

Peer Reviewed Paper **openaccess**

Restoration of defaced serial numbers using lock-in infrared thermography (Part I)

Ikwulono Unobe,^a Lisa Lau,^b John Kalivas,^{c,*} Rene Rodriguez^{d,*} and Andrew Sorensen^e

^aCollege of Science and Engineering, Idaho State University, 921 S. 8th Avenue, Stop 8023, Pocatello, ID, 83209-8023, USA.
E-mail: unobikwu@isu.edu, <https://orcid.org/0000-0001-7145-5027>

^bCollege of Science and Engineering, Idaho State University, 921 S. 8th Avenue, Stop 8023, Pocatello, ID, 83209-8023, USA.
E-mail: laulisa@isu.edu, <https://orcid.org/0000-0001-8872-3390>

^cCollege of Science and Engineering, Idaho State University, 921 S. 8th Avenue, Stop 8023, Pocatello, ID, 83209-8023, USA.
E-mail: kalijohn@isu.edu, <https://orcid.org/0000-0001-7056-976X>

^dCollege of Science and Engineering, Idaho State University, 921 S. 8th Avenue, Stop 8023, Pocatello, ID, 83209-8023, USA.
E-mail: rodrrene@isu.edu, <https://orcid.org/0000-0003-4878-0299>

^eDepartment of Civil and Environmental Engineering, Utah State University, 4110 Old Main Hill, Logan, Utah, 84322-4110, USA.
E-mail: andrew.sorensen@usu.edu, <https://orcid.org/0000-0001-9998-2021>

Infrared thermal imaging is an evolving approach useful in non-destructive evaluation of materials for industrial and research purposes. This study investigates the use of this method in combination with multivariate data analysis as an alternative to chemical etching; a destructive method currently used to recover defaced serial numbers stamped in metal. This process involves several unique aspects, each of which works to overcome some pertinent challenges associated with the recovery of defaced serial numbers. Infrared thermal imaging of metal surfaces provides thermal images sensitive to local differences in thermal conductivity of regions of plastic strain existing below a stamped number. These strains are created from stamping pressures distorting the atomic crystalline structure of the metal and extend to depths beneath the stamped number. These thermal differences are quite small and thus not readily visible from the raw thermal images of an irregular surface created by removing the stamped numbers. As such, further enhancement is usually needed to identify the subtle variations. The multivariate data analysis method, principal component analysis, is used to enhance these subtle variations and aid the recovery of the serial numbers. Multiple similarity measures are utilised to match recovered numbers to several numerical libraries, followed by application of various fusion rules to achieve consensus identification.

Keywords: serial number restoration, lock-in infrared thermography, principal component analysis, Zernike moments, similarity measure

Introduction

Stamp marks, which provide a means of unique identification for firearms and automobiles, are regularly defaced for criminal activities. The stamping of a serial number onto metals causes a permanent change in

shape, primarily due to the inability of regions of crystalline arrangement within localised grains to resist the induced stress of stamping, leading to an alteration of the structure and, in extension, the interlocking

Correspondence

John Kalivas (kalijohn@isu.edu) and Rene Rodriguez (rodrrene@isu.edu)

Received: 8 April 2019

Revised: 20 October 2019

Accepted: 1 November 2019

Publication: 18 November 2019

doi: 10.1255/jsi.2019.a19

ISSN: 2040-4565

Citation

I. Unobe, L. Lau, J. Kalivas, R. Rodriguez and A. Sorensen, "Restoration of defaced serial numbers using lock-in infrared thermography (Part I)", *J. Spectral Imaging* 8, a19 (2019). <https://doi.org/10.1255/jsi.2019.a19>

© 2019 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper in this journal is given, the use is not for commercial purposes and the paper is not changed in any way.



grain boundaries. This resulting deformation extends to some depth below the stamped mark and is known as the zone of plastic strain.^{1–6} Although the shape of this underlying region will be complex depending on various factors, it can be considered as a sphere to make an estimate of the depth with its radius equating to the depth of the zone of plastic strain.⁶ This area has physical and chemical properties that differ from those of non-stamped metal due to the changes in its microstructure. Approaches used for the recovery of obliterated serial numbers leverage such differences to recover the numbers.^{1–6}

A widely used approach is chemical etching. This process utilises the change in chemical potential that makes the defaced area more reactive to acids, allowing recovery of serial numbers.^{1–4} The method, however, is a highly controlled and destructive recovery process that requires delicacy and some expertise in applying and removing the etchant. Additionally, because it is a destructive method, the test can only be run once and the test specimen is permanently altered. This shortcoming is the impetus for developing a non-destructive, reproducible method of defaced serial number identification. A non-destructive method that has been tested for this is magnetic particle inspection.³ However, this method has limited applicability (ferromagnetic alloys and magnetic metals). Other non-destructive methods being tested, including scanning acoustic microscopy⁷ and spectral imaging,⁸ are experimental and their effectiveness is yet to be established.

One non-destructive technique that has proven useful in the detection of defects in metals and holds promise for use in the identification of defaced serial numbers, is infrared thermal imaging.^{9–13} Thermal imaging captures the infrared radiation from the surface of an object and converts this to visible images scaled by temperature, allowing visualisation of the temperature evolution over time. Thermal imaging is analogous to spectral imaging as both methods seek to obtain images of objects at changing parameters; temperature/time for thermal imaging and intensity/wavelength for spectral imaging. Thermal imaging seeks to locate and characterise flaws in an object by measuring, under controlled conditions, their effect on heat flow through the object by observation of the propagation of applied thermal energy.¹⁰ Local regions of plastic strain can be detected as the temperature gradient therein will differ from the rest of the surface due to the local change in thermal conductivity. Infrared

thermography involves the use of an infrared camera and heating apparatus to capture this effect. However, the data acquired through thermographic imaging techniques can be noisy due to undesired signals from several factors including unevenly heated surfaces, radiation from the heated surface and local emissivity variations. One way to mitigate these effects is by using the lock-in infrared thermography technique.^{9,14–17}

Lock-in thermography (LIT) involves inducing sinusoidal energy waves into the surface of a sample at a lock-in frequency, and collecting the infrared thermal images of the sample surface over the entire period of the pulsed wave.^{14,15,17} Digitally processed according to the lock-in principle, the thermographic image sequence is analysed and compressed into an amplitude image and a signal phase image.^{14,15,17} The developed phase and amplitude images have some surface features suppressed, making them more robust than raw thermal images for analysing sub-surface features.^{14,15,17,18} This process, being sensitive to variations in thermal conductivity, is evaluated for recovering defaced serial numbers.

A possible pitfall of using LIT for serial number recovery is the sensitivity of the imaging apparatus to visibly capture what might be infinitesimal changes in temperature gradient across the surface. Another is the capture of other features (such as possible material inhomogeneities) within the thermal depth range.¹⁵ A conceivable way to overcome these challenges is using multivariate image analysis (MIA), a computational process that allows for the examination of collected images to help with detection and analysis of possible variations within the data. Successfully applied in diverse fields,^{19–24} a very popular form of MIA is principal component analysis (PCA); a technique used for identifying patterns in data and expressing them in a way that highlights the importance of elements within the data.²⁵ When used for images, the resulting principal components can be formed into score images that are of the same spatial dimensions as the original images and provide a visual representation of the information retained in respective principal components. This method has been successfully utilised in some thermal imaging applications for non-destructive testing to capture the thermal contrast evolution in time, enhancing thermal signatures of sub-surface defects and simultaneously reducing optical effects that could lead to false alarms.^{26,27} Using this method, it is conceivable that the unique thermal gradient of the zone of plastic strain across the phase and/or amplitude images will be

identified and clustered in score images to reproduce the defaced number.

Other methods for processing thermographic data include differentiated absolute contrast (DAC),^{28,29} thermographic signal reconstruction (TSR)^{30,31} and pulsed phase thermography (PPT).³² These are, however, limited in needing user intervention for slope calculation (DAC), dependence of the performance on polynomial order and insensitivity to subtle defects (TSR) and the effect of high frequency noise from raw data on the results (PPT).

However, even after recovering the removed numbers, it is imperative to conclusively identify the number without bias via an automated process. This identification can be achieved by comparing score images of recovered numbers to a library of unaltered digital number images to determine the best match based on similarity measures.^{33–35} Such comparisons between score images of defaced numbers and library images come with several challenges including image size differences, as well as possible translation, scale and rotation differences between the images being compared. These challenges can be overcome by comparing image features extracted via orthogonal moments, in lieu of the actual images. For this study, Zernike moments³⁶ are used to characterise the score images into vectors of a predetermined size. Invariant to rotation, Zernike moments have been proven to outperform other orthogonal moments including general orthogonal moments, Legendre moments and Jacobi Fourier moments as a global shape descriptor for object classification and retrieval.^{36–42} Other orthogonal moments could be used such as wavelet moments,⁴³ Krawtchouk moments,⁴⁴ Tchebichef moments⁴⁵ and pseudo Zernike moments,⁴² but these moments were not studied.

The smaller the resulting output from a similarity comparison, the closer the corresponding compared objects are within the characteristics captured by the similarity measure used. Several similarity measures are used in this study to compare the developed score images to several digitally developed libraries of number images to provide a holistic and unbiased comparison. As such, it is pertinent to combine the individual similarity values in a way that establishes an objective consensus to the comparison results. Data fusion, a technique used in combining individual sources of information into a single informative output with improved reliability and less ambiguity,^{46,47} is used to achieve this consensus. A high-level or decision level fusion approach is used in

this study as it involves the fusion of several parameters (similarity values), each of which can be used to make an independent identification,^{47,48} to yield a final inference.

To test the process, a stainless steel plate with several numbers stamped into it is used. Each number on the sample was machined off to successively lower depths and specifically tested to ascertain the validity of the entire process. Results are presented for this preliminary study.

Materials and methods

Samples

Stainless steel plates with numbers

The experimental process was carried out on a clean stainless steel surface with no numbers stamped into it, to serve as a control experiment. Another stainless steel plate (72 mm × 25 mm) stamped with several numbers to depths of approximately 0.2 mm was used to simulate the mechanically stamped serial numbers. This sample, obtained as a test piece from Precision Forensic Testing,⁴⁹ was originally 6.35 mm thick. A uniform thickness of 0.18 mm was mechanically shaved off the top surface, leaving the numbers barely visible. Each number except the first was then progressively shaved off to a depth 0.03 mm beyond the previous. The first number on the plate is left visible to serve as a control. Figure 1 shows the original sample. After the defacing, it is impossible to visually identify the numbers present before the material is removed. The sample was sanded using a 600 grit sand paper to polish off the surface and thoroughly cleaned with ethanol to remove any residue from the sanding process, ensuring a smooth surface as shown in Figure 2. The boxes show the areas around each defaced number isolated and individually analysed. Although the thermal images were initially collected by the infrared camera to cover areas over two numbers, the thermal image datasets were split to contain the areas over each number and analysed separately.

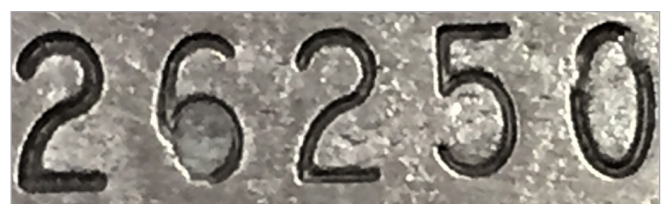


Figure 1. Stainless steel sample with stamped numbers.

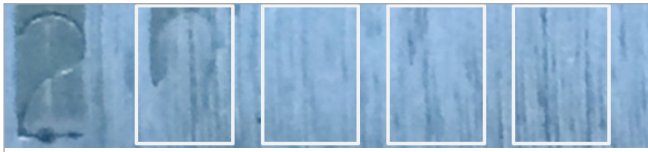


Figure 2. Stainless steel sample with numbers defaced (boxes show areas extracted and individually analysed for each number).

Sample preparation

The surface area around the defaced numbers was covered with a light coat of India ink ($\sim 500\mu\text{m}$) to reduce the possible adverse effects from surface emissivity on the quality of the captured images.

Lock-in infrared thermal imaging

Data collection

The experimental system consists of an infrared camera, a function generator and thermal energy source as shown in Figure 3. The primary thermal energy source, a 5 W Ar-ion cw laser, operating in all-lines mode, is chopped with a Uniblitz mechanical shutter to apply pulsed energy to the sample. The laser was set up to put out 3 W of thermal energy and losses through the top-hat filter which transforms the Gaussian shape of the beam into a uniform intensity beam as well as the prisms account for about half of the power, allowing approximately 1.5 W to the sample. This power level gave a good trade-off between inducing a temperature swing in the sample over the pulsing cycle, while avoiding overheating the sample and thus losing the ability to capture the small thermal gradients characteristic of the zone of plastic strain. A FLIR SC6700 infrared camera is used to collect thermal images of the sample surface and a Stanford Research Systems model DS340 15 MHz function generator synched with the camera is used to control the pulse rate of energy from the laser. A synchronising circuit is

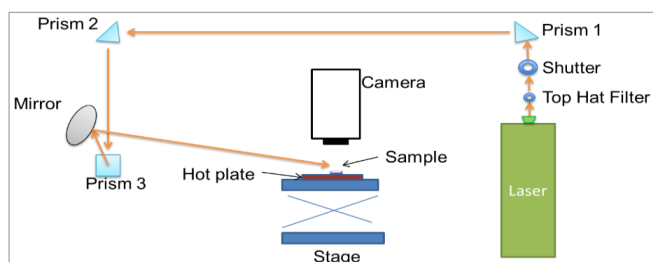


Figure 3. Experimental setup.

used to synchronise the laser pulses with the camera's imaging apparatus ensuring a consistent rate of capturing images over an entire pulse irrespective of the cycle time used. A digital hot plate is used to heat up the sample to an initial temperature of approximately 73°C as measured by a digital thermometer. This is done to ensure that the sample is at an even initial temperature and to enhance the contrast of the thermal gradient across the surface to improve contrast.¹⁰ The laser beam is also pulsed into the sample a few times prior to experimentation to further ensure even distribution of temperature through the sample. During the experiments, the sample on the hot plate is placed in the focus range of the camera and thermal energy is pulsed into the surface of the sample under observation using the laser at a defined modulated (lock-in) frequency of 0.05 Hz. The temperature distribution on the heated surface is monitored using the infrared camera with images over the pulsing cycle.

Images are simultaneously collected at a frame rate of 32 frames per pulse cycle. This process is carried out for several cycles. If the heat propagates irregularly through the plastic deformation areas below the stamped numbers, whereas propagation of the heat through non-deformed areas is more regular, then over the course of several heat pulses, differences in the thermal images between the irregular propagation in the deformed areas and the regular propagation in the non-deformed should be increased by including more pulses. The white noise in the thermal images associated with the regular propagation is essentially averaged out by including more pulses, and 15 pulse cycles were used for these studies. Figure 4 is a mean temperature-time plot across all pixels of the defaced area showing the pulsing of thermal energy through the sample over time. The values over each peak show the maximum temperature value in that cycle. Control of the heating is difficult and slight variations in the maximum temperature over the course of the 15 cycles is a typical occurrence. The effect of the pulse-to-pulse temperature variation is currently under further investigation.

For the sample used in this study, the experiment setup is adjusted with the camera field of view narrowed to capture areas over two of the removed numbers. This is done both to enhance the detail captured with more pixels across a single number, as well as to ensure an even spread of the pulsed thermal energy from the laser beam to avoid possible lateral thermal gradients. The

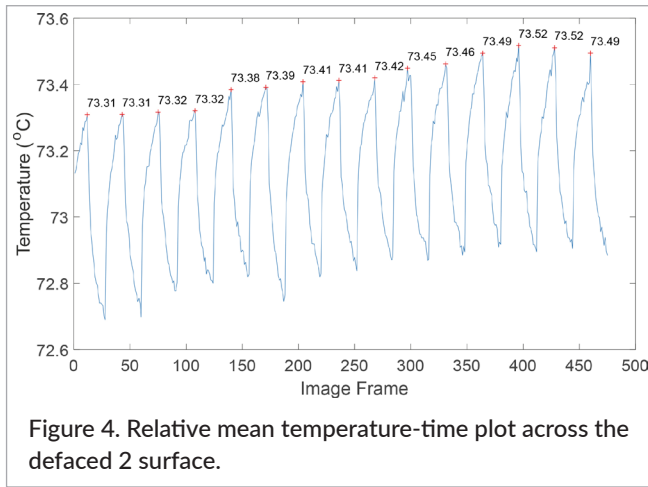


Figure 4. Relative mean temperature-time plot across the defaced 2 surface.

captured images are then digitally split to have data-sets that are focused on the areas around each removed number.

Lock-in analysis

The process used to develop phase and amplitude images using LIT is as described in Reference 14. This process involves applying sinusoidal thermal energy waves into the surface of a sample at a defined frequency (lock-in frequency), and collecting the infrared thermal images of the sample surface over the entire period of the pulsed wave. These thermal images are processed via a two-channel image correlation process to develop an amplitude image A , and a signal phase image ϕ using Equations 1–4:¹⁴

$$S^{0^\circ} = \frac{1}{k} \sum_{i=1}^k 2 * \sin(t) * F(i) \quad (1)$$

$$S^{-90^\circ} = \frac{1}{k} \sum_{i=1}^k -2 * \cos(t) * F(i) \quad (2)$$

$$A = \sqrt{(S^{0^\circ})^2 + (S^{-90^\circ})^2} \quad (3)$$

$$\phi = \tan^{-1} \left(\frac{-S^{-90^\circ}}{S^{0^\circ}} \right) \quad (4)$$

where k is the number of frames per lock-in period, $F(i)$ is the i^{th} thermal image collected by an infrared camera, $t = [2\pi(i - 1)] / k$, $\sin(t)$ and $\cos(t)$ are weighing factors and S^{0° and S^{90° are the in-phase and quadrature signals, respectively.

Alternatively, the phase image can be developed by inverting the in-phase and quadrature signals as shown in Equation 5.¹⁶

$$S^{0^\circ/-90^\circ} = \frac{S^{0^\circ}}{S^{-90^\circ}} \quad (5)$$

The image formed using this inverted equation holds some advantages over regular phase images including improved spatial resolution while maintaining the emissivity corrected property of phase images.¹⁶ As in cases like this study when it is suspected that the thermal gradient differences are very small, using this image might improve contrast helping to recover the number. Consequently, this method is used in this study to develop the phase images. Also, to improve contrast in the amplitude images, the fourth power was used instead of the square of the in-phase and quadrature signals. This reduces the effects of the emissivity multiplier for each signal and improves the output image.

A major drawback of LIT is the presence of blind frequencies (i.e. pulsing frequencies at which the contrast between clean areas and defective areas is minimal). Defect detection in phase images developed at these frequencies is usually difficult and thus, such frequencies need to be avoided. To address this shortcoming, phase difference plots are developed^{9,18} to determine the best test frequencies for collecting data. This is done by calculating phase images at several frequencies and determining the phase differences between defaced and non-defaced areas. The frequencies at which these differences are maximised are selected to be used for the experimentation. An example of this analysis is presented in the Results and discussion.

Each pulse cycle is used to develop one amplitude and one phase image in which the contrast change across the surface due to the variation in thermal gradient from the defects is improved. Fifteen of these images are developed, each at an increasing temperature from the input energy pulse. The variation in degree of phase shift between defaced areas and clean non-defaced areas captured in the amplitude and phase images at different cycles due to changes in the thermal gradient may be small and thus difficult to observe. As such, PCA of the images should bring more focus on the propagation differences, reproducing it in a single score image independent of others.

PCA is carried out independently on the matrix of the phase images and that of the amplitude images and respective score images are visually inspected to determine if phase or amplitude score images best reproduces the defaced number.

Principal component analysis

The input matrix for PCA consists of a collection of filtered phase or amplitude images (referred to henceforth as input images) processed from the respective thermal image sequences captured using an infrared camera. The input images were filtered using a moving average filter to enhance the contrast between pixels by averaging the value of pixels in a rectangle of specified length and width (12×12 pixels for this study) around a particular pixel and replacing its intensity value with the average. This window size was chosen as it represents about half of the pixel number that spans the width of a number stamp cut for the sample used in this study. For the edges of the images, the window is taken with the entire length and/or width of the rectangle on the inward side. This window increases as the number of pixels closest to the edge increases until the full length can be used on all sides. The input images for the area around the defaced 0 were further preprocessed using standard normal variate⁵⁰ to remove some light scatter effects caused by some surface unevenness still existing even after polishing that were observed on visual inspection of the input images. This was not done for the images for the other defaced numbers as there was no observed scatter in their input images.

The PC²⁵ carried out on this input matrix is briefly defined:

- Each input image is unfolded (row-wise or column-wise) and concatenated as a single column of a new representative matrix $\mathbf{X}$ as shown in Figure 5.
- The covariance matrix $\mathbf{C}$ is determined for the matrix.

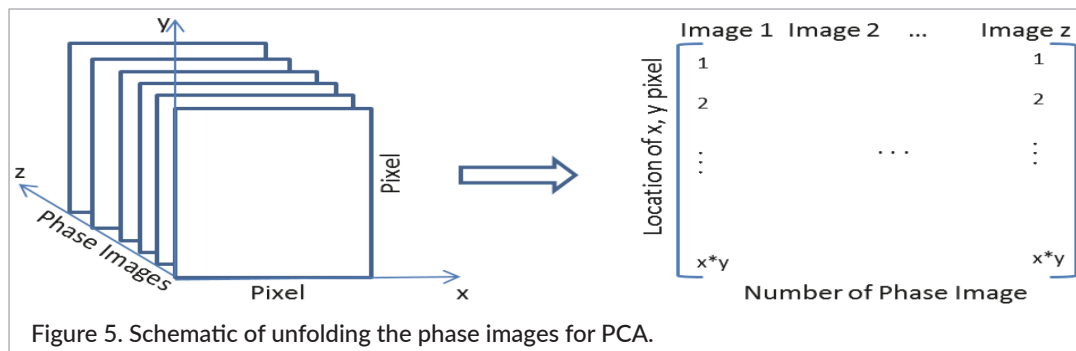
- The eigenvectors ($\mathbf{V}$) and their corresponding eigenvalues (Σ) are determined for the covariance matrix.
- Each eigenvector ($\mathbf{v}$) multiplies the representative matrix $\mathbf{X}$ to develop a score vector which captures a percentage of the variance within the original images given in its corresponding eigenvalue.
- Each principal component score is then folded into a score image to visualise the information captured as a linear combination across the input images.

In projecting the input images into orthogonal directions, each principal component captures variation across the phase or amplitude images. Among these variations is the change in the rate of heat propagation between the defaced areas and clean areas, which appear as subtle signatures in the input images. These variations, largely identical in defaced areas, are mainly captured by a single principal component, independent of other linear variations across the input images.

Although identical in the process of developing, there are some differences in the amplitude and phase images developed from thermal images collected over succeeding periods. The transient flow of thermal energy over time through the sample will cause changes in the phase shift in different cyclic periods until thermal equilibrium is achieved. Accordingly, each developed amplitude and phase image will be unique in the degree of phase shift between defaced and clean areas. As such, several amplitude and phase images covering different periods will capture this changing phenomenon and PCA is expected to leverage these changes to reproduce the defaced number from the zone of plastic strain.

Feature extraction

Binarisation of score and numerical library images
The reference library images were created by first typing out the numbers in Microsoft Word to the different fonts



used. These were then expanded so each number occupied an entire page (11×8.5 inches). The MS Word document was then converted to individual bitmap images and imported into MATLAB to create digital image libraries of sizes 200×150 pixels. Figure 6 shows an example of the numbers for the Times New Roman font library images.

The PCA score images are initially preprocessed into binary black-and-white images to match the black-and-white library images. The binarisation process involves selecting an optimal threshold of pixel intensities to best separate background from object pixels. This threshold is selected for each individual image iteratively by using an initial assumption of the four-pixel mean intensity from each corner of respective score images. The mean value of pixels above this initial threshold (assumed to represent objects within the image) and those equal to or below the threshold (assumed to be represent the background) are separately calculated and then averaged together. This average value is used as a new threshold. This iterative process continues until there is no change in the determined threshold. The corresponding score image is then binarised with respect to the threshold, with 1 for above the threshold as white and 0 or below being black.

Zernike moments

Computation of Zernike moments for binarised score and numerical library images involves four steps: preprocessing the image to make it scale and translation invariant, computation of radial polynomials, computation of Zernike basis functions (Zernike polynomials) and finally computation of Zernike moments by projecting an image onto the Zernike basis functions.^{36,39–42}

Preprocessing the image to make it scale and translation invariant involves normalising the image using its regular moments as described in Reference 42. Scale invariance is achieved by altering the object size such that its area (pixel count for binary images) is set to a predetermined value (β). This value resizes the object in

the image, either increasing its size or reducing it with respect to its original area and needs to be determined empirically to ensure that the entire object still lies within the image and is large enough to be discernible. For all image objects in this study (score and numerical library pixels set to 1 by the binarisation process), the β value is set to 5000. Scale invariance is achieved by Equations 6 and 7.

$$g(x, y) = f\left(\frac{x}{a}, \frac{y}{a}\right) \quad (6)$$

$$a = \sqrt{\frac{\beta}{m_{00}}} \quad (7)$$

where (x, y) are the pixel coordinates of each pixel in the image, m_{00} is the area (pixel count for binary image) and $g(x, y)$ is the invariant image pixel coordinates.

Translation involves moving the centre of mass of the object in the image to the centre of the image. This is achieved as shown in Equation 8.

$$h(x, y) = g(x + \bar{x}, y + \bar{y}) \quad (8)$$

where $h(x, y)$ is the translation invariant image pixel coordinates and $\bar{x}, \bar{y}$ is the centroid of the object (all pixels with intensity value of 1) in the binary image.

Zernike basis functions and consequently Zernike moments are developed from these normalised images. Computing the Zernike basis functions and subsequent Zernike moments for an image requires mapping the image to reside in a unit circle with the origin at the centre of the image. Pixels outside a unit circle are not used in computing the moments. Thus, to ensure that all score and library image pixels are captured within the unit circle, the entire binary image is resized into a larger square image of dimensions with each side equal to $2(n_x^2 + n_y^2)^{1/2}$ where n_x and n_y are, respectively, one half of the number of x and y pixels. This larger image is padded on each side of the original image with pixels with intensity 0 for black to create a new image of the larger dimensions.

The polar value ρ and corresponding polar angle θ for any pixel (x, y) in polar coordinates is calculated as:

$$\rho = \sqrt{\left(\frac{x - \bar{x}}{d}\right)^2 + \left(\frac{y - \bar{y}}{d}\right)^2} \quad (9)$$

$$\theta = \tan^{-1}\left(\frac{x - \bar{x}}{y - \bar{y}}\right) \quad (10)$$



Figure 6. Images for reference numbers library created to use in identifying recovered numbers.

where (x,y) are the pixel coordinates of each pixel for the square image, $(\bar{x},\bar{y})$ are the pixel coordinates of the centre of the image and d is half the dimension of the image.

Equations 9 and 10 map each pixel coordinates to the unit circle in polar coordinates. The radial polynomials $[R_{nm}(\rho)]$ necessary for the determination of the Zernike basis polynomials are calculated using the polar value ρ and defined order n as shown in Equation 11:

$$R_{nm}(\rho) = \sum_{s=0}^{n-m/2} (-1)^s * \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n+|m|}{2} - s\right)!} * \rho^{n-2s} \quad (11)$$

where n is a number of basis polynomials that define the order of Zernike moments, m is a positive or negative integer bound by the constraints $n - |m| = \text{even}$ and $|m| \leq n$.

The Zernike basis functions (polynomials) and the corresponding moments are determined from these real valued polynomials and the input images using Equations 12 and 13:

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho) \exp(jm\theta) \quad (12)$$

$$A_{nm} = \frac{n+1}{\pi} \sum_{x=1}^N \sum_{y=1}^N f(x,y) V_{nm}^*(\rho,\theta); \quad x^2 + y^2 \leq 1 \quad (13)$$

where V_{nm} is the Zernike polynomial, A_{nm} is the Zernike moment, $f(x,y)$ is the intensity value of an image pixel with coordinates x,y , $j = \sqrt{-1}$ and $V_{nm}^*(\rho,\theta)$ is the complex conjugate of $V_{nm}(\rho,\theta)$.

The Zernike moment vectors are developed by concatenating the Zernike moments (A_{nm}) at each order into a single vector. This is carried out for all score and numerical library images. An image pixel can be recreated using the Zernike basis polynomial and the Zernike moments up to the defined order as shown in Equation 14.

$$\hat{f}(x,y) = \sum_{n=0}^n \sum_{m=0}^n A_{nm} V_{nm}(\rho,\theta) \quad (14)$$

where $\hat{f}(x,y)$ is a reconstructed pixel in the translated (centred) representation. The pixels added previously to ensure the original image is inscribed in a unit circle are removed, leaving the reconstructed image of the same size as the original image.

The process of extracting Zernike moments requires the definition of a finite order of Zernike basis (n) to be used in extracting Zernike moments from an image. Higher order Zernike moments are more susceptible

to noise and thus undesirable for image representation.⁴² Conversely, higher number moments also better capture the finer details of an image and cannot totally be discarded. As such, it is necessary to determine an order of moments that will both adequately capture the characteristics of the input image and yet be robust to noise. A good way of estimating such an order is by its reconstruction error.^{36,38,39} This requires the comparison of the original input image to its reconstructed version from a set of Zernike moments at a specified Zernike polynomial order. The lower the reconstruction error, the better the order captures image features. The mean square reconstruction error (ϵ) between the original image and its reconstructed versions at increasing orders is used as a comparative measure to determine this difference as shown in Equation 15:

$$\epsilon = \frac{\sum_{i=1}^N \sum_{j=1}^N [f(x_i, y_j) - \hat{f}(x_i, y_j)]^2}{\sum_{i=1}^N \sum_{j=1}^N [f(x_i, y_j)]^2} \quad (15)$$

where N is the number of pixels in the x and y directions.

As such, the difference between an image and its reconstructed version from a set of Zernike moments at a defined order can be used as a measure of the validity of the order as representative of the features of the original image. To determine the appropriate order, the contribution of each respective order moments to the reconstruction process is measured by the difference in its reconstruction error from the next order (n) and this difference $C(n)$ is computed as shown in Equation 16:^{36,38}

$$C(n) = \epsilon_{n-1} - \epsilon_n \quad (16)$$

The cumulative sum of these differences for an increasing order is used to estimate the orders that fully capture the image features.³⁶ Examples of this process based on cumulative sum plots are shown in Figure 7 for two library number images and a recovered number score image. Similar plots are obtained for other library and score images.

The insets in Figure 7 show the plots for the Zernike moment vectors at basis polynomial orders of 20 and 40, respectively, for a reference library number 2. Vectors such as these from each image are used in recreating the image and calculating the mean square errors. They are subsequently used for comparing the recovered images to the reference library images.

Figure 7 shows the image representation ability of Zernike basis polynomials from 1 through 50 determined

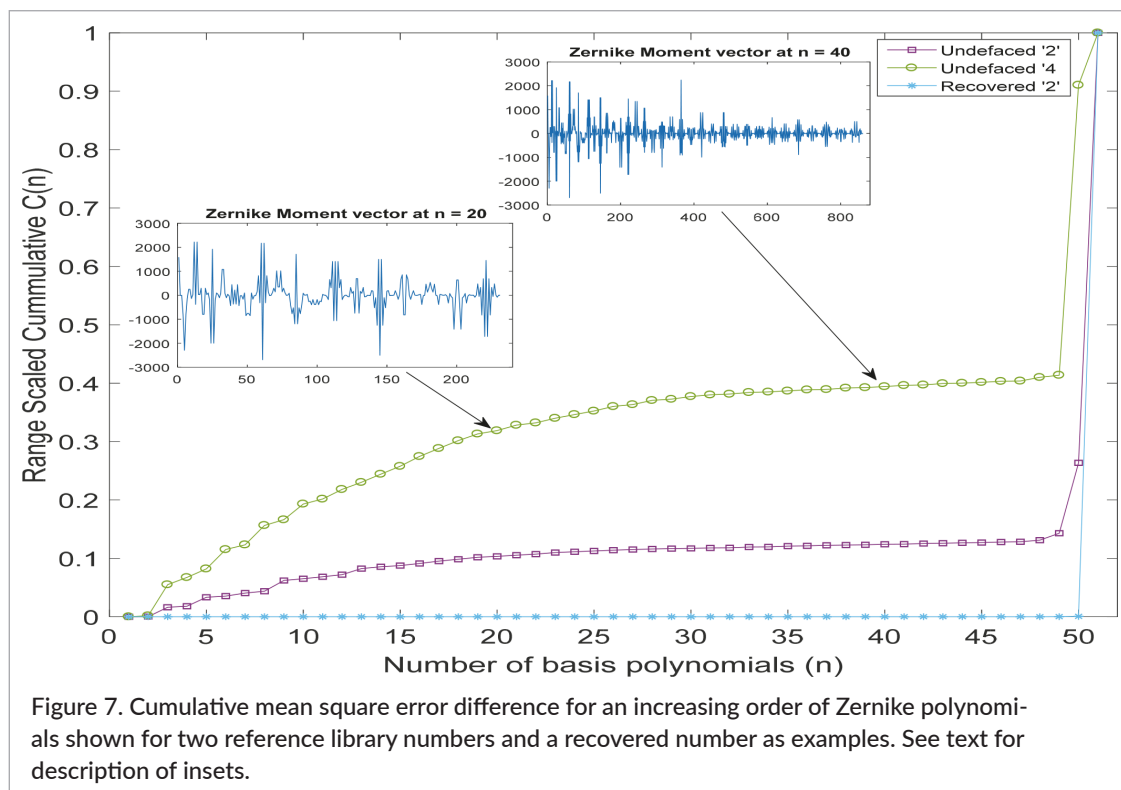


Figure 7. Cumulative mean square error difference for an increasing order of Zernike polynomials shown for two reference library numbers and a recovered number as examples. See text for description of insets.

by the cumulative sum. The cumulative sum increases steadily until all major features have been captured and then it slowly increases, indicating little change in the mean square errors for consecutive orders and subtle features are now being included. Ultimately, there is a sudden increase at higher orders indicating all image features have been captured and higher order polynomials are adding non-feature related elements and distorting the reconstructed image. From the plot, this increase occurs at order 51 and thus, it can be deduced that beyond order 50, the reconstructed image is distorted and the Zernike moment vectors will contain non-feature related elements that may impair the classification.

Similarity measures

Fifteen similarity measures are used in this study to compare the Zernike moment vectors of the score images to those for 10 reference libraries made up of unaltered numbers. Zernike moment vectors are determined for each library image, to compare to the Zernike moment vectors of the score images. Several comparisons are made, one for each Zernike polynomial order from 15 to 50 in steps of 5 for 8 sets of Zernike moment vector comparisons. Table A1 in the Appendix shows the simi-

larity measures used and their equations, each designed to define a perfect match between the two vectors as zero and increasing in value with a decrease in similarity.

All score images developed for a defaced number are evaluated irrespective of whether or not a recovered number can visually be recognised. Using all score images avoids the necessity of preselecting score images to compare and ensures that parts of a defaced number that may be partially reproduced in a range of score images are utilised in making the identification.

Data fusion

With 15 score images being compared with 8 sets of Zernike moment vectors across 10 libraries using 15 similarity measures each, there is a need to determine a consensus across the resulting 18,000 ($15 \times 8 \times 10 \times 15$) similarity values obtained. To achieve consensus, a high-level data fusion approach is applied. Multiple fusion rules are also used where each rule is applied to extract a consensus value from all 18,000 similarity values relative to each number identification possibility. These rules are sum, geometric mean, harmonic mean, median, L-2 norm, L-1 norm, truncated geometric mean and truncated harmonic mean. Most rules are applied twice, once to

raw values and another time to rank values. Descriptions of the rules are given in Appendix I. Briefly, all similarity values for a given defaced number are assembled into an $18,000 \times 10$ matrix where the columns correspond to the number identification targets 0 to 9. Each row is normalised to unit length to remove magnitude variations between the values from respective similarity measures. Each fusion rule is applied to the columns of this matrix as well as to the matrix after transformation to rank values. The smaller the fusion rule value for a particular column, the more similar the defaced number is to the corresponding target library number. By using normalised and rank values, there are 14 fusion rules.

The defaced number is matched to the target number that is most consistently (eight of the fourteen rules, i.e. majority vote) ranked lowest across the fusion rules and whose sum across the ranked fusion rules is lowest. A value "NC" is assigned if no particular rank has a majority across the fusion rules.

Algorithms

Algorithms for lock-in analysis, PCA, Zernike moments, similarity measures and fusion rules were written by the authors using MATLAB 9.2.

Results and discussion

Stainless steel plate with numbers

A lock-in frequency of 50mHz was selected to be used in pulsing the thermal energy through the sample. This lock-in frequency was selected using phase difference plots developed using the procedure in References 9 and 18. Figure 8 shows the phase difference plots for the measured defaced areas. The phase differences were calculated as the difference between the phase values

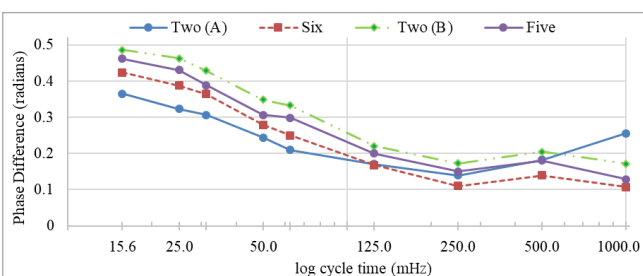


Figure 8. Phase difference plots for each defaced number on stainless steel sample.

within the defaced region where the number previously existed and average phase values of a clean non-stamped area for various lock-in frequencies. From this figure, it can be observed that while there is no blind frequency for the numbers, suboptimal frequencies exist that give the least contrast range from 250mHz to 1000mHz where the phase differences oscillate in a range of less than 0.2 rad. Phase differences increase afterwards allowing for the selection of optimal frequencies for all the defaced numbers observed to range from 50mHz to 15.6 mHz for this sample.

In selecting a modulation frequency to use in the experiments, the depth of thermal diffusion was also taken into consideration. A function of the modulation frequency, the depth of thermal diffusion is the depth to which the phase images capture characteristics within an object. With decreasing modulation frequency, the depth of thermal diffusion increases. The sample used in these experiments is just over 6 mm after defacing and the 50mHz modulation frequency has a thermal diffusion length of 4 mm in stainless steel, which is still within the limits of the sample thickness. As a result, 50mHz was used as the modulation frequency for the experiments carried out.

Figure 9 shows the mean infrared thermal image for one pulse cycle of a clean stainless steel surface left intact with no number stamped into it. The image is comparable to those shown in Figure 10 for the areas around the

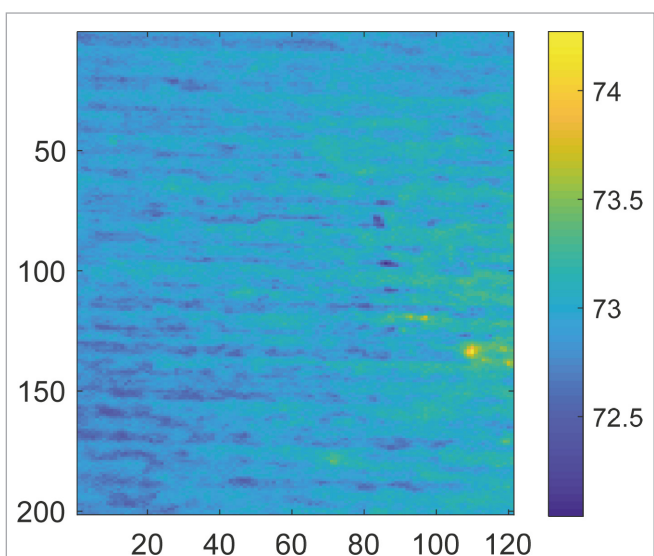


Figure 9. Raw thermal image of a clean surface (axes represent the pixel coordinates and the colour bar shows the temperature range ($^{\circ}\text{C}$) of the sample).

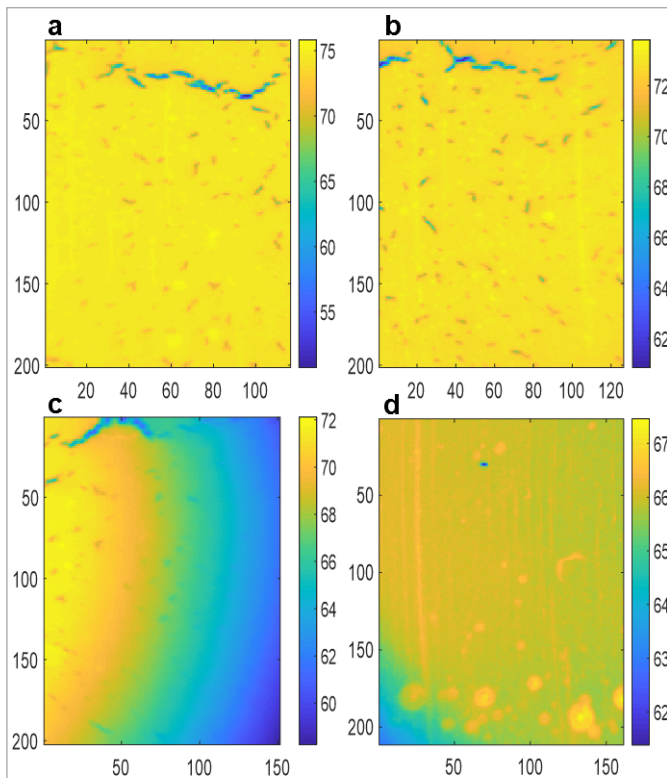


Figure 10. Raw thermal images of areas around defaced numbers: a) six, b) two, c) five and d) zero (axes represent the pixel coordinates and colour bars show the temperature range ($^{\circ}\text{C}$) of the sample).

defaced numbers (the boxed in areas in Figure 5). As with the sample at room temperature, the numbers removed cannot be uniquely distinguished from a clean surface in the infrared thermal images.

Figures 11 and 12 show phase images over a single pulse cycle for the clean region and defaced numbers, respectively. Although the phase images have effectively removed such unwanted features such as inhomogeneous illumination as well as surface radiation and local emissivity variations from the images, it is still impossible to identify the serial numbers in Figure 12.

Score images developed separately from datasets of the phase and amplitude images were visually inspected to determine which dataset best reproduces the defaced number. This process showed amplitude images for defaced 6, 2 and 0 and phase images for defaced 5 to be best. This difference is likely caused by local changes in the thickness of the paint layers due to the manual process utilised in applying the paint. A relatively thicker paint layer over a number like the defaced 5 would mean the phase image would contain more information about

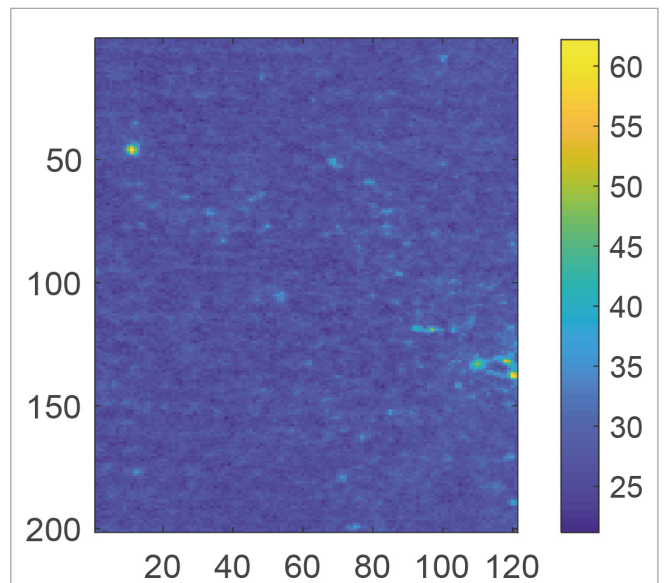


Figure 11. Phase image of a clean surface (axes represent the pixel coordinates and the colour bar shows the degree of phase shift).

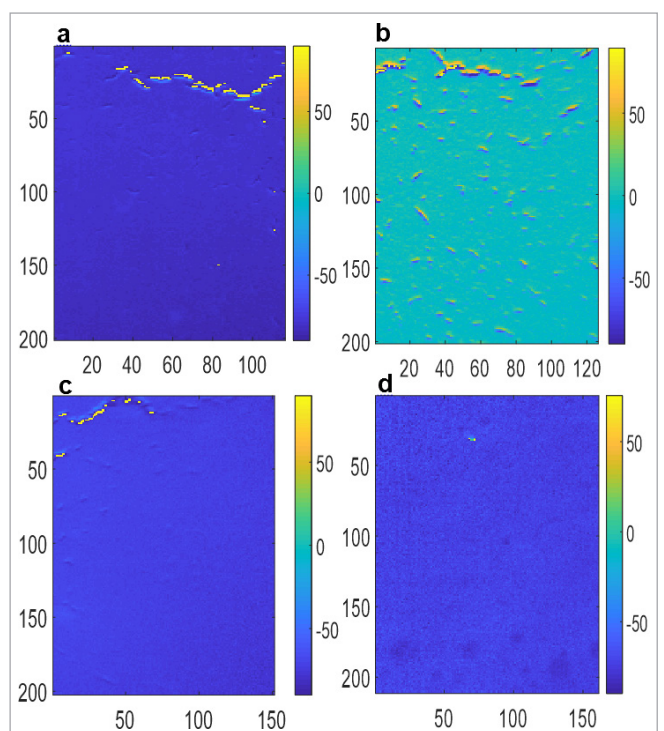


Figure 12. Phase images of areas around defaced numbers: a) six, b) two, c) five and d) zero (axes represent the pixel coordinates and the colour bar shows the degree of phase shift).

the defaced 5, because phase images can minimise the environmental background interference and capture sub-surface discontinuities at greater depths than possible

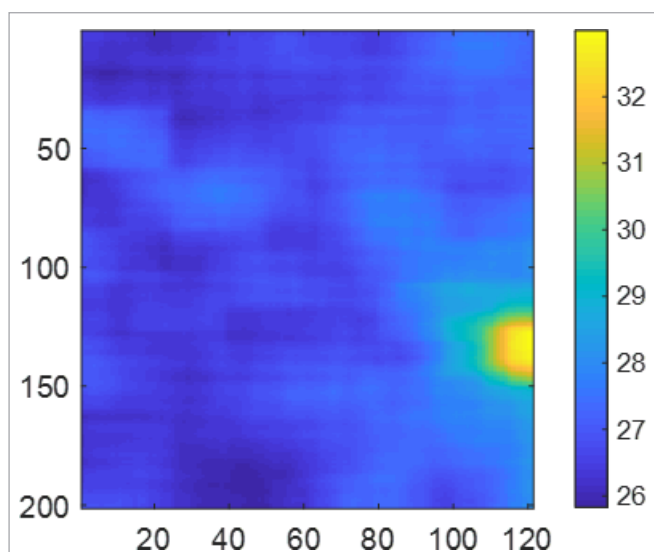


Figure 13. Filtered phase image of a clean surface (axes represent the pixel coordinates and the colour bar shows the filtered phase values).

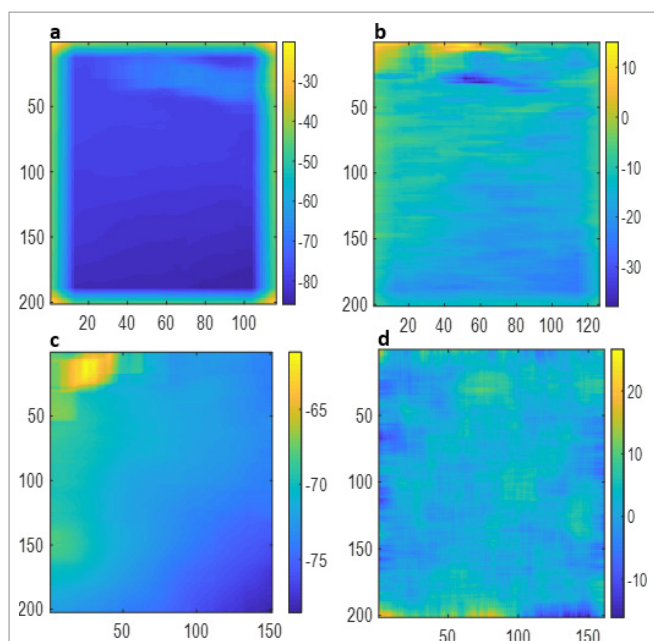


Figure 14. Filtered phase images of areas around defaced numbers: a) six, b) two, c) five and d) zero (axes represent the pixel coordinates and the colour bar shows the filtered phase values).

with the amplitude images. Thus, if the thickness of the paint layer for the defaced 5 is significantly thicker than for the defaced number 2, the phase image of the 5 would better characterise the initial stamped number in this area.

Table 1 shows the percent variance of information explained in each PC. As by convention, PC1 explains the most amount of variance and the other PCs explain some of the remaining variance present in decreasing order. The shaded cells indicate the PC where the defaced number was best reproduced for each dataset.

Figure 15 presents the score images from the clean undefaced region. PC1 is responsible for 99% of the variance within the images. PC2 to PC15 are responsible for the remaining 1% of the variance. The score images show no identifiable structure although with some random high intensity areas due to surface roughness of the sample. This is expected, as no crystalline deformation exists within the sample in this region.

Figure 16 shows the score images from the amplitude images of the area around the number “6”. The first principal component PC1 is responsible for 98% of the variance and PC2 to PC15 for the remaining 2%. However, unlike for the clean area, one of the score images (PC13) shows defined intensity contrast localised around the section where the number previously existed. This PC accounts for only 0.0003% of the variance across the phase images. This can be attributed to how relatively small the phase shift due to thermal gradient differences in the zone of plastic strain is compared to other sample features within the thermal depth range, highlighting the robustness of PCA in recognising and characterising such small variances across the pixels.

Figure 17 similarly shows the refolded score images (contrast adjusted) from PCA of the regions around the numbers “2”, “5” and “0”. PC11 presented in Figure 17a shows some contrast in the region where the number existed. This PC accounts for only 0.003% of the variance. Likewise, for the defaced 5, PC10 shown in Figure 17b shows some contrast in intensity allowing for possible identification of the number that was defaced.

PC1, accounting for 49% of the variance, shows intensity contrast to possibly identify the number 0 that was defaced. This can be seen in Figure 17c. The relatively larger value in the variance accounted for in the PC that reproduces this number is due to the fact that the number was initially stamped deeper than the others and so had a more extensive zone of plastic strain. This allowed for the thermal gradient difference in this region to account for a larger percentage of the variance in the thermal depth range.

Also, the difference in the PC best identifying the thermal gradient variation and reproducing the defaced

Table 1. Percentage of variance in each PC for defaced numbers on the stainless steel sample (shaded boxes represent the PC where the respective numbers were best reproduced).

PC	Percent variance (%)				
	Clean area	Six	Two	Five	Zero
1	99.0328	98.5497	82.2910	99.8550	48.6725
2	0.5639	1.0724	16.8583	0.0820	8.9673
3	0.2580	0.2219	0.7022	0.0375	5.5631
4	0.1152	0.0993	0.0624	0.0125	5.4441
5	0.0108	0.0312	0.0283	0.0060	4.7530
6	0.0064	0.0142	0.0162	0.0017	4.4526
7	0.0027	0.0052	0.0127	0.0013	3.7963
8	0.0023	0.0027	0.0074	0.0010	3.6741
9	0.0019	0.0010	0.0055	0.0008	3.2420
10	0.0014	0.0009	0.0046	0.0006	2.8108
11	0.0012	0.0006	0.0034	0.0004	2.5601
12	0.0009	0.0004	0.0029	0.0004	2.2362
13	0.0008	0.0003	0.0021	0.0003	1.9693
14	0.0007	0.0002	0.0016	0.0002	1.1534
15	0.0006	0.0002	0.0016	0.0002	0.7053

number can be attributed to discrepancies in the surface conditions around each number from the machining process used in removing them as well as some variation in the depth of stamped marks and consequently the depth of the zone of plastic strain.

Although hard to visually observe, PCA, in projecting the phase and amplitude images into different orthogonal directions, reproduces the variation in degree of phase shift between the input images independent of other linear variations in a single score.

The binarised score images used in the image classification process used to verify the recovered numbers are shown in Appendix B.

Figure 18 images the results of each fusion rule applied across the merits from comparing the recovered number “6” to each digit. This image gives a visual representation of the results used to independently verify the recovered numbers. From the figure, it is seen that the process works efficiently in matching the recovered number to a “6”, consistently having the lowest rank across the fusion rules. Table 2 shows the final results with using both majority vote and sum across the fusion rules. This table reiterates the information in Figure 18 in showing that there is an overall consensus across all the fusion rules

matching the recovered number to a 6. Again, in order to correctly match a defaced number to a target number, the defaced number must be ranked lowest by a majority vote of the fusion rules and have the lowest sum of the ranked fusion rule values. Otherwise, the defaced number is non-identifiable.

The process is similarly efficient in matching the other recovered numbers to their corresponding digits as shown in Tables 3–5. However, as seen in Table 4, for the identification of the defaced “5”, there is an increase in the sum across the fusion rule ranks. This, as can be seen in the visual representation of the fused merits across the fusion rules in Figure 19, is a result of some of the fusion rules misclassifying the number, possibly due to some score images being misidentified. This highlights the necessity of utilising a consensus among the fusion rules. Both of the consensus methods applied correctly matched the recovered number to the right digit despite the contrasting results from some fusion rules.

Classification results are not shown for the clean area as it is known that this area had no number stamped into it and thus attempting to classify its score images to possible number digits (0–9) will result in misleading classifications of the dataset to a particular digit.

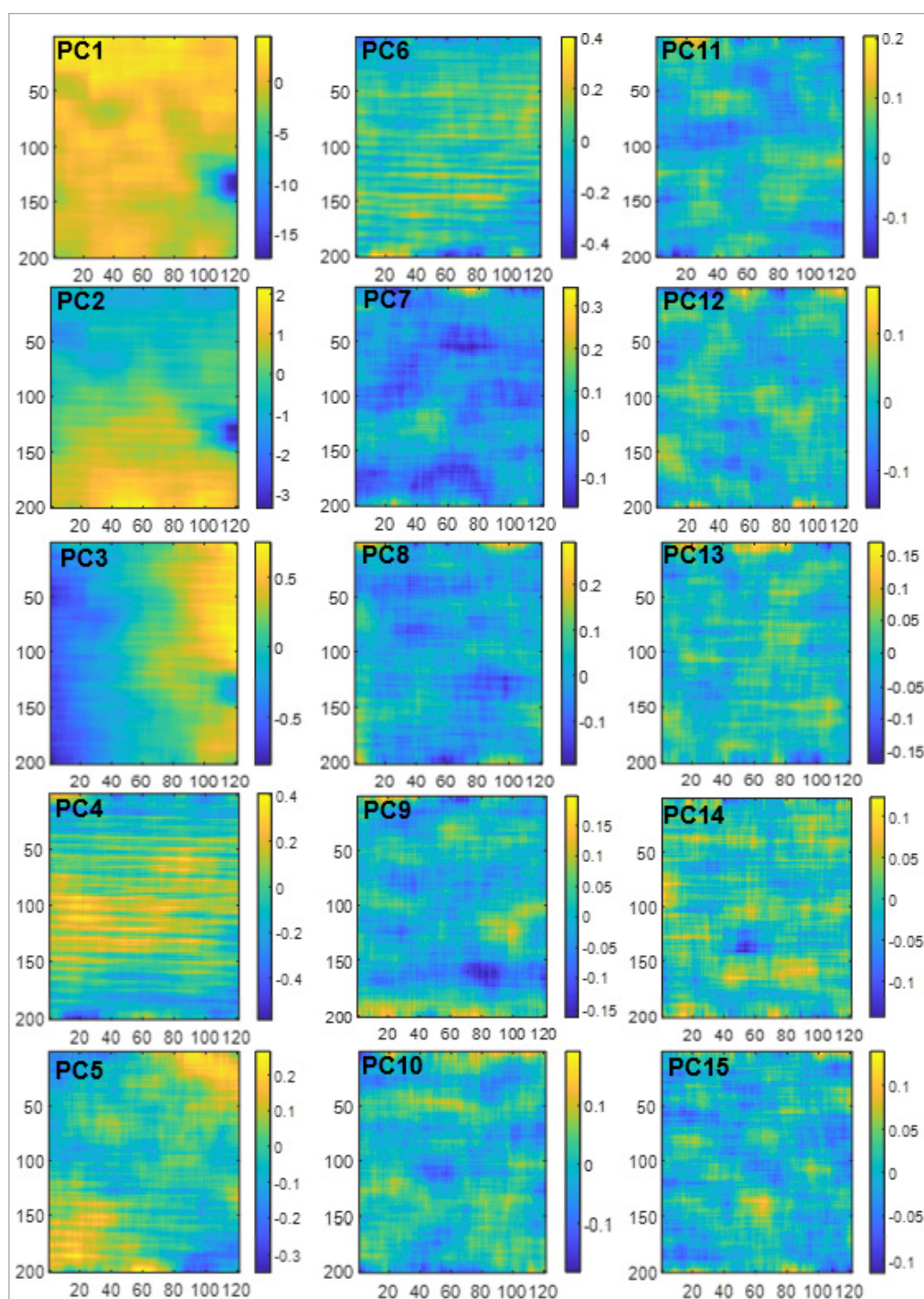


Figure 15. Score images for clean surface (axes represent the pixel coordinates and the colour bar shows the range of score values).

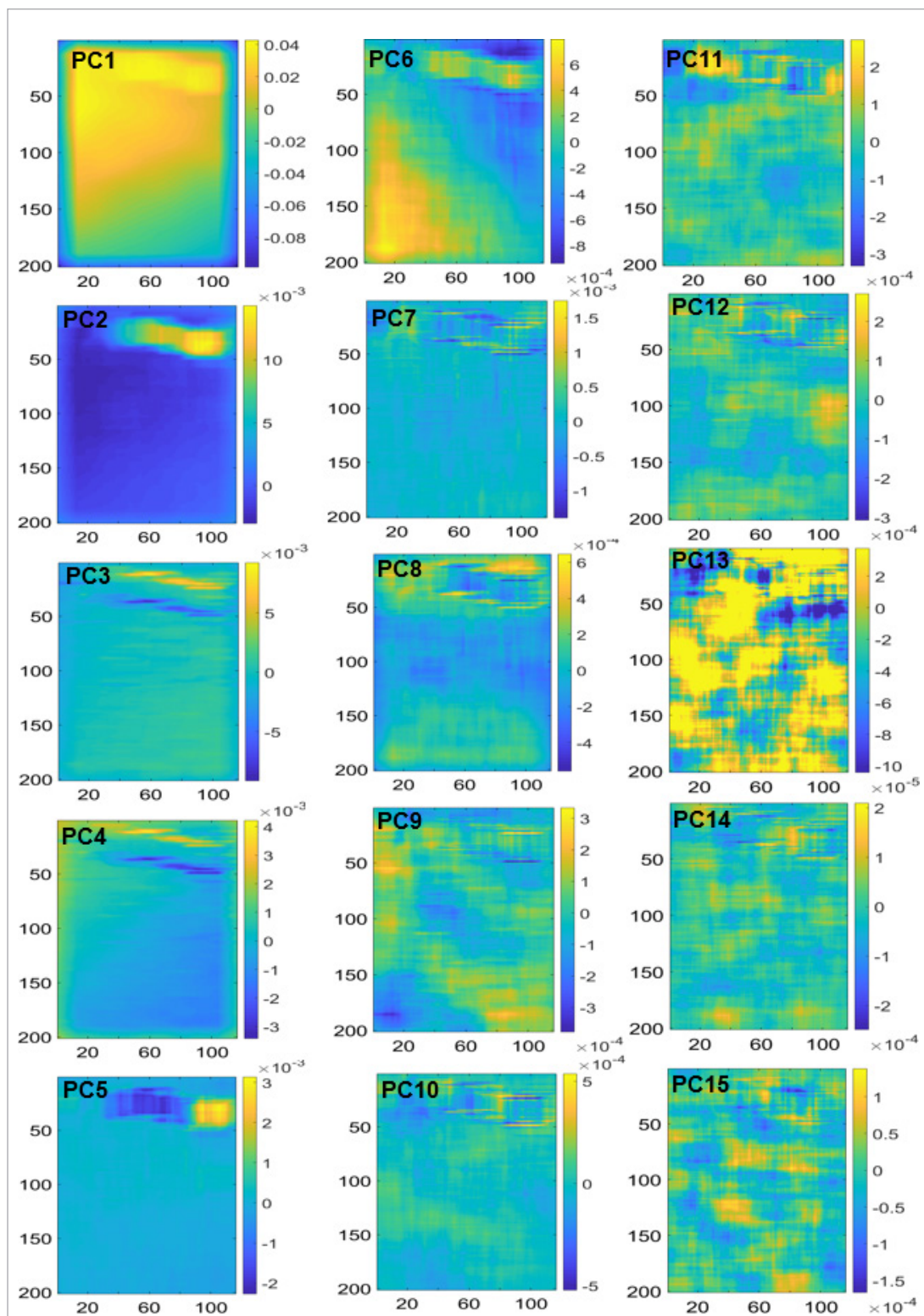


Figure 16. Score images for area around defaced number six (axes represent the pixel coordinates and the colour bar shows the range of score values).

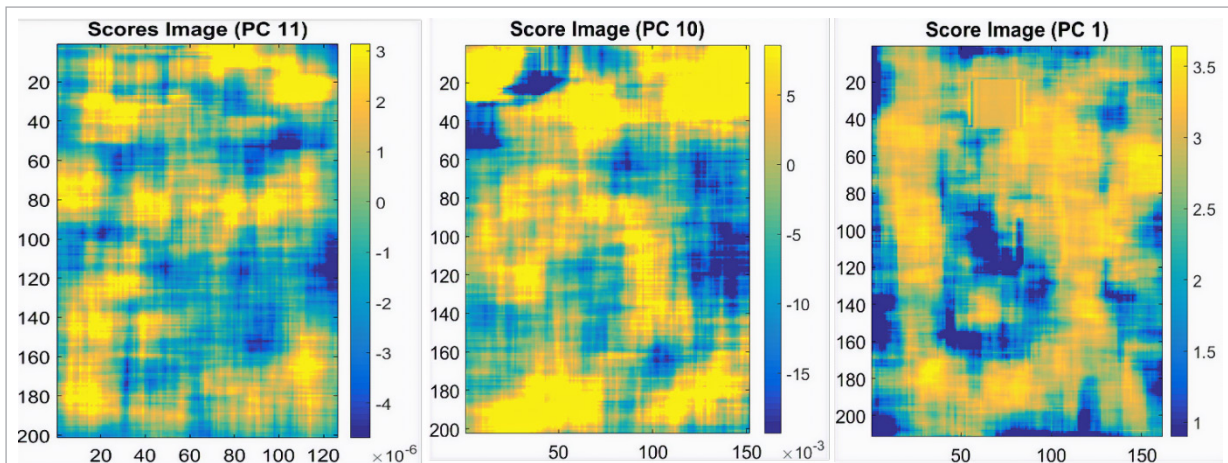


Figure 17. Score images showing recovered numbers: a) two, b) five and c) zero (axes represent the pixel coordinates and the colour bar shows the range of score values).

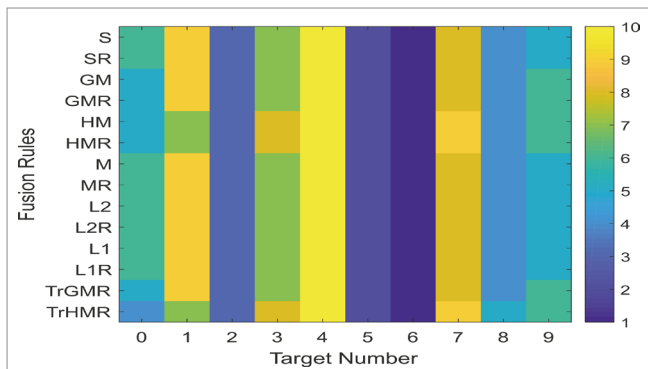


Figure 18. Image of fused measures from all fusion rules for the recovered number 6 (See Appendix for rule definitions. Colour bar shows rank from 1 to 10).

Conclusion

Using a non-destructive process, the object of the study was to determine the possibility of recovering defaced serial numbers and independently verifying the authenticity of the recovered numbers. The developed process, successfully employed the LIT technique in combination with PCA, to effectively recover the defaced numbers by utilising the thermal gradient differences in the zone of plastic strain exposed when the numbers were mechanically removed. This process is quite general and should be applicable across different items including firearms, automobiles etc. To verify the identification of the recovered numbers, invariant features are extracted

Table 2. Fusion of similarity measure values for area around defaced number six.

Target number	0	1	2	3	4	5	6	7	8	9
Majority vote	6	9	3	7	10	2	1	8	4	5
Sum	77	120	42	101	140	28	14	115	57	76

Table 3. Fusion of similarity measure values for area around defaced number two.

Target number	0	1	2	3	4	5	6	7	8	9
Majority vote	5	0	1	8	10	2	4	NC*	NC*	7
Sum	72	115	14	112	138	36	52	87	38	106

*NC indicates there was no consensus rank for the digit across the fusion rules

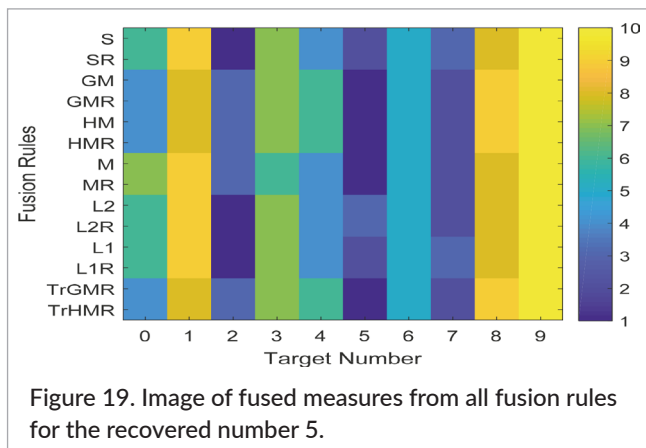
Table 4. Fusion of similarity measure values for area around defaced number five.

Target number	0	1	2	3	4	5	6	7	8	9
Majority vote	NC*	9	3	7	4	1	5	2	8	10
Sum	74	120	30	96	68	22	70	32	118	140

*NC indicates there was no consensus rank for the digit across the fusion rules

Table 5. Fusion of similarity measure values for area around defaced number zero.

Target number	0	1	2	3	4	5	6	7	8	9
Majority vote	1	7	2	10	3	5	6	4	9	8
Sum	14	104	28	138	48	70	84	50	128	106



from the images using Zernike moments. These invariant features are then compared to those of a clean number library. The comparison is done using several similarity measures to capture different possible characteristic patterns within the vectors. As such, it was pertinent to achieve a consensus among the measures and several fusion rules are used to achieve consensus to the library number. Such a process ensures independence from human bias in the identification of the recovered number.

Acknowledgements

This work was supported by the National Institute of Justice, Office of Justice Programs (grant numbers 2013-R2-CX-K012 and 2015-R2-CX-0017). The opinions, findings and conclusions or recommendations expressed in this exhibition are those of the authors and do not necessarily reflect those of the Department of Justice.

Declarations of interest

Patent pending

References

1. P.B. Wilson, "The restoration of erased serial identification marks", *Police J. Theory, Pract. Princ.* **52**(3), 233 (1979). <https://doi.org/10.1177/0032258X7905200304>
2. L. Nickolls, *The Scientific Investigation of Crime*. Butterworth, London (1956).
3. D.E. Polk and B. Giessen, "Metallurgical aspects of serial number recovery", *AFTE J.* **21**(2), 174 (1989).
4. G. Peeler, S. Gutowski and S. Wrobel, "The restoration of impressed characters on aluminum alloy motorcycle frames", *J. Forensic Identif.* **58**(1), 27 (2008).
5. R. Kuppaswamy, "Metallographic etching of aluminium and its alloys for restoration of obliterated marks in forensic science practice and investigations", in *Aluminium Alloys, Theory and Applications*. InTech (2011). <https://doi.org/10.5772/15693>
6. G. Wightman and J. Matthew, "Restoration of stamp marks on steel components", *Forensic Sci. Int.* **180**(1), 32 (2008). <https://doi.org/10.1016/J.FORSCIINT.2008.06.017>
7. H. Katterwe, "Restoration of serial numbers", in *Forensic Investigation of Stolen-Recovered and Other Crime-Related Vehicles*. Elsevier, p. 177 (2006). <https://doi.org/10.1016/B978-012088486-5/50041-7>
8. C. Parisien, G. Kolhatkar, F. Crispino, A. Lajeunesse and A. Ruediger, "Reconstruction of obliterated characters in polycarbonate through spectral imaging", *Anal. Chem.* **89**(21), 11648 (2017). <https://doi.org/10.1021/acs.analchem.7b03069>
9. M. Choi, K. Kang, J. Park, W. Kim and K. Kim, "Quantitative determination of a subsurface defect of reference specimen by lock-in infrared thermography", *NDT E Int.* **41**(2), 119 (2008). <https://doi.org/10.1016/J.NDTEINT.2007.08.006>
10. A. Killey and J.P. Sargent, "Analysis of thermal nondestructive testing", *J. Phys. D. Appl. Phys.*

- 22(1), 216 (1989). <https://doi.org/10.1088/0022-3727/22/1/032>
11. C.M. Sayers, "Detectability of defects by thermal non-destructive testing (retroactive coverage)", *Br. J. Non-Destr. Test.* **26(1)**, 28 (1984).
12. K. Cramer, P. Howell and H. Syed, "Quantitative thermal imaging of aircraft structures", in *Thermosense XVII: An International Conference on Thermal Sensing and Imaging Diagnostic Applications*. p. 226 (1995).
13. Y.A. Çengel and A.J. Ghajar, *Heat and Mass Transfer: Fundamentals & Applications*. McGraw-Hill Higher Education (2015).
14. O. Breitenstein, W. Warta and M. Langenkamp, *Lock-in Thermography: Basics and Use for Evaluating Electronic Devices and Materials*. Springer (2010). <https://doi.org/10.1007/978-3-642-02417-7>
15. A. Gleiter, C. Spießberger and G. Busse, "Lockin thermography with optical or ultrasound excitation", *J. Mech. Eng.* 619 (2010).
16. O. Breitenstein, J.P. Rakotoniaina, F. Altmann, T. Riediger and M. Gradhand, "New developments in IR lock-in thermography", in *Proc. 30th ISTFA*, p. 595 (2004).
17. O. Breitenstein, C. Schmidt and D. Karg, "Thermal failure analysis by IR lock-in thermography", *Microelectron. Fail. Anal. Desk Ref. Fifth Ed. EDFAS* (2004).
18. W. Bai and B.S. Wong, "Evaluation of defects in composite plates under convective environments using lock-in thermography", *Meas. Sci. Technol.* **12(2)**, 142 (2001). <https://doi.org/10.1088/0957-0233/12/2/303>
19. J.M. Prats-Montalbán, A. de Juan and A. Ferrer, "Multivariate image analysis: a review with applications", *Chemometr. Intell. Lab. Syst.* **107(1)**, 1 (2011). <https://doi.org/10.1016/J.CHEMOLAB.2011.03.002>
20. M. Kamruzzaman, G. ElMasry, D.-W. Sun and P. Allen, "Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis", *Anal. Chim. Acta* **714**, 57 (2012). <https://doi.org/10.1016/J.ACA.2011.11.037>
21. S. Chevallier, D. Bertrand, A. Kohler and P. Courcoux, "Application of PLS-DA in multivariate image analysis", *J. Chemometr.* **20(5)**, 221 (2006). <https://doi.org/10.1002/cem.994>
22. E. Marengo, M. Manfredi, O. Zerbinati, E. Robotti, E. Mazzucco, F. Gosetti, G. Bearman, F. France and P. Shor, "Technique based on LED multispectral imaging and multivariate analysis for monitoring the conservation state of the Dead Sea Scrolls", *Anal. Chem.* **83(17)**, 6609 (2011). <https://doi.org/10.1021/ac201068s>
23. M. Vidal and J.M. Amigo, "Pre-processing of hyperspectral images. Essential steps before image analysis", *Chemometr. Intell. Lab. Syst.* **117**, 138 (2012). <https://doi.org/10.1016/j.chemolab.2012.05.009>
24. W.F. de Carvalho Rocha, G.P. Sabin, P.H. Março and R.J. Poppi, "Quantitative analysis of piroxicam polymorphs pharmaceutical mixtures by hyperspectral imaging and chemometrics", *Chemometr. Intell. Lab. Syst.* **106(2)**, 198 (2011). <https://doi.org/10.1016/j.chemolab.2010.04.015>
25. I.T. Jolliffe, *Principal Component Analysis*. Springer (2002).
26. S. Marinetti, E. Grinzato, P.G. Bison, E. Bozzi, M. Chimenti, G. Pieri and O. Salvetti, "Statistical analysis of IR thermographic sequences by PCA", *Infrared Phys. Technol.* **46(1-2)**, 85 (2004). <https://doi.org/10.1016/j.infrared.2004.03.012>
27. S. Hermosilla-Lara, P.Y. Joubert, D. Placko, F. Lepoutre and M. Piriou, "Enhancement of open-cracks detection using a principal component analysis/wavelet technique in photothermal nondestructive testing", in *Proceedings of the 2002 International Conference on Quantitative InfraRed Thermography* (2002).
28. C. Ibarra-Castanedo, N.P. Avdelidis, M. Grenier, X. Maldague and A. Bendada, "Active thermography signal processing techniques for defect detection and characterization on composite materials", in *Thermosense XXXII* **7661**, 76610O (2010). <https://doi.org/10.1117/12.850733>
29. C. Ibarra-Castanedo, J.-M. Piau, S. Guilbert, N.P. Avdelidis, M. Genest, A. Bendada and X.P.V. Maldague, "Comparative study of active thermography techniques for the nondestructive evaluation of honeycomb structures", *Res. Nondestruct. Eval.* **20(1)**, 1 (2009). <https://doi.org/10.1080/09349840802366617>
30. F. López, C. Ibarra-Castanedo, V.P. Nicolau and X. Maldague, "Comparative study of thermographic signal reconstruction and partial least-squares thermography for detection and evaluation of sub-surface defects", in *12th International Conference on*

Quantitative Infrared Thermography, Bordeaux, France (2014).

31. D.L. Balageas, J.-M. Roche, F.-H. Leroy, W.-M. Liu and A.M. Gorbach, "The thermographic signal reconstruction method: a powerful tool for the enhancement of transient thermographic images", *Biocybern. Biomed. Eng.* **35**(1), 1 (2015). <https://doi.org/10.1016/J.BBE.2014.07.002>
32. X. Maldague, F. Galmiche and A. Ziadi, "Advances in pulsed phase thermography", *Infrared Phys. Technol.* **43**(3–5), 175 (2002). [https://doi.org/10.1016/S1350-4495\(02\)00138-X](https://doi.org/10.1016/S1350-4495(02)00138-X)
33. D.S. Crosby, L.C. Breaker and W.H. Gemmill, "A proposed definition for vector correlation in geophysics: theory and application", *J. Atmos. Ocean. Technol.* **10**(3), 355 (1993). [https://doi.org/10.1175/1520-0426\(1993\)010<0355:APDFVC>2.0.CO;2](https://doi.org/10.1175/1520-0426(1993)010<0355:APDFVC>2.0.CO;2)
34. C. Cheng and M.-S. Chiu, "A new data-based methodology for nonlinear process modeling", *Chem. Eng. Sci.* **59**(13), 2801 (2004). <https://doi.org/10.1016/J.CES.2004.04.020>
35. C.E. Anderson and J.H. Kalivas, "Fundamentals of calibration transfer through Procrustes analysis", *Appl. Spectrosc.* **53**(10), 1268 (1999). <https://doi.org/10.1366/0003702991945515>
36. A. Khotanzad and Y.H. Hong, "Invariant image recognition by Zernike moments", *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(5), 489 (1990). <https://doi.org/10.1109/34.55109>
37. C. Singh and P. Sharma, "Performance analysis of various local and global shape descriptors for image retrieval", *Multimed. Syst.* **19**(4), 339 (2013). <https://doi.org/10.1007/s00530-012-0288-7>
38. M.R. Teague, "Image analysis via the general theory of moments", *J. Opt. Soc. Am.* **70**(8), 920 (1980). <https://doi.org/10.1364/JOSA.70.000920>
39. C. Singh, E. Walia and R. Upneja, "Accurate calculation of Zernike moments", *Inform. Sciences* **233**, 255 (2013). <https://doi.org/10.1016/J.INS.2013.01.012>
40. C. Wolf, G. Taylor, J. Jolion, B. Verne and A. Einstein, "Learning individual human activities from short binary shape sequences". Technical Report LIRIS (2011). <https://liris.cnrs.fr/Documents/Liris-5294.pdf>
41. A. Tahmasbi, F. Saki and S.B. Shokouhi, "Classification of benign and malignant masses based on Zernike moments", *Comput. Biol. Med.* **41**(8), 726 (2011). <https://doi.org/10.1016/j.compbiomed.2011.06.009>
42. C.-H. Teh and R.T. Chin, "On image analysis by the methods of moments", *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(4), 496 (1988). <https://doi.org/10.1109/34.3913>
43. J. Flusser, B. Zitova and T. Suk, *Moments and Moment Invariants in Pattern Recognition*. John Wiley & Sons (2009). <https://doi.org/10.1002/9780470684757>
44. P.-T. Yap, R. Paramesran and S.-H. Ong, "Image analysis by Krawtchouk moments", *IEEE Trans. Image Process.* **12**(11), 1367 (2003). <https://doi.org/10.1109/TIP.2003.818019>
45. R. Mukundan, S.H. Ong and P.A. Lee, "Image analysis by Tchebichef moments", *IEEE Trans. Image Process.* **10**(9), 1357 (2001). <https://doi.org/10.1109/83.941859>
46. D.L. Hall and J. Llinas, "An introduction to multisensor data fusion", *Proc. IEEE* **85**(1), 6 (1997). <https://doi.org/10.1109/5.554205>
47. E. Borràs, J. Ferré, R. Boqué, M. Mestres, L. Aceña and O. Busto, "Data fusion methodologies for food and beverage authentication and quality assessment – a review", *Anal. Chim. Acta* **891**, 1 (2015). <https://doi.org/10.1016/J.ACA.2015.04.042>
48. A.J. Tencate, J.H. Kalivas and A.J. White, "Fusion strategies for selecting multiple tuning parameters for multivariate calibration and other penalty based processes: a model updating application for pharmaceutical analysis", *Anal. Chim. Acta* **921**, 28 (2016). <https://doi.org/10.1016/J.ACA.2016.03.046>
49. Precision Forensic Testing, *Stamped Serial Number Restoration Metal – Precision Forensic Testing*. <https://www.precisionforensictesting.com/collections/crime-laboratory-products/products/stamped-serial-number-restoration-metal> [accessed 20 June 2018].
50. Å. Rinnan, F. van den Berg and S.B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra", *TrAC Trends Anal. Chem.* **28**(10), 1201 (2009). <https://doi.org/10.1016/j.trac.2009.07.007>

Appendix A

Similarity measures

Correlation coefficient

Pearson correlation

This is a measure of the linear relationship between two vectors. A perfect value of 1 means that both vectors have a perfect correlation between them and 0 means there is no correlation between them. In keeping with the convention adopted for this study, the correlation coefficient is subtracted from 1 so that a smaller value will indicate a higher correlation between the vectors.

$$CC = 1 - \frac{S_{ls}}{S_l S_s} = 1 - \frac{\sum_{i=1}^n (x_{si} - \bar{x}_s)(x_{li} - \bar{x}_l)}{\sqrt{\sum_{i=1}^n (x_{si} - \bar{x}_s)^2} \sqrt{\sum_{i=1}^n (x_{li} - \bar{x}_l)^2}} \quad (A1)$$

Spearman correlation coefficient

For the Spearman correlation, the values within each vector are first ranked before the correlation is calculated as above.

Euclidean distance

Vector to vector Euclidean distance

The Euclidean distance is a measure of the distance in space between two vectors. This is determined as the square root of the square of the difference between corresponding points in both vectors being compared. The smaller the value of the resulting merit, the higher the degree of similarity between them.

$$ED = \sqrt{(x_l - x_s)(x_l - x_s)^T} \quad (A2)$$

Vector outer product Euclidean distance

This is similar to Euclidean distance except the outer vector products are used for the comparison ($X_l = x_l x_l^T$ and $X_s = x_s x_s^T$). The outer products are unfolded to form a vector and Equation A2 is used.

Angle between vectors

This similarity measure is determined as the cosine of the angle between two vectors for a shape comparison. The value is subtracted from 1 for a smaller value indicating a higher degree of similarity between vectors.

$$\cos \theta = 1 - \frac{|x_l^T x_s|}{x_l x_s} \quad (A3)$$

Square of the angle

This measure involves taking the square of the value for the angle between the vectors and then subtracting from 1 for a smaller value for a higher degree of similarity.

Determinant

This similarity measure is a measure of the space size formed by two vectors and calculated by Equation A4. The smaller the resulting value, the smaller space and hence, the more similar the vectors.

$$Det = \begin{vmatrix} x_l^T & x_s^T \\ x_l & x_s \end{vmatrix} = (x_l x_s \sin \theta)^2 \quad (A4)$$

Unconstrained Procrustes analysis

This measure is a transformation process that, as used in this study, determines how much transformation is required to make the matrix formed by one outer vector product similar to another outer vector product. Transformation occurs in the form of rotation, dilation and translation of each vector to make them match. A Frobenius norm (F) of the matrix difference between the transformation matrices for the two vectors is used as a final measure of the similarity between them. This value represents the matrix from the difference between the transformation matrices as a scalar with a smaller value indicating a higher degree of similarity.

$$\begin{aligned} X_l &= X_s F_{sl} \\ F_{sl} &= X_s^+ X_l \\ X_s^+ &= \frac{X_s}{X_s^4} \\ F_{ss} &= X_s^+ X_s \\ F &= F_{sl} - F_{ss} F \end{aligned} \quad (A5)$$

where F is a transformation matrix necessary to make X_s most similar to X_l ; $F_{sl} - F_{ss} F$ is the Frobenius norm for the matrix difference between the two transformation matrices F_{sl} and F_{ss} calculated as the square root of the sum of squares of each value in the difference matrix.

Four similarity merits are obtained using Procrustes analysis in this study. Two by using the mean centred outer products of each of the vectors (X_l and X_s) being compared to define the transformation matrix and another two by not mean centring the outer product matrices before determining the transformation matrices.

With mean centring, translation correction is included and without mean centring, only rotation and dilation are evaluated.

Constrained Procrustes analysis

This transformation measure defines individually the degrees of dilation and rotation required after translation to make the two vectors similar. As with unconstrained Procrustes analysis, the Frobenius norm is used to determine a final scalar similarity merit, a smaller value indicating a higher degree of similarity.

$$\begin{aligned} X_{sl} &= X_s^T X_l = U_{sl} \Sigma_{sl} V_{sl}^T \\ H_{sl} &= u_{sl} v_{sl}^T \\ X_{ss} &= X_s^T X_s = U_{ss} \Sigma_{ss} V_{ss}^T \\ H_{ss} &= u_{ss} v_{ss}^T \\ H &= H_{sl} - H_{ss}F \end{aligned} \quad (A6)$$

where U is the eigenvectors of matrix XX^T , u is the first eigenvector in U , Σ is the diagonal matrix of singular values, V is the loading matrix = eigenvectors of matrix $X^T X$ and v is the first eigenvector in V .

Two possible sets of similarity merits are obtained, with one set involving mean centred outer product arrays

(translation is included) and the other set without mean centring (only rotation and dilation).

Mahalanobis distance

This is a distance measure that determines the distance of a vector from a space formed by an array of vectors. To use this measure as a distance measure between two vectors, the pseudo-inverse of the outer product of one of the vectors is taken as the usual covariance matrix used in the Mahalanobis distance estimation. A smaller merit value in this measure indicates a smaller distance in space, and hence a higher degree of similarity between the vectors.

$$MD = \sqrt{(x_s - x_l)^T X_s^+ (x_s - x_l)} \quad (A7)$$

A second similarity merit is obtained using the Mahalanobis distance measure defined in Equation A7 by exchanging the outer vector product matrix X_s for X_l .

Table A1 gives a summary of these similarity measures and their corresponding equations.

Fusion rules

The equations below show the fusion rules to combine the $n = 18,000$ similarity values into a single defining

Table A1. Similarity measures used.

Similarity measure ^a	Equation ^b
Correlation coefficient (2)	$CC = \frac{S_{ls}}{S_l S_s}$
Euclidean distance (2)	$ED = \sqrt{(x_l - x_s)(x_l - x_s)^T}$
Angle (2)	$\cos \theta = 1 - \frac{ x_l^T x_s }{x_l x_s}$
Determinant (1)	$Det = \left \begin{pmatrix} x_l^T \\ x_s^T \end{pmatrix} \begin{pmatrix} x_l & x_s \end{pmatrix} \right = (x_l x_s \sin \theta)^2$
Procrustes analysis (PA) (4)	$F = F_{sl} - F_{ss}F$
Constrained PA (2)	$H = H_{sl} - H_{ss}F$
Mahalanobis distance (2)	$MD = \sqrt{(x_s - x_l)^T C_s^+ (x_s - x_l)}$

^aValues in parenthesis indicate the number of variations

^bSee respective descriptions for definitions of equation

x_l = feature vector of reference library image

x_s = feature vector of score image

value. In these equations S_i denotes each similarity value for the i^{th} measurement where i varies relative to the particular Zernike moment vector score image and library font. Table A2 shows the notations used in Figures 16 and 17 for the fusion rules.

$$\text{Sum} = \sum_{i=1}^n S_i \quad (\text{A8})$$

$$\text{Median} = \text{med}\{S_1 \dots S_n\} \quad (\text{A9})$$

$$\text{L-2 norm} = \sqrt{\sum_{i=1}^n S_i^2} \quad (\text{A10})$$

$$\text{L-1 norm} = \sum_{i=1}^n |S_i| \quad (\text{A11})$$

$$\text{Geometric mean} = \sqrt[n]{S_1 * S_2 * \dots * S_n} \quad (\text{A12})$$

$$\text{Harmonic mean} = \left(\frac{1}{n} \sum_{i=1}^n S_i^{-1} \right)^{-1} \quad (\text{A13})$$

The S_i values are either raw values (normalised as described in the Similarity measures section) or rank values. Using raw and rank values makes for 12 fusion rules. Two additional rules are only used on the raw normalised values. These are:

- 1) Truncated geometric mean rule: the truncated geometric mean utilises a defined percentage of the lowest values to determine the geometric mean instead of using all values. The lowest 75% of values are used in this study.
- 2) Truncated harmonic mean rule: this rule is similar to the truncated geometric mean except it utilises the harmonic mean of the truncated values.

Table A2. Fusion rule notation for figures.

Rule	Normalised raw values notation	Rank values notation
Sum	S	SR
Median	M	MR
L-2 norm	L2	L2R
L-1 norm	L1	L1R
Geometric mean	GM	GMR
Truncated geometric mean	TrGM	—
Harmonic mean	HM	HMR
Truncated harmonic mean	TrHM	—

Appendix B: Binarised score images

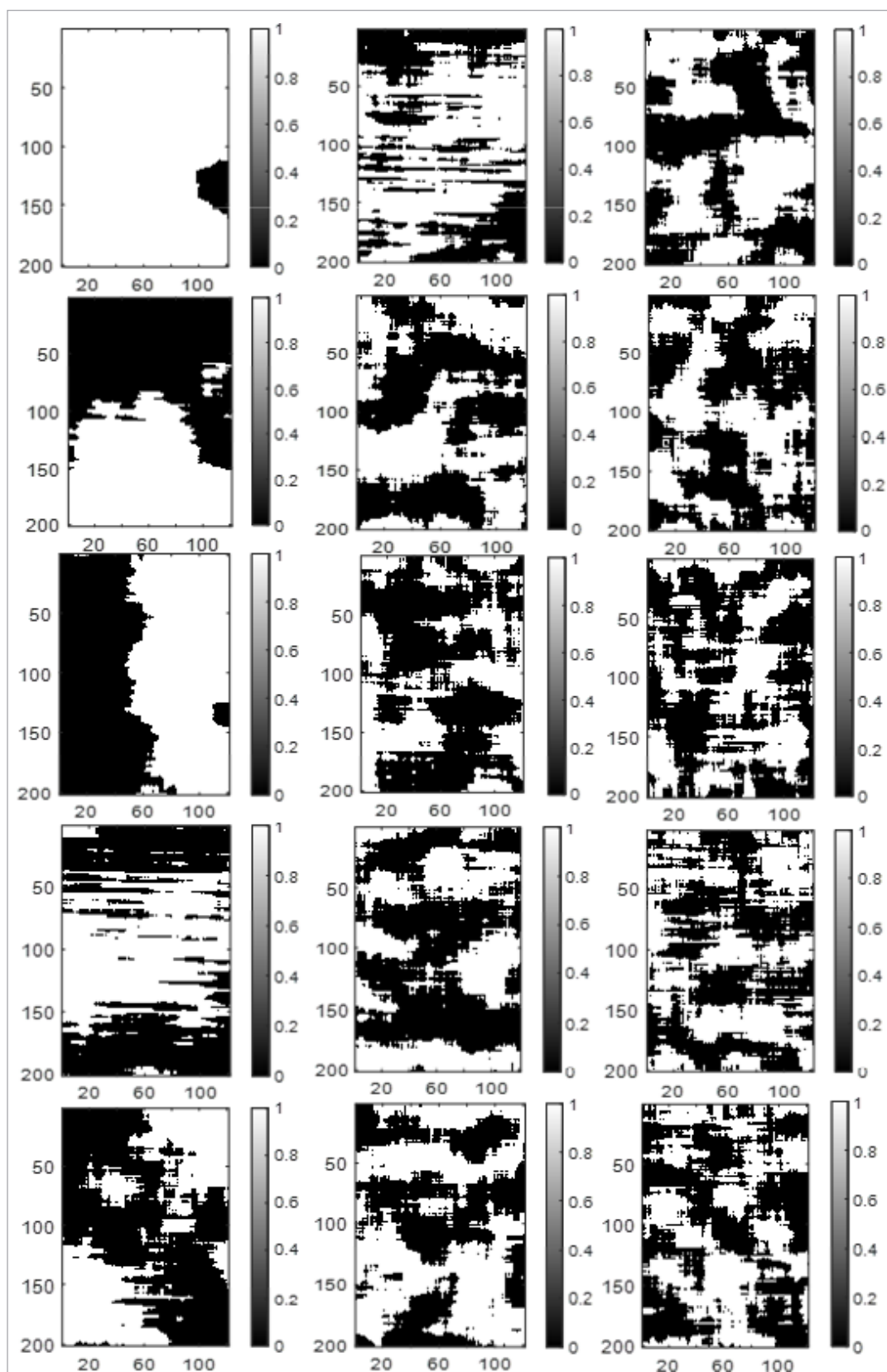


Figure B1. Binarised score images for clean surface.

