barrier coatings* (TBCs). The test specimens are affected by partial detachment of the TBC. The aim pursued, is the detection, localization and sizing of such defects.

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

1.1. Theoretical background

In long pulsed thermography, a high-energy pulse (during from few hundreds of milliseconds up to a few seconds) is projected onto the surface under analysis, by means of one or more than one flash lamps. Figure 1 illustrates a typical experimental test bench normally used in thermographic inspections.

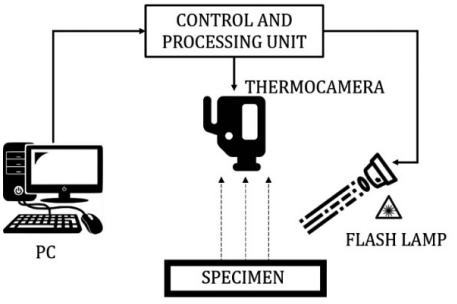

Fig. 1. Experimental assembly for the pulsed thermographic inspection.

The time duration of the exciting pulse influences the thermal behaviour of the entire specimen under test. Once the heat flux is absorbed by the sample surface, the heat conduction through the material takes place. Therefore, the propagation of the heat inside the specimen leads to a progressive decrease of temperature surface. An IR thermal camera, controlled by a computer, acquires the surface temperature response.

2. Materials and experimental procedure

2.1. Tested specimens description

The specimens used are three discs (diameter 25 mm and thickness 5 mm), with metallic bondcoat and a ceramic topcoat (thickness 200 μm) which is *yttria stabilized zirconia* (YSZ), deposited with the *high velocity oxy fuel* (HVOF) technique. The specimens, developed and exploited by the same authors in [18], have been realized keeping the same thickness of the substrate and of the TBC coating, but with an adhesion defect (Fig. 2a). The method used for the realization of the specimens with detachment of TBC involves the following implementation phases: drilling of the basis metal specimens, insertion of a copper pin, moulding and sandblasting

a) b)

Fig. 2. Tested samples during the production of the adhesion defect (a), picture of one of the reference-coated samples (b), [18].

of the surface, application of the bondcoat and topcoat, pin removal, insertion of base material plugs (Fig. 2a). Figure 2b shows the final sample.

In this paper, a circular defect is simulated (at the centre of the sample), with 2 mm, 3 mm and 5 mm. The base specimen is identical for the three case studies. The specimens are identified with letters A, D and E, respectively.

2.2. Experimental setup

The test instrumentation, already employed in a previous work by the same authors [18], comprises a 1064 nm wavelength Ytterbium pulsed fibre laser generating a pulse excitation of duration 500 ms and the collimated laser spot with a diameter of about 8 mm.

Figure 3 shows the used experimental setup, comprising the thermal camera, the laser lamp and the tested specimen (on the left side of the picture). Table 1 lists the main features of the used thermal camera.

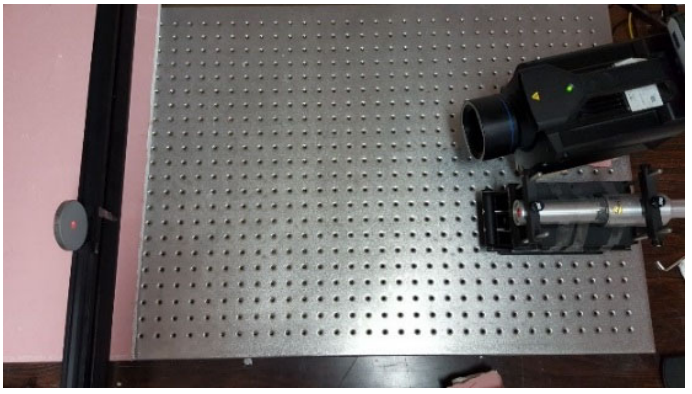

Fig. 3. Experimental assembly for the pulsed thermographic inspection, [18].

Table 1. FLIR X6540 SC features.

Feature	Specification
Detector type	Focal plane array
Sensor material	InSb
IR max. resolution	640 × 512
Applied windowing	176 × 136
Frame rate	500 Hz
Spectral range	3–5 μm
Thermal sensitivity	< 25 mK (18 mK typical)
Temperature measurement accuracy	±1°C (1.8°F) or ±1%
Lens	SWB 50 mm f/3 Bayonet
Focal length	50 mm (±5%)
Field of view in degrees	10.97
Distance to target	0.70 ± 0.07 m
Instantaneous field of view	0.25 ± 0.03 mm/pixel

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

2.3. Measurement procedure

In this sub-section, the authors introduce the overall measurement procedure performed on specimens A, D and E (of diameter size 2 mm, 3 mm and 5 mm respectively).

For each tested specimen, a repeatability test has been performed. Performing repeatability tests is an essential part of the estimation of the uncertainty granted by the introduced algorithm and it is aimed at evaluating the degree of precision of the outcome of the algorithm.

In this work, the authors performed five repetitions for each specimen. Each repetition consists in the application of a thermal cycle to the analysed specimen and the collection of temperature data of its surface, during the cooling stage. The consecutive cycle is applied once the tested sample has reached a standard temperature referred to the laboratory conditions. Therefore, for each specimen, five data sets are obtained.

When performing a repeatability test, the experimenters made sure that the repetitions are performed under the following conditions:

- Same method.
- Same operator.
- Same test location.
- Same experimental equipment.
- Same environmental conditions.
- Same measurement configuration.
- Same test specimen.

Cycles repeated over a relatively short period of time (compatible with the achievement of the cooling completion of the specimen between each repetition). Then, the proposed algorithm is implemented in each dataset, in order to obtain the information about defect location and size pertaining to each test and analysed specimen. Type A uncertainty evaluation approach is adopted, in order to estimate the repeatability of the technique, [23].

3. Description of the proposed algorithm

In this section, the authors describe the proposed algorithm and the steps leading to the defects detection and characterization. The algorithm consists in a series of operations to be directly implemented on the thermographic images.

Figure 4 shows a synthetic chart of the introduced algorithm, with a brief description of the implemented steps, which are going to be discussed in the following sections.

3.1. Data pre-processing

The data pre-processing steps are basically aimed at the raw data arrangement in order to make them suitable for the implementation of the main algorithm. The main purposes achieved are the de-noising of the raw thermograms and the automatic selection of the *region of interest* (ROI), referring to the effective specimen area where the main algorithm is going to be implemented. The target of this sub-section is to illustrate the operations implemented in the data pre-processing stage. The figures shown below are relative to one of the tests concerning specimen A (circular defect of 2 mm size).

In the de-noising operation, a Gaussian-weighted moving average over each window of a single raw thermographic image is considered. Figure 5 and Fig. 6 show, respectively, the time sequence of the raw thermograms and the Gaussian smoothing window employed for the smoothing of the

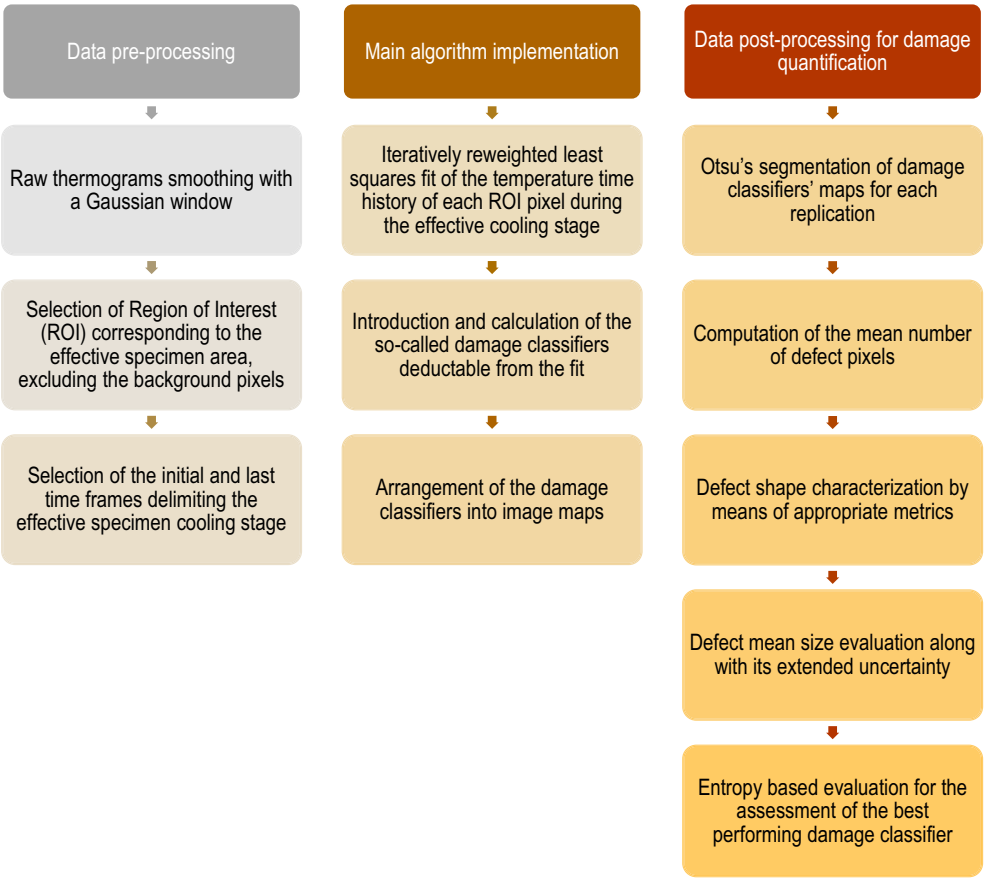

Fig. 4. Flowchart of the proposed algorithm.

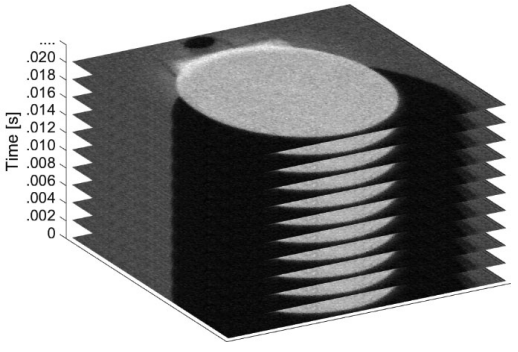

Fig. 5. Time sequence of raw thermograms.

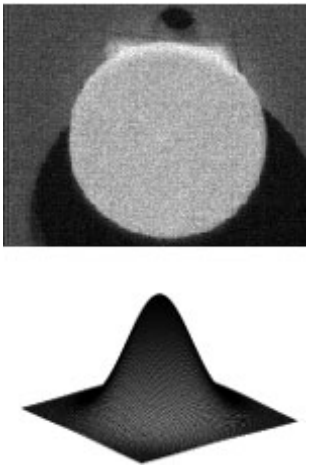

Fig. 6. Single raw thermogram with Gaussian window for image smoothing.

www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl
Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

images. Figure 7 and Fig. 8 illustrate the comparison between an arbitrary raw thermogram and the relative smoothed version. The smoothing operation is desirable, although the overall trend of the surface temperature is predictable (and monotonically decreasing during the post-excitation period). Indeed, the same single pixel may not behave as predicted in two consecutive times, because of inevitable and undesirable noise influence. These undesirable effects are typically composed of high frequency components and they should be removed in order to couple each pixel with a physically consistent monotonically decreasing temperature trend.

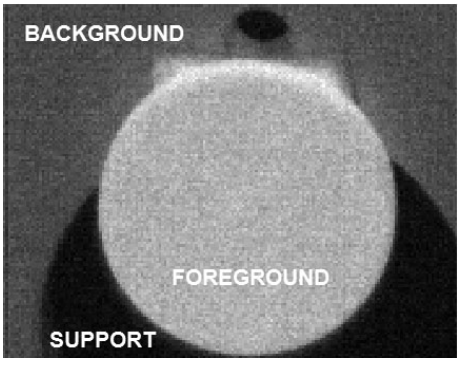

Fig. 7. Raw noisy thermogram.

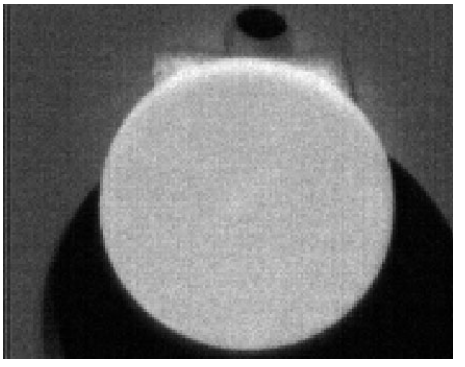

Fig. 8. Filtered thermogram (Gaussian window, 7×7 pts).

A further crucial operation here considered refers to the identification of the group of pixels relative to the effective specimen area. The selection of the ROI is fundamental because it allows to isolate the specimen area from the surrounding image background, with the advantage of considerably limiting the number of pixels subjected to the main algorithm.

The ROI selection is automatically performed and it is based on the mean temperature of the time history of the acquired thermograms prior to the laser excitation of the sample under analysis. Figure 9 illustrates the time history of a single pixel on the specimen surface. The red dashed line indicates the time (prior to the laser excitation) considered for the computation of the pixel mean temperature. Figure 10 shows the mean temperature map prior to the thermal excitation, along with the relative colormap with the temperature levels indicated.

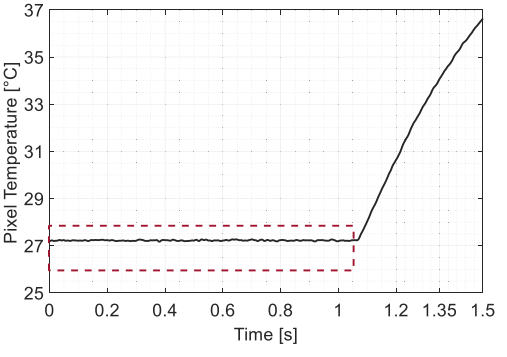

Fig. 9. Time history selection for a single specimen pixel.

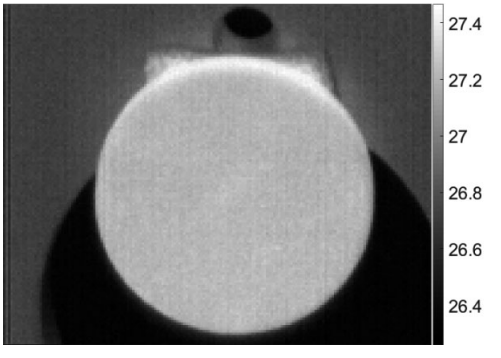

Fig. 10. Mean Temperature map prior to laser excitation.

It is worth stressing here that the figures contained in this subsection, are relative to a single repetition of the excitation of specimen A. Since the main object of this section is to illustrate and present the algorithm, the images for specimens D and E (tested in this work) are omitted.

The ROI extraction from the mean temperature map is based on Otsu's segmentation algorithm [24]. Basically, the algorithm is based on the image threshold approach. The image pixels are grouped into $m = 3$ classes, each of them representing, respectively, the background, the support and the specimen. The implementation of the algorithm is automatic, once the number of classes (*i.e.* the objects to detect) is stated. Then, the algorithm computes the pixel threshold values separating the three classes. Such automatic computation is accomplished once the minimum of the *intra-class variance* is achieved, [25].

Figure 11 identifies the three areas (background, specimen's support and the main specimen) on the basis of the Otsu's algorithm. The indication of the mean temperatures relative to each area is reported in Fig. 12.

Fig. 11. Identification of the three areas based on Otsu's algorithm.

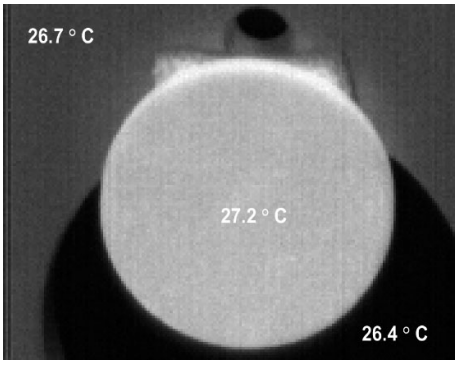

Fig. 12. Mean temperatures of the three detected areas.

The foreground area extracted from the segmentation operation (Fig. 12), is not circular like it should be. Indeed, the upper portion of the segmented area, is irregular and its pixels exhibit a temperature that makes them belong to the material. This discrepancy is due to an anomalous temperature distribution among those pixels, which have to be considered as artefacts. So, the first step of the algorithm implementation deals with the removal of the pre-excitation time history. Briefly, the mean pre-excitation temperature (from the previous step) is subtracted from the effective temperature time history. This operation (performed for each pixel) provides the term ΔT (difference between the actual temperature and the average temperature prior to the thermal excitation). Figure 13a shows the whole-time temperature trend of a pair of those artefact pixels, compared to another pair of pixels actually belonging to the material. Figure 13b shows their locations. The remarkable differences in those trends allow to set a criterion of further restriction of the ROI area, limiting the ROI to the pixels exhibiting a temperature trend that is physically consistent with heating and cooling stages. In addition, Fig. 13b shows the actual ROI area.

Figure 14 is the chart of the temperature trend of a single pixel, Fig. 15 shows the same trend in the log-log scale. In Fig. 15 the *pre-excitation stage* (P.S.), the effective cooling (*intermediate stage*, I.S.) and the *very late stage* (L.S.) are shown. With reference to Fig. 15, the whole trend of ΔT is significantly varying. The Fourier heat diffusion law (valid for semi-infinite solid

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

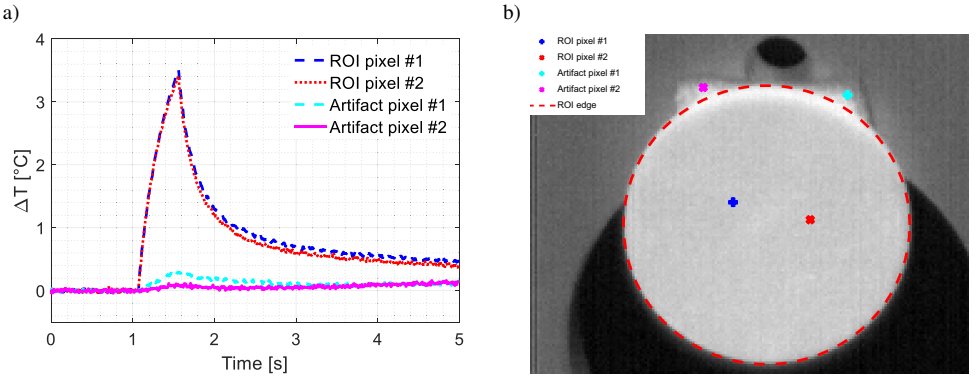

Fig. 13. Temperature time history of two arbitrary ROI pixels and two artefact pixels (a), their spatial location (b).

instantaneously heated) is strictly valid in the intermediate stage, where the actual sample cooling is occurring.

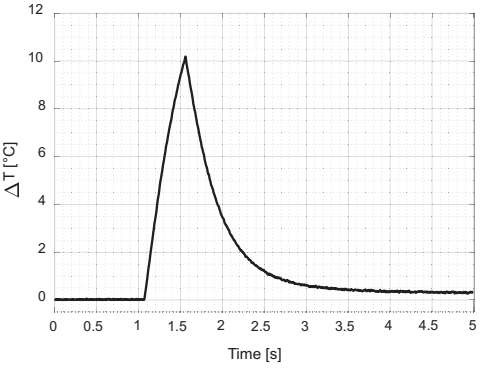

Fig. 14. Entire ΔT time history of an arbitrary ROI pixel.

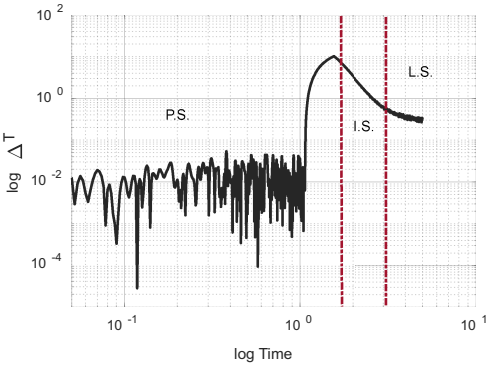

Fig. 15. Log-Log ΔT time history for the same pixel.

Since the analysis should be performed in the actual cooling stage, the appropriate removal of the early and the last stages is fundamental. In addition, the detection of the beginning and the ending frames of the intermediate stage may lead to deceptive results. Indeed, the accurate selection of the beginning of the cooling stage (in particular immediately after the laser excitation) may be compromised because of the possible infrared camera saturation, exacerbating the non-linear behaviour and invalidating the hypothesized model. In the proximity of the late stage the hypothesized trend is less clear because of the noise due to non-negligible convection and/or stray radiation mechanisms.

The first cooling frame (of each specimen pixel) has been set fifty frames after the maximum temperature (for each pixel) is reached, to ensure that the temperature decay has actually started. The selection of the last frame has been performed on the basis of statistical considerations.

The last frame is in correspondence with the time when the ΔT curve crosses the band shown in Fig. 16 for the first time.

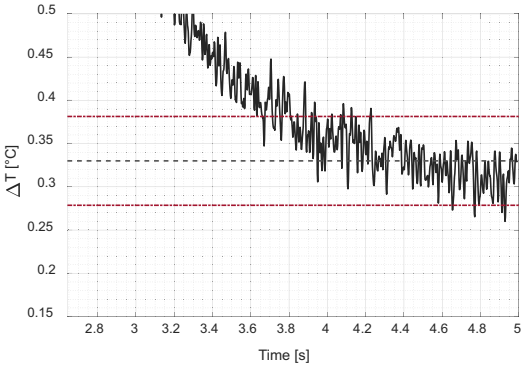

Fig. 16. Last frame selection. Mean ΔT, the upper and lower bounds about mean ΔT are represented. Confidence interval of 95% $\bar{\Delta T} \pm 2\sigma$, σ standard deviation.

3.2. Main algorithm description

The main feature of the proposed algorithm is to simplify and automate the detection of defects affecting samples with thermal barrier coatings. As pointed out in the introductory section, many algorithms have been introduced in the NDE field. Surely, the *thermographic signal reconstruction* (TRS) [8], is among the most performing techniques. Along with the main technique employed for the thermographic analysis, a detailed and careful analysis of the algorithm results should be considered. Indeed, only a correct interpretation of the analysis outcomes can lead to an accurate and precise localization and quantification of the defects.

The proposed technique is based on a series of enhancements of the linear fit of the log-log time history of the surface ΔT, [11, 17]. The previously described automatic selection of the ROI and the actual cooling stage of each pixel (within the ROI) represent fundamental steps towards the efficiency of the proposed technique.

The idea inspiring this work is to implement a linear least squares fit of the actual cooling stage of each pixel representative of the specimen surface. From the consideration made in the introductory section, it can be argued that sound ("good") pixels should have a slope of -0.5 that is independent of material properties. Pixels exhibiting anomalies of thermal diffusion mechanisms (symptomatic of defects) may deviate from such a trend. Therefore, a simple linear least squares analysis for each pixel is useful for the discrimination of defect pixels. Figure 17 and Fig. 18 illustrate the temperature data and the linear fit for the same defect pixels and sound pixels, respectively. The differences between the trends in Fig. 17 and Fig. 18 support the idea of grouping defect and non-defect pixels, according to the values taken by certain fit features.

In this paper, the authors introduce a robust version of least squares technique. The algorithm takes advantage of *iteratively reweighted least squares* (IRLS) fit, using a Talwar weighting function, [26–27]. Generally, the IRLS technique addresses the issue of data containing significant outliers. In this study, the relative importance of possible outliers (predominantly in the beginning and the end frames of the log-log ΔT time trend and possibly due to incorrect selection of both frames) could lead to misleading fit estimation. Therefore, an iterative weighted regression (accomplished by iteratively weighting each data point) ensures better and more precise results.

The authors have extracted the distributions of fit slope and determination coefficient for each specimen pixel. Additionally, the authors also introduce the fit standard deviation (S_ρ) and intercept (q) distributions. The following figures present the robust regression results (set in figure maps) for the three tested specimens (A, D and E).

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

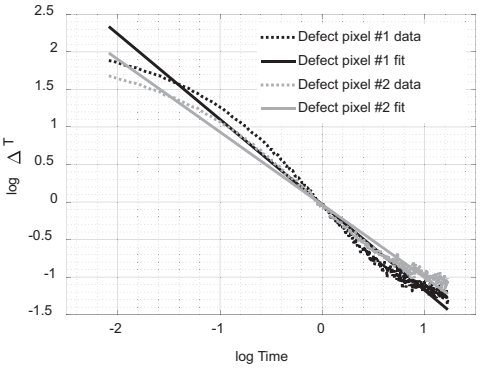

Fig. 17. Data and linear fit comparison plots for defect pixels #1 and #2.

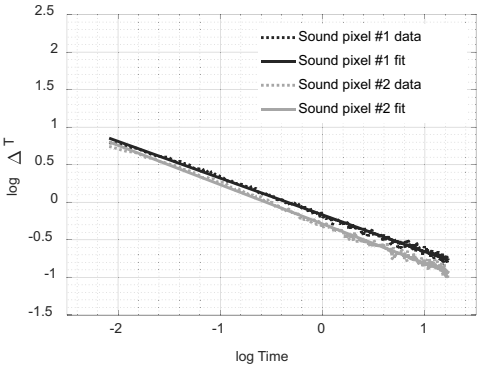

Fig. 18. Data and linear fit comparison plots for sound pixels #1 and #2.

Figure 19 up to Fig. 22 show, respectively, the slope, the R^2, the intercept and the standard deviation maps of the fit of each pixel relative to specimen A.

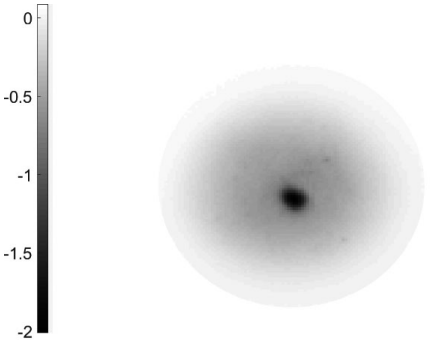

Fig. 19. Map of fit slope of specimen A.

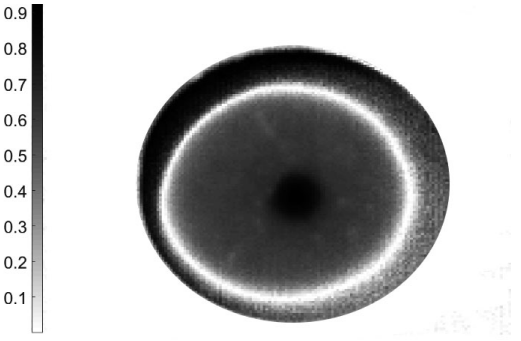

Fig. 20. Map of determination coefficient, R^2, specimen A.

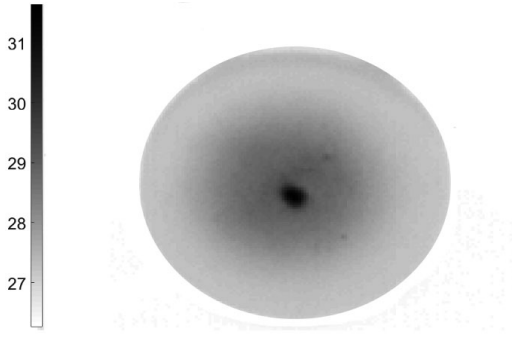

Fig. 21. Map of fit intercept, q, specimen A.

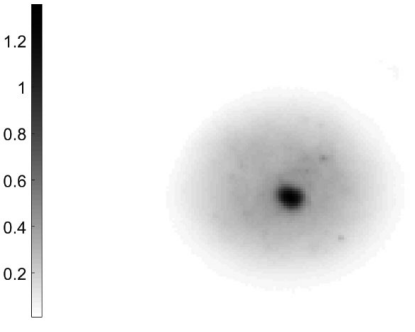

Fig. 22. Map of fit standard deviation, S_ρ, specimen A.

Figure 23 up to Fig. 26 show the map of fit slope, the determination coefficient distribution, the fit intercept map and the fit standard deviation for specimen D.

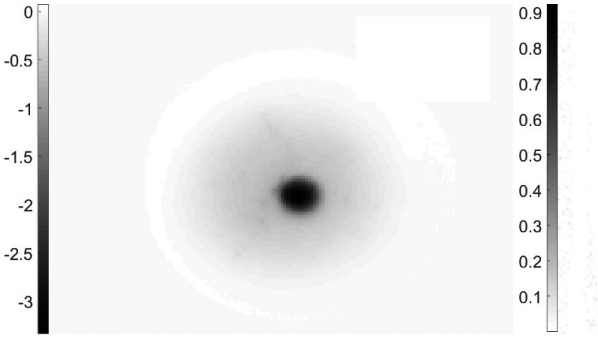

Fig. 23. Map of fit slope of specimen D.

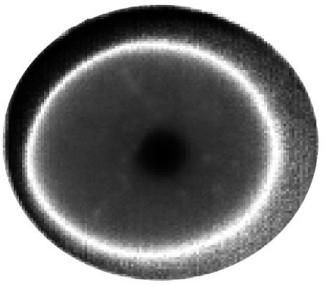

Fig. 24. Map of determination coefficient, R^2, specimen D.

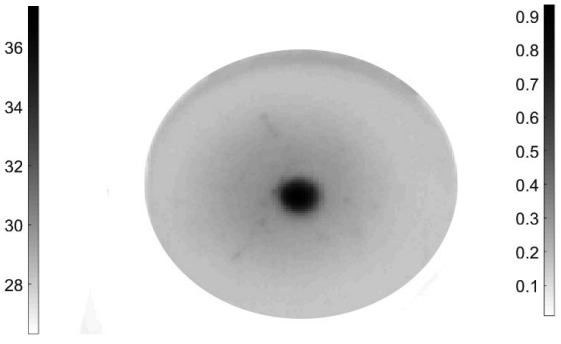

Fig. 25. Map of fit intercept, q, specimen D.

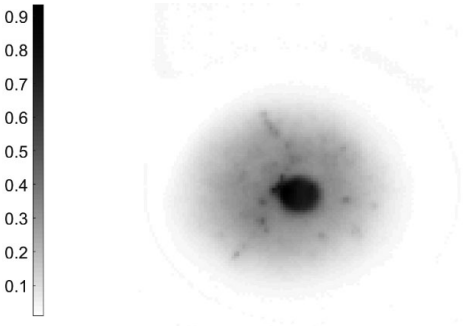

Fig. 26. Map of fit standard deviation, S_ρ, specimen D.

Figure 27 up to Fig. 30 show, respectively, the slope, the R^2, the intercept and the standard deviation maps of the fit of each pixel relative to specimen E.

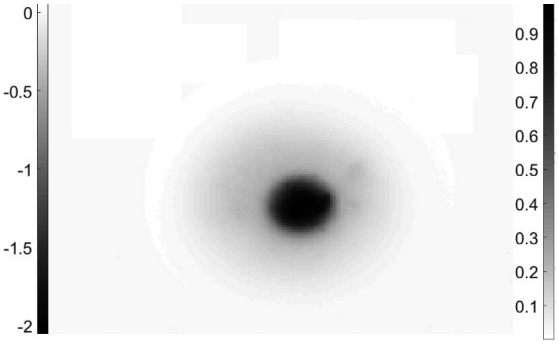

Fig. 27. Map of fit slope of specimen E.

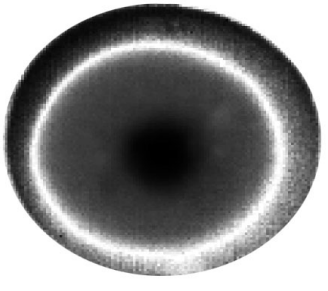

Fig. 28. Map of determination coefficient, R^2, specimen E.

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

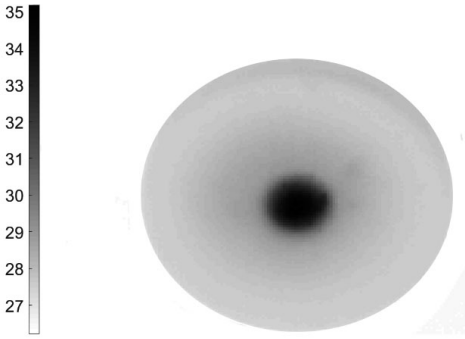

Fig. 29. Map of fit intercept, q, specimen E.

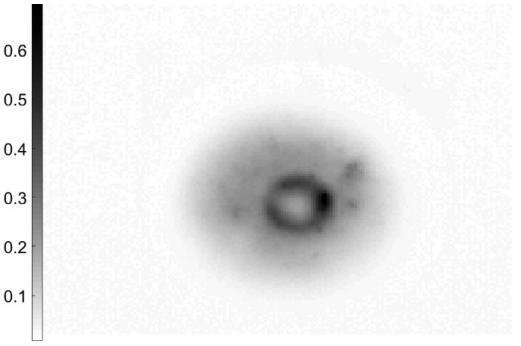

Fig. 30. Map of fit standard deviation, S_ρ, specimen E.

The presented maps give interesting qualitative indications of both location and size of the defects in the three tested specimens. Along with the optimization of the regression technique (for the mitigation of the bias effect introduced by the not proper selection of the first and last frames) and the introduction of the fit intercept and standard deviation as defect markers, the authors propose a further segmentation of the fit maps. This additional step allows to distinguish between sound and defect pixels within the ROI and quantify the size of the detected flaws.

3.3. Data post-processing

This sub-section is devoted to the description of post-processing operations to implement on the image maps of the fit parameters introduced in the previous sections. These operations, representing the main peculiarity of the proposed algorithm, allow to elaborate on the information and results derived from the preceding steps, and put them onto images defining the damage classifier maps.

The damage classifier maps have been obtained by means of an Otsu's segmentation of the fit parameter images. Since the information brought by the images pertains only to the ROI pixels, it is possible to divide the ROI domain into two classes: the sound pixels and the defect pixels.

The threshold value distinguishing both sub-domains is computed by maximizing the *between-classes variance* (σ_B^2). By way of example, Fig. 31 and Fig. 32 illustrate the histogram of pixel counts and the σ_B^2 function for fit slope, intercept and standard deviation relative to a test performed

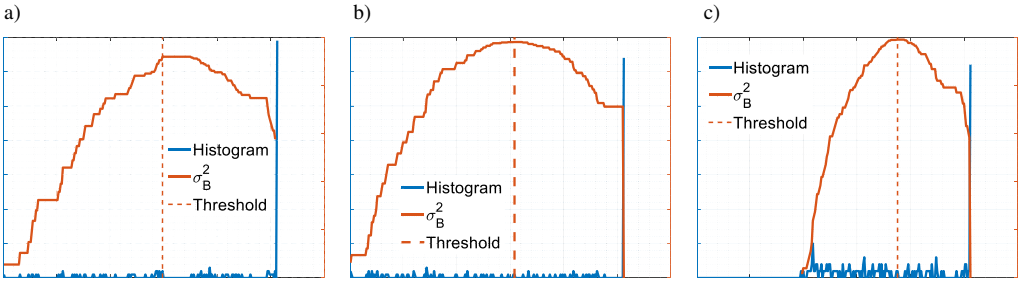

Fig. 31. Otsu's algorithm implementation for m (a), q (b) and S_ρ (c) maps, specimen A.

on specimens A and D. In these cases, the results relative to fit determination coefficient are not reported because of the segmentation give unintelligible results.

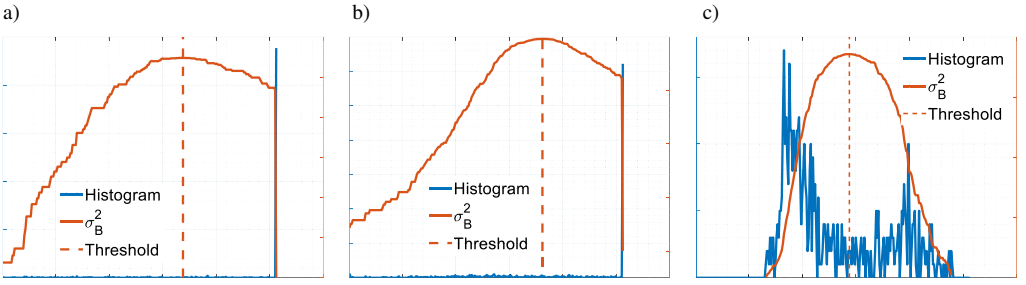

Fig. 32. Otsu's algorithm implementation for m (a), q (b) and S_ρ (c) maps, specimen D.

Figure 33 illustrates the Otsu's segmentation results for specimen E.

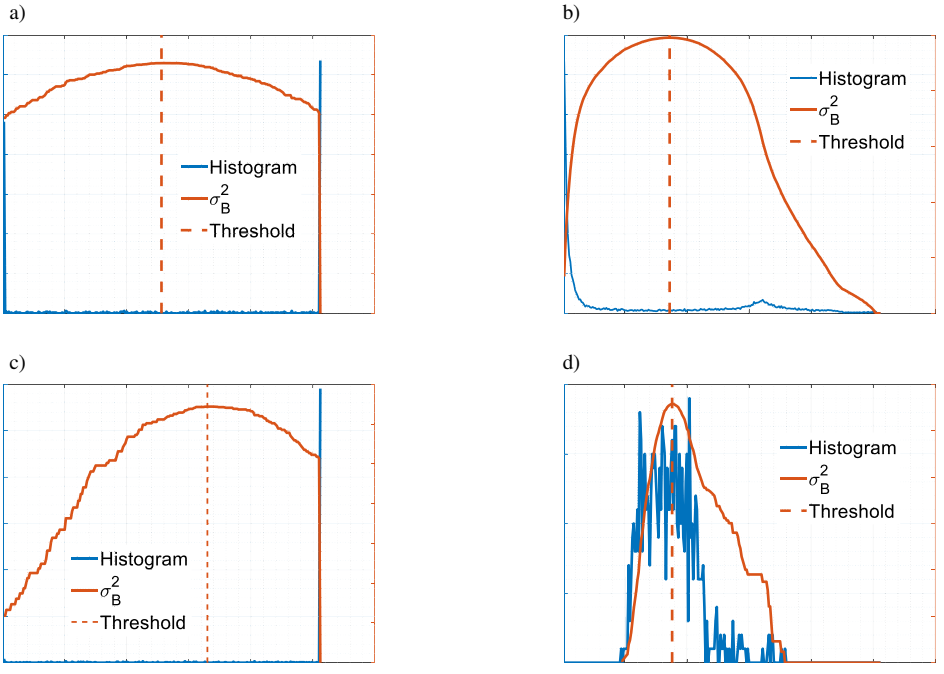

Fig. 33. Otsu's algorithm implementation for m (a), R^2 (b), q (c) and S_ρ (d) maps, specimen E.

Figure 34a, Fig. 34b and Fig. 34c, show, respectively, the damage classifier maps for fit slope, fit intercept and fit standard deviation attributable to one test performed on specimen A.

Figure 35a, Fig. 35b and Fig. 35c illustrate the same damage classifier maps for specimen D. It is worth pointing out that in both cases, the map pertaining to the fit determination coefficient is omitted because of poor quality segmentation.

Figure 36a, Fig. 36b and Fig. 36c show, respectively, the damage classifier maps for fit slope, determination coefficient, fit intercept and fit standard deviation for specimen E.

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

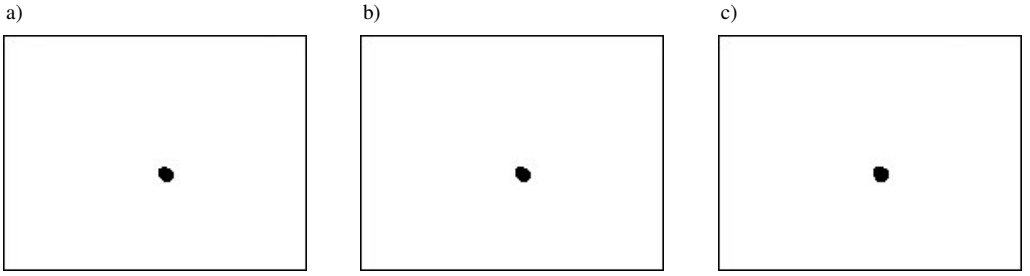

Fig. 34. Damage classifier maps m (a), q (b) and S_ρ (c). Specimen A.

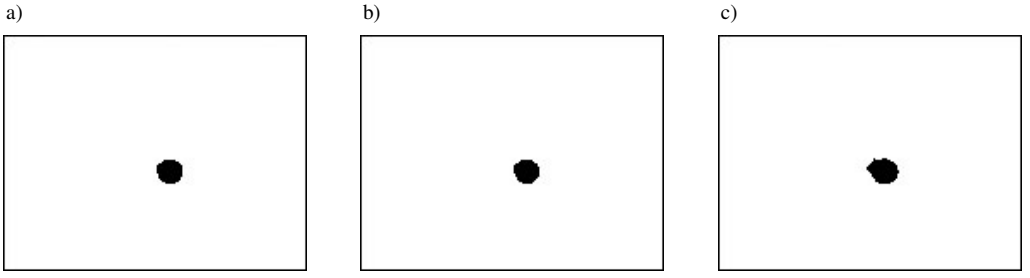

Fig. 35. Damage classifier maps m (a), q (b) and S_ρ (c). Specimen D.

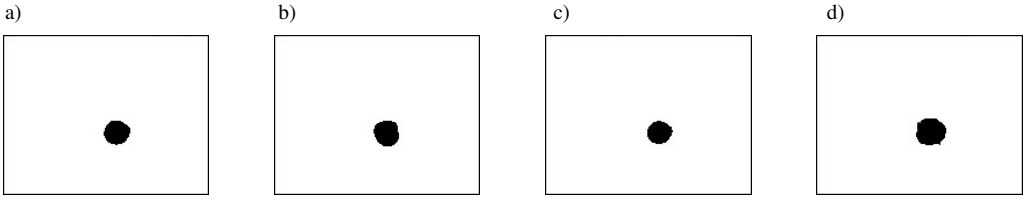

Fig. 36. Damage classifier maps m (a), R^2 (b), q (c) and S_ρ (d). Specimen E.

4. Results

A set of five repetitions is available for each specimen. As pointed out in section 2.3, the repetitions are needed in order to assess the repeatability and the stability of the performance of the method. The outcome of the algorithm consists of a series of damage classifier maps, pertaining to each fit parameter. The graphic representation from Otsu's segmentation allows the computation of the number of pixels pertaining to the detected defect region. Therefore, the inner core, where the defect is located, can be quantitatively defined.

Table 2 lists the mean value and the extended uncertainty of the number of pixels pertaining to the defect area in specimen A.

Table 2. Results for specimen A, units in pixel.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Deviation S_ρ**
Mean Value	53.8	NC	49.4	49.8
Ext. Unc. 95%	2.9	NC	0.5	0.8

The determination coefficient results are omitted because of the poor segmentation quality. The fit intercept and standard deviation parameters are characterized by the lowest uncertainties among the five preformed replications.

Table 3 and Table 4 report the number of pixels of the defect area for specimen D and E respectively.

Table 3. Results for specimen D, units in pixel.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Deviation S_ρ**
Mean Value	111.6	NC	112.6	112.0
Ext. Unc. 95%	1.1	NC	0.9	1.2

Table 4. Results for specimen E, units in pixel.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Deviation S_ρ**
Mean Value	314.2	307.2	314.8	314.6
Ext. Unc. 95%	1.2	1.6	0.4	0.5

It is worth pointing out that an increase of the flaw main size ensures a more accurate and precise detection. In addition, as the defect diameter increases, the mean values resulting from the fit slope, intercept and standard deviation maps get quite similar. Only the fit determination coefficient R^2 does not fulfil such behaviour.

4.1. Defect shape metrics

An accurate and appropriate characterization of the detected flaws plays an important role in the assessment of the detectability effectiveness granted by the proposed technique. The ability to bring out pixels (pertaining to presumptive defects) resembling connected areas having recognizable shapes is essential. This subsection deals with a series of considerations aimed at defining the detected defect by a morphological point of view. Therefore, some morphology indices are introduced in Table 5.

Table 5. Shape metrics definitions.

Index	**Definition**	
Circularity	$C = 4\pi \dfrac{A}{P^2}$	(1)
Solidity	$\Psi = \dfrac{A}{CA}$	(2)
Eccentricity	$\Xi = \dfrac{df}{MA}$	(3)
Aspect Ratio	$AR = \dfrac{NA}{MA}$	(4)
Roundness	$\Omega = 4\dfrac{A}{\pi MA^2}$	(5)

The circularity C mathematically describes the degree of similarity of the detected pixel ensemble to a perfect circle. In (1), the term A is the number of pixels and P the perimeter,

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

calculated as the distance between each adjoining pair of pixels around the border of the region. A unitary value is representative of a perfect circle. As C approaches 0, the shape is less circular.

The solidity Ψ describes how much a shape is convex or concave. In (2), the denominator CA is the homologous area enclosed by a convex hull. A unitary solidity indicates a perfect convex shape.

Equation (3) introduces eccentricity Ξ and it is computed as the ratio between the distance between the foci (df) of an ellipse and its major axis length (MA). The value is between 0 and 1, (0 and 1 are degenerate cases). An ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment.

A similar metric indicating the extent to which a shape can be approximated to an ellipse is the aspect ratio, AR, introduced by (4). It is the ratio between the minor axis length (NA) and the major axis length.

Equation (5) introduces the roundness Ω. Roundness is similar to circularity but it is more robust towards possible irregular borders along the perimeter as it involves the major axis length instead of the perimeter.

Figure 37 shows the fit intercept map for a test performed on specimen E. In particular, the defect pixel ensemble (in grey), the convex hull and the perimeter are shown. Figure 38 illustrates, for the same case, the ellipse circumscribing the defect, along with its major and minor axes.

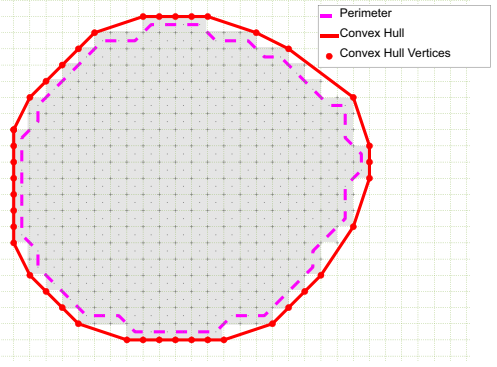

Fig. 37. Detected defect shape descriptors: blob's perimeter and solidity (intercept fit map, spec. E, test #1).

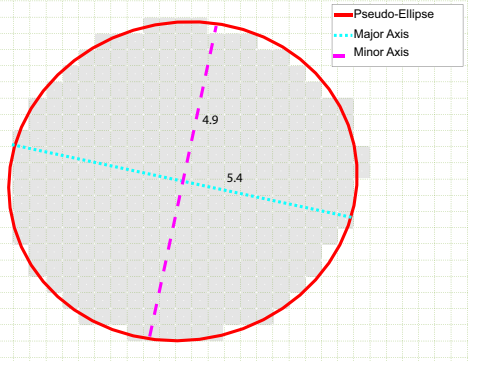

Fig. 38. Detected shape descriptors (intercept fit map, spec. E, test #1): aspect ratio, major and minor axes, units in mm.

Table 6, Table 7 and Table 8 introduce the mean values (deriving from the five replications) of the damage classifiers for specimens A, D and E.

Table 6. Shape metrics (average values) of the segmented defect area, specimen A.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
C	0.88	NC	0.98	0.88
Ψ	0.95	NC	0.98	0.97
Ξ	0.53	NC	0.35	0.47
AR	0.79	NC	0.88	0.85
Ω	0.77	NC	0.87	0.88

Table 7. Shape metrics (average values) of the segmented defect area, specimen D.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
C	0.84	NC	0.92	0.86
Ψ	0.97	NC	0.99	0.95
Ξ	0.44	NC	0.31	0.37
AR	0.89	NC	0.95	0.93
Ω	0.89	NC	0.95	0.90

Table 8. Shape metrics (average values) of the segmented defect area, specimen E.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
C	0.91	0.87	0.97	0.91
Ψ	0.96	0.98	1.00	0.91
Ξ	0.44	0.45	0.17	0.39
AR	0.98	0.89	0.99	0.92
Ω	0.89	0.88	0.93	0.89

The segmentation of the image maps of the introduced fit indices leads to the detection of defect areas (for the three tested specimens), which can be considered circular with excellent approximation. Also in this case, the results relative to the fit determination coefficient for specimens A and D are omitted because of the poor segmentation results. For all the other cases, the circularity, eccentricity, aspect ratio and roundness confirm the effective defect shape obtainable from m, q and S_ρ. In addition, the solidity values and, therefore, the deviation of the detected main area from the homologous convex one, indicate shapes characterized by highly regular borders.

4.2. Defect size calculation

This sub-section presents the results for the calculation of the detected defect size. Once the segmented defects have been obtained, the diameter is estimated.

The authors introduce the Feret diameter for the evaluation of defect main size, [28]. For each segmented image pertaining to a defined damage classifier, many Feret diameters, computed at different orientations, are considered. Such an approach allows the estimation of the *directional uncertainty*, which considers the imperfect circularity of the detected shapes. In addition, the performed repetitions for each specimen, allow the estimation of the *repeatability uncertainty*.

Therefore, for each damage classifier, the diameter overall uncertainty takes into account the standard deviation of the Feret diameters (as the considered orientation changes) within a single repetition and the standard deviation of the Feret diameter mean values among the repetitions. The first contribution is referred to as *within standard deviation* and the second contribution is referred to as *between standard deviation*.

Figure 39 shows the maximum and the minimum Feret diameters of detected defect from the fit intercept map pertaining to a test on specimen A. Figure 40 and Fig. 41 illustrate the same information for specimens D and E.

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

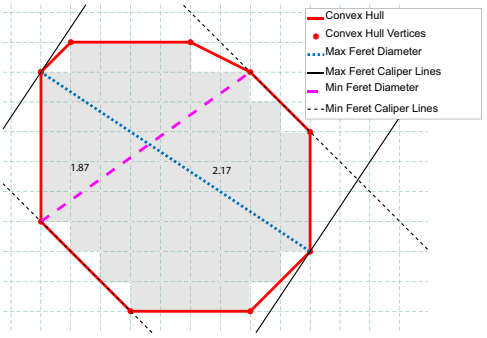

Fig. 39. Feret diameters computed from q map (specimen A, 1st repetition).

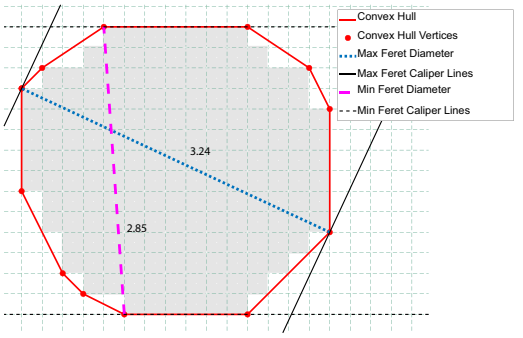

Fig. 40. Feret diameters computed from q map (specimen D, 1st repetition).

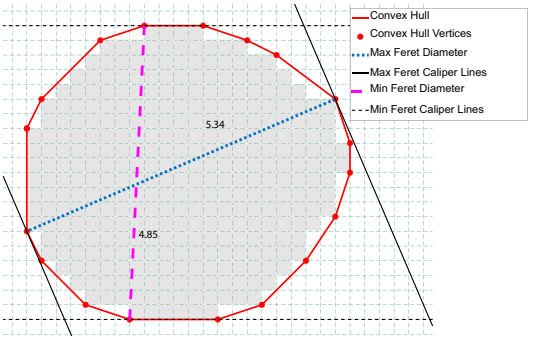

Fig. 41. Feret diameters computed from q map (specimen E, 1st repetition).

The defect mean diameter is estimated as follows:

$$\left\langle \bar{d} \right\rangle = \frac{\sum_{m=1}^{5} N_m \bar{d}_m}{\sum_{m=1}^{5} N_m} . \tag{6}$$

In (6), the subscript m refers to the m-th repetition, N_m the number of estimated Feret diameters for a single defect image and $\bar{d}_m$ the mean Feret diameter for the m-th repetition.

The standard uncertainty contribution due to the within standard deviation is given by the following equation:

$$u_w = \frac{\sqrt{\dfrac{\sum_{m=1}^{5} \nu_m S_m^2}{\sum_{m=1}^{5} \nu_m}}}{\sqrt{\sum_{m=1}^{5} N_m}} . \tag{7}$$

In (7) the term ν_m is the degree of freedom number of the m-th repetition and S_m^2 is the Feret diameter variance of the m-th repetition. The standard uncertainty related to the between standard

deviation is as follows:

$$u_b = \frac{\sqrt{\frac{1}{M-1}\sum_{m=1}^{5}\left(\bar{d}_m - \langle\bar{d}\rangle\right)^2}}{\sqrt{M}}. \tag{8}$$

In (8), M is the number of repetitions. The overall standard uncertainty is computed by combining u_w and u_b according to the *Root Sum of Squares* (RSS) rule. Therefore, it is extended to 95% level of confidence according to the Welch–Satterthwaite formulation, [23]. Table 9, Table 10 and Table 11 list the final results for specimens A, D and E, respectively.

Table 9. Defect mean diameter and extended uncertainty (95%) for specimen A, units in mm.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
Mean Diameter	2.07	NC	1.99	1.98
Ext. Uncertainty	0.17	NC	0.16	0.16

Table 10. Defect mean diameter and extended uncertainty (95%) for specimen D, units in mm.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
Mean Diameter	2.98	NC	2.99	2.98
Ext. Uncertainty	0.24	NC	0.09	0.12

Table 11. Defect mean diameter and extended uncertainty (95%) for specimen E, units in mm.

	Fit Slope m	**Det. Coef R^2**	**Fit Intercept q**	**Fit Standard Dev. S_ρ**
Mean Diameter	5.03	4.94	5.00	5.00
Ext. Uncertainty	0.24	0.98	0.02	0.12

Also in this case, the results for the fit determination coefficient are not reported because of very poor segmentation quality of R^2 maps for specimens A and D. The fit zero-intercept damage classifiers are characterized by the best accuracy and precision. Therefore, an analysis involving only the uncertainties of the segmentation results, suggests that the q parameter is the best performing one.

4.3. Segmentation goodness evaluation metrics

In this sub-section, the authors suggest an alternative and more interesting approach to the evaluation of the quality of information carried by each introduced damage classifier. Such approach is based on the evaluation of the effectiveness of the segmentation images of each damage classifier.

In this paper, the authors define a segmentation evaluation method which is objective (*i.e.* quantitative) and unsupervised. The need for an unsupervised evaluation method arises from the lack of a reference image (the so-called ground image). The method is based on the concept of information entropy, [29–31].

The evaluation parameter having a good segmentation maximizes the uniformity of pixels within each segmented region and minimizes the uniformity across the pixels of contiguous

Metrol. Meas. Syst., Vol. 27 (2020), No. 2, pp. 219–242
DOI: 10.24425/mms.2020.132771

regions. Therefore, since the entropy provides a measure of the degree of disorder within a region, it is an excellent descriptor of overall segmentation quality.

Given a segmented image, the entropy of a generic j-th region is:

$$H_j \triangleq -\sum_m P(x_i) \log P(x_i) \,. \tag{9}$$

In (9), the term $P(x_i)$ is the probability of state x_i, *i.e.* the probability that a pixel has a value of x_i. Once the entropy of each segmented region and the entropy of the overall images are calculated, the *Segmentation Entropy* (SE) is computed as follows:

$$SE \triangleq \frac{\sum_{j=1}^{N} H_j - H_f}{H_f} \,. \tag{10}$$

In (10), H_f is the entropy of the whole image and the term N is the total number of segmented regions two in this case study). Therefore, the so-defined *Segmentation Entropy* gives an insight into the degree of uniformity (for relatively low SE) or non-uniformity (for relatively large SE) of the luminance value of pixels grouped in each segmented region.

Table 12 lists the SE average values of the damage classifiers for each tested specimen. In all the cases, the fit intercept parameter is characterized by the lowest values.

Table 12. Segmentation entropy of the damage classifiers.

Specimen	Fit Slope m	Det. Coef R^2	Fit Intercept q	Fit Standard Dev. S_ρ
A	2.39	NC	0.092	0.13
D	0.11	NC	0.064	0.11
E	0.14	1.4	0.05	0.12

A comprehensive analysis of computed mean diameter along with its uncertainty, shape factors and segmentation entropy seem to suggest that for such a kind of test specimen, characterized by a simulated topcoat detachment defect, the zero-intercept value of the fit of the log-log temperature time history (during the cooling stage) is the best performing damage classifier.

5. Conclusions

This paper proposes a novel algorithm for the detection and quantification of subsurface defects representative of possible partial detachments of TBCs from the underlying metallic substrate.

The algorithm is based on a series of operations to be implemented on thermograms that are representative of the time evolution of the surface temperature of the test specimen under analysis. The thermal excitation of the specimen is carried out by means the Long Pulsed Thermography technique, which is particularly effective for TBC's defect detection.

The method develops and enhances the polynomial fit of the logarithmic time history of the surface temperature during its cooling stage. The enhancements of the polynomial fit technique, are aimed at improving the accuracy of the detection by means of the introduction of certain damage classifiers, obtainable from the fit analysis and related to the fit slope, determination coefficient, intercept and standard deviation. Additionally, the algorithm consists of a series of

steps, to be implemented automatically without any *a priori* knowledge about the distribution of pixels between the background and the foreground regions and the distribution of pixels between defect and sound regions within the identified foreground. In fact, the main feature of the algorithm is based on a series of automatic image segmentations, which have the main purpose of identifying the ROI (and then, the effective ensemble of specimen pixels to which to implement the computations) and the defect pixels inferable from the damage classifiers' maps.

In order to analyse the algorithm performance, the authors developed an experimental setup, comprising the IR instrumentation and the data acquisition and processing system. Several specimens have been tested, based on the same component carrying a defect of three different sizes.

Additionally, the authors developed a measurement plan for the assessment of the algorithm repeatability, consisting of a set of five repetitions for each specimen, with a detailed characterization of the involved uncertainties.

The classifiers deducted from each test are organized in images. These are conveniently post-processed by means of an automatic segmentation. The set of connected defect pixels gives a visual representation of the defect shape and extent. In addition, the segmentation allows for the direct evaluation of the defect main size along with its estimated uncertainty.

In order to further validate the introduced technique, the authors introduce a series of shape descriptors in such a way that the morphological coherence of the detected defect areas can be assessed.

Furthermore, the authors suggest an additional indicator, for the identification of the best performing damage classifier in the presented case study. The indicator is based on the information entropy within each segmented region and across the segmented regions deduced from the damage classifiers maps.

References

[1] Meola, C., Carlomagno, G.M. (2004). Recent advances in the use of infrared thermography. *Measurement Science and Technology*, 15(9), 27–58.

[2] Vavilov, V., Burleigh, D.D. (2015). Review of Pulsed Thermal NDT: Physical principles, Theory and data processing. *NDT & E International*, 73, 28–52.

[3] Balageas, D.L., Roche, J.M. (2014). Common tools for quantitative time-resolved pulse and step-heating thermography – part I: theoretical basis. *Quantitative InfraRed Thermography Journal*, 11(1), 43–56.

[4] De Capua, C., Morello, R., Jablonski, I. (2018). Active and eddy current pulsed thermography to detect surface crack and defect in historical and archaeological discoveries. *Measurement*, 116, 676–684.

[5] Zheng, K., Chang, Y.S., Yao, Y. (2015). Defect detection in CFRP structures using pulsed thermographic data enhanced by penalized least squares methods. *Composites Part B: Engineering*, 79, 351–358.

[6] Shepard, S.M. (2001). Advances in pulsed thermography. *Proc. of SPIE Thermosense XXIII*, 4360, 511–515.

[7] Shepard, S.M., Lhota, J.R., Rubadeux, B.A., Ahmed, T., Wang, T. (2002). Enhancement and reconstruction of thermographic NDT data. *Proc. SPIE Thermosense XXIV*, 4710, 531–535.

[8] Maldague, X., Galmiche, F., Ziadi, A. (2002). Advances in pulsed phase thermography. *Infrared Physics & Technology,* 43(3-5), 175–181.

[9] Ibarra-Castanedo, C., Maldague, X. (2004). Pulsed phase thermography reviewed. *Quantitative InfraRed Thermography Journal*, 1(1), 47–70.

[10] Rajic, N. (2002). Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures. *Composite Structures*, 58(4), 521–528.

[11] Palumbo, D., Galietti, U. (2016). Damage investigation in composite materials by means of new thermal data processing procedures: damage investigation with stimulated thermography. *Strain*, 52(4), 276–285.

[12] Roche, J.M., Balageas, D.L. (2015). Common tools for quantitative pulse and step-heating thermography – part II: experimental investigation. *Quantitative InfraRed Thermography Journal*, 12(1), 1–23.

[13] Vavilov, V. (1980). Infrared non-destructive testing of bonded structures: aspects of theory and practice. *British Journal of Non-destructive Testing*, 22(4), 175–183.

[14] Vavilov, V., Taylor, R. (1982). Theoretical and practical aspects of the thermal non-destructive testing of bonded structures. *Research Techniques in Nondestructive Testing*, 5, 239–279.

[15] Almond, D.P., Delpech, P., Beheshtey, M.H., Wen, P. (1996). Quantitative determination of impact damage and other defects in carbon fiber composites by transient thermography. *Proc. of SPIE Nondestructive Evaluation of Materials and Composites*, 2944, 256–264.

[16] Grys, S., Minkina, W., Vokorokos, L. (2015). Automated characterisation of subsurface defects by active IR thermographic testing – Discussion of step heating duration and defect depth determination. *Infrared Physics & Technology*, 68, 84–91.

[17] Palumbo, D., Tamborrino R., Galietti, U. (2017). Coating defect evaluation based on stimulated thermography. *Proc.of SPIE Thermosense: Thermal Infrared Applications XXXIX*, 10214X.

[18] Dinardo, G., Fabbiano, L., Tamborrino, R., Vacca, G. (2019). Automatic defect detection from thermographic non-destructive testing. *Journal of Physics: Conference Series*, 1249(1), 012010.

[19] Preetha, M.M.S.J., Suresh, L.P., Bosco, M.J. (2012). Image segmentation using seeded region growing. *Proc. of International Conference on Computing, Electronics and Electrical Technologies (ICCEET)*, 576–583.

[20] Huang, M., Yu, W., Zhu, D. (2012). An improved image segmentation algorithm based on the Otsu method. *Proc. of 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel Distributed Computing*, 135–139.

[21] Feng, Q., Gao, B., Lu, P., Woo, W.L., Yang, Y., Fan, Y., Qiu, X., Gu, L. (2018). Automatic seeded region growing for thermography debonding detection of CFRP. *NDT & E International*, 99, 36–49.

[22] Grys, S. (2018). Determining the dimension of subsurface defects by active infrared thermography – experimental research. *Journal of Sensors and Sensor Systems*, 7(1), 153–160.

[23] JCGM 100:2008 (2008) Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement. Joint Committee for Guides in Metrology.

[24] Otsu, N. (1979). A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66.

[25] Yang, X., Shen, X., Long, J., Chen, H. (2012). An improved median-based Otsu image thresholding algorithm. *AASRI Procedia*, 3, 468–473.

[26] Holland, P.W., Welsch, R.E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics – Theory and Methods*, 6(9), 813–827.

[27] Hinich, M.J., Talwar, P. P. (1975). A simple method for robust regression. *Journal of the American Statistical Association*, 70(349), 113–119.

[28] Dražić, S., Sladoje,N., Lindblad, J. (2016). Estimation of Feret's diameter from pixel coverage representation of a shape. *Pattern Recognition Letters*, 80, 37–45.

[29] Borsotti, M., Campadelli, P., Schettini, R. (1998). Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19, 741–747.

[30] Gao, H., Tang, Y., Jing, L., Li, H., Ding, H. (2017). A novel unsupervised segmentation quality evaluation method for remote sensing images. *Sensors MDPI*, 17(10), 2427.

[31] Hao, J., Shen, Y., Xu, H., Zou, J. (2009). A region entropy based objective evaluation method for image segmentation. *Proc. of IEEE Instrumentation and Measurement Technology Conference*, 373–377.