PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Research

Cite this article: Wu T, Gao B, Woo WL. 2020 Hierarchical low-rank and sparse tensor micro defects decomposition by electromagnetic thermography imaging system. *Phil. Trans. R. Soc. A* **378**: 20190584.
http://dx.doi.org/10.1098/rsta.2019.0584

Accepted: 11 May 2020

One contribution of 15 to a theme issue 'Advanced electromagnetic non-destructive evaluation and smart monitoring'.

Subject Areas:
image processing, computer vision, electromagnetism

Keywords:
inductive thermal video, tensor decomposition, hierarchical sparse, defect defection

Author for correspondence:
Bin Gao
e-mail: bin_gao@uestc.edu.cn

Hierarchical low-rank and sparse tensor micro defects decomposition by electromagnetic thermography imaging system

Tongle Wu[1], Bin Gao[1] and Wai Lok Woo[2]

[1]Department of Automation Engineering, University of Electronic Science and technology, Chengdu, People's Republic of China
[2]Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, UK

BG, 0000-0003-3377-6895; WLW, 0000-0002-8698-7605

With the advancement of electromagnetic induction thermography and imaging technology in non-destructive testing field, this system has significantly benefitted modern industries in fast and contactless defects detection. However, due to the limitations of front-end hardware experimental equipment and the complicated test pieces, these have brought forth new challenges to the detection process. Making use of the spatio-temporal video data captured by the thermal imaging device and linking it with advanced video processing algorithm to defects detection has become a necessary alternative way to solve these detection challenges. The extremely weak and sparse defect signal is buried in complex background with the presence of strong noise in the real experimental scene has prevented progress to be made in defects detection. In this paper, we propose a novel hierarchical low-rank and sparse tensor decomposition method to mine anomalous patterns in the induction thermography stream for defects detection. The proposed algorithm offers advantages not only in suppressing the interference of strong background and sharpens the visual features of defects, but also overcoming the problems of over- and under-sparseness suffered by similar state-of-the-art algorithms. Real-time natural defect detection experiments have been conducted to verify that the proposed algorithm is more efficient and accurate than existing algorithms in terms of visual presentations and evaluation criteria.

THE ROYAL SOCIETY PUBLISHING

This article is part of the theme issue 'Advanced electromagnetic non-destructive evaluation and smart monitoring'.

1. Introduction

Non-destructive testing (NDT) is an indispensable and effective tool for industrial development. To some extent, it reflects the level of industrial development of a country. The importance of NDT has been recognized in recent years. These mainly include radiographic testing (RT) [1], ultrasonic testing (UT) [2], magnetic particle testing (MPT) [3] and liquid penetrate testing (LPT) [4]. Comparing with traditional methods, inductive thermography (IT) focuses the heat on the defect due to friction or eddy current distortion, which increases the temperature contrast between the defective and defect-free regions. Thermal pattern contrast [5] is a relatively novel NDT method which has the advantages of being fast, non-contact and non-interaction and provides full field visual information.

NDT technology of an IT-integrated thermal image diagnosis system has been urgently demanded in the manufacturing industry and railway domain. Gao *et al.* [6] proposed a method to separate anomalous patterns from the transient thermal pattern by applying IT. Cheng *et al.* [7] applied IT to detect and separate the impact damage. Genest [8] used IT to detect crack defects. Netzelmann [9] used IT to study the external influence of the magnetic field for the thermal contrast of crack-type defects. He *et al.* applied IT for evaluating impact in CFRP laminates and detecting corrosion blisters [10].

Considering that the defect size is distributed in a sparse way within the whole specimen, from an imaging viewpoint this corresponds to an image with an anomalous thermal pattern embedded in the background of a normal thermal pattern spanned with a low-rank property. In this situation, the sparse and low-rank-based learning methods have become very useful tools as post-processing algorithms. In order to overcome the sensitivity of the commonly used principal component analysis (PCA) methods to abnormal outliers, the low-rank background and sparse foreground separation algorithm based on a robust PCA (RPCA) [11] has been proposed. Some related references can be found in [12–14]. In addition, some other RPCA-based variants have emerged which include variational RPCA [15], non-convex RPCA [16], Online RPCA [17]. Gao *et al.* [18] proposed a sparse non-negative matrix decomposition under a variational Bayesian framework to extract defect features. To avoid tuning model parameters, Gao *et al.* [19] proposed a variational Bayesian approach to extract a sub-group of sparse components for diagnostic imaging. Wang *et al.* [20] proposed a thermal pattern-based contrast enhancing algorithm based on a sparse abnormal optical flow field. Nevertheless, the aforementioned algorithms require the reshaping of a raw three-dimensional stream of video data into a two-dimensional matrix form, which can destroy the essential structure spatial–temporal information [21]. The authors in [22–25] attempted to investigate the adaptation of structured norms to avoid destroying the essential structure of the spatial–temporal information under certain circumstances in the matrix-wise decomposition. These alternating methods use the low-rank matrix and structured sparseness to model the spatial–temporal information in the dataset.

Instead of vector or matrix representation, a higher-order tensor represented as a multidimensional array provides a more faithful representation of the intrinsic structure underlying such data ensembles. In the machine learning field, CANDECOMP/PARAFAC and Tucker factorizations [26] are prominent baseline algorithms used in the tensor decomposition approach. With the same aim to reduce the sensibility to sparse outliers, Zhao *et al.* [27] presented the Bayesian robust tensor factorization (BRTF) algorithm for incomplete tensor data. The method provides good results on background subtraction and object recognition such as human faces. Zhou *et al.* [28] proposed an outlier-robust tensor PCA (OR-TPCA) for simultaneous low-rank tensor recovery and outlier detection. Lu *et al.* proposed TRPCA [21] based on t-SVD with a new tensor nuclear norm. Recently, based on the excellent performance of non-convex model,

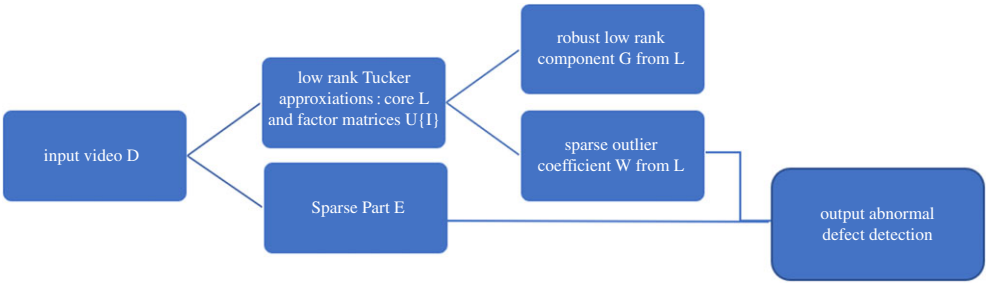

Figure 1. Flowchart of proposed HLSTD. (Online version in colour.)

Xu *et al.* [29] proposed non-convex rank minimization to extract low-rank tensor. Chen *et al.* [30] used the non-convex empirical Bayes method to model low rank tensor. In addition, based on the application in flow data, stochastic [31], incremental tensor [32] and recursive [33] tensor decomposition algorithms have also been proposed. In the field of NDT applications, Gao *et al.* [34] used the CP decomposition to model thermography spatial-transient stage and material property characterization. Gao *et al.* [35] proposed multidimensional tensor-based inductive thermography method for gear inspection. In order to fully mine spatio-temporal video information, Lu *et al.* [36] proposed ensemble variational Bayesian tensor factorization for super resolution of CFRP debond detection.

Comparing with many emerging tensor-based algorithms in the field of machine learning, they are less widely applied in the field of NDT. In NDT, the surface morphology of the tested specimen is different, and the defect is extremely tiny and irregular. Thus, it is difficult to capture the characteristics of the defect due to the complex background and strong noise. These factors render the existing tensor machine learning algorithms handicapped when applied to defect detection. These algorithms cannot accurately separate the defects from the background and noise when the sparseness characteristics of the defects failed to be estimated correctly. This yields erroneous interpretation of the defects whether these defects are attributed to a single profile or multiple profiles of sparse components. In this paper, we propose a novel hierarchical low-rank and sparse tensor decomposition (HLSTD) algorithm. This makes full exploitation of the sparse and low-rank component by robust factorization for a core tensor. The separated foreground information with sparse outliers is embedded in the background of low-rank coefficient representations. Topics in this article are listed below.

(1) The proposed tensor factorization framework flowchart in figure 1. The time series tensor of the input three-dimensional infrared heat map is decomposed in the framework of the Tucker model for the sparse core tensor, the low-rank representation coefficient of the background is further extracted and the abnormal sparse coefficient is embedded, which is called hierarchical tensor decomposition. The robust defect extraction effect is achieved by combining the sparse components of two layers. The weak and small size sparse defects can be effectively extracted from the intense complex thermal patterns. The visual contrast with the surrounding normal background is significantly improved along with the well-preserved background information. This helps the users or the automated diagnostic system to evaluate the presence of defects and to determine the optimal sparseness in order to obtain a faithful representation of the video data factorization.
(2) Development of the augmented Lagrangian alternative direction minimization (AL-ADM) [37] numerical method to optimize the parameters of the proposed model. This enables efficient implementation of the proposed algorithm for effective detection task.

The rest of this paper has been organized as follows: the proposed algorithm is described in §2. The experimental set-up and introduction of the specimens are given in §3. Analysis of results are elaborated in §4. Finally, conclusions are drawn in §5.

2. Methodology

(a) Main notation

$\langle n \rangle$	mode-n matricization	$^{\mathrm{T}}$	matrix transpose
$\times_n$	nth mode product of a tensor by a matrix	$\otimes$	Kronecker product
$\| \|_*$	tensor nuclear norm	$\| \|$	Frobenius norm
$\| \|_1$	l_1 norm	$[]_+$	$[a]_+ = 0$, if $a < 0$, else $[a]_+ = a$

(b) Proposed model

The electromagnetic thermal image sequence can be modelled as a three-dimensional tensor structure $\mathcal{D} \in \mathbb{R}^{W \times H \times T}$, where W and H denote the resolution of image length and width, respectively. T denotes the number of frames in a whole thermal video. Under the existing low-rank and sparse tensor decomposition frameworks, the mathematical model is expressed as follows:

$$\left.\begin{aligned} &\min_{\mathcal{L},\mathcal{E}} \|\mathcal{L}\|_* + \lambda\|\mathcal{E}\|_1 \\ &\text{s.t.} \quad \mathcal{D} = \mathcal{L} + \mathcal{E} \end{aligned}\right\}, \tag{2.1}$$

where $\| \|_*$ and $\| \|_1$ can be regarded as low-rank and sparse constraints to extract the low-rank tensor L and the sparse tensor E, respectively. In the field of NDT applications, the image of defects is embedded in the background of the specimen in the form of sparse distribution. The aim is to extract the sparse defect thermal pattern $\mathcal{E}$ by using the decomposition model. However, this kind of direct decomposition model is affected by the parameters λ, resulting in either under-sparse or over-sparse detection results. Therefore, by integrating the above robust decomposition methods, we propose the following hierarchical low-rank sparse tensor decomposition based on HOSVD (high-order singular value decomposition) [26]:

$$\left.\begin{aligned} &\min_{\mathcal{L},\mathcal{G},\mathcal{W},\mathcal{E},\boldsymbol{U}^{(1)},\boldsymbol{U}^{(2)},\boldsymbol{U}^{(3)}} \|\mathcal{G}\|_* + \lambda\|\mathcal{W}\|_1 + \|\mathcal{E}\|_1 \\ &\text{s.t.} \quad \mathcal{D} = \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} + \mathcal{E} \\ &\qquad \mathcal{L} = \mathcal{G} + \mathcal{W} \\ &\qquad \boldsymbol{U}^{(1)^{\mathrm{T}}}\boldsymbol{U}^{(1)} = \boldsymbol{I} \\ &\qquad \boldsymbol{U}^{(2)^{\mathrm{T}}}\boldsymbol{U}^{(2)} = \boldsymbol{I} \\ &\qquad \boldsymbol{U}^{(3)^{\mathrm{T}}}\boldsymbol{U}^{(3)} = \boldsymbol{I}, \end{aligned}\right\} \tag{2.2}$$

where $\mathcal{D}$ denotes three-dimensional thermography sequence and $\mathcal{L}$, $\boldsymbol{U}^{(1)}$ denote core tensor, factor matrices under robust Tucker decomposition framework, respectively. $\mathcal{E}$ indicates the sparse outliers component and $\mathcal{G}$, $\mathcal{W}$ denote the low-rank part, sparse component under tensor principle component analysis framework, respectively. Comparing with the state-of-the-art RTPCA models such as (2.1), it separates the sparse defect thermal pattern into component $\mathcal{E}$ by default. Nevertheless, if λ is set too large it will result in over-sparseness in $\mathcal{E}$, which means that only the high heat thermal mode is extracted without background reference. If λ is set too small, it will result in under-sparseness which subsequently causes a considerable amount of thermal pattern artefacts (or interference) remaining in $\mathcal{E}$. According to [21] advice for setting $\lambda = 1/\sqrt{T_{\max}(W,H)}$, existing algorithms always get over-sparse results. The proposed method uses $\mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} + \mathcal{E}$ as a final result of defect detection. $\mathcal{W} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}$

denotes background of supplementary certification defect detection or a weak tiny defection signal. $\mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}$ denotes complex and high thermal interference caused by strong noise and specimen characteristics. The hyperparameter we need to tune in our algorithm is the mode of core tensor, $r_1, r_2, r_3, \mathcal{L} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$. Combining setting λ as $1/\sqrt{r_3 \max(r_1, r_2)}$ with the additive representation of two tensors for the final result, the proposed method could robustly extract tiny and weak defection image signal.

(c) Optimization steps

In this paper, we will use augmented Lagrangian alternating direction minimization (AL-ADM) for tuning the model parameters. The T-SVD-based new tensor nuclear norm $\|\|_*$ as a low-rank constraint could better capture the global spatio-temporal information [21]. Thus, we adopt a similar tensor nuclear norm for the formulation of the proposed model. Reformulating the constraint optimization problem (2) into unconstraint formulation by AL-ADM [38] gives the following:

$$\begin{aligned} L_\rho(\mathcal{L}, \mathcal{G}, \mathcal{W}, \mathcal{E}, \boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \boldsymbol{U}^{(3)}; \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= \|\mathcal{G}\|_* + \lambda\|\mathcal{W}\|_1 \\ &+ \|\mathcal{E}\|_1 + \boldsymbol{\lambda}_1(\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}) + \boldsymbol{\lambda}_2(\mathcal{L} \\ &- \mathcal{G} - \mathcal{W}) + \frac{\rho}{2}[\|\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}\|^2 \\ &+ \|\mathcal{L} - \mathcal{G} - \mathcal{W}\|^2]. \end{aligned} \tag{2.3}$$

The above augmented Lagrangian function transforms the original equality constrained optimization problem into a more attainable one to solve the unconstrained optimization problem by adding Lagrangian multipliers $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$ and quadratic penalty term constraints to each equation with penalty coefficient ρ. The convergence speed of the algorithm can be accelerated by increasing ρ to a certain extent size [38]. Under the proximal operator and SVD decomposition, the above decoupled variables are solved by the alternating minimization method which yields a closed-form solution. The solution to (3) is obtained using the following steps.

(1) To update $\boldsymbol{U}^{(i)}$, for $i = 1, 2, 3$, and extract items related to $\boldsymbol{U}^{(i)}$, we just need to minimize the following formula:

$$\begin{aligned} \boldsymbol{U}^{(i)} &= \underset{\boldsymbol{U}^{(i)^{\mathrm{T}}}\boldsymbol{U}^{(i)}=\boldsymbol{I}}{\arg\min}\ \boldsymbol{\lambda}_1(\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}) \\ &\quad + \frac{\rho}{2}\|\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}\|^2 \\ &= \underset{\boldsymbol{U}^{(i)^{\mathrm{T}}}\boldsymbol{U}^{(i)}=\boldsymbol{I}}{\arg\min}\ \frac{\rho}{2}\left\|\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E} + \frac{\boldsymbol{\lambda}_1}{\rho}\right\|^2. \end{aligned} \tag{2.4}$$

Because $\boldsymbol{U}^{(1)}$, $\boldsymbol{U}^{(2)}$, $\boldsymbol{U}^{(3)}$ have similar variable status. Here we only deduce the optimization steps in detail for $\boldsymbol{U}^{(1)}$ and $\boldsymbol{U}^{(2)}$ and $\boldsymbol{U}^{(3)}$ can be analogized.

$$\begin{aligned} \boldsymbol{U}^{(1)} &= \underset{\boldsymbol{U}^{(1)^{\mathrm{T}}}\boldsymbol{U}^{(1)}=\boldsymbol{I}}{\arg\min}\ \left\|\boldsymbol{D}_{\langle 1\rangle} - \boldsymbol{U}^{(1)}\boldsymbol{L}_{\langle 1\rangle}\left[\boldsymbol{U}^{(3)} \bigotimes \boldsymbol{U}^{(2)}\right]^{\mathrm{T}} - \left(\boldsymbol{E} - \frac{\boldsymbol{\lambda}_1}{\rho}\right)_{\langle 1\rangle}\right\|^2 \\ &= \underset{\boldsymbol{U}^{(1)^{\mathrm{T}}}\boldsymbol{U}^{(1)}=\boldsymbol{I}}{\arg\max}\ \mathrm{Tr}\left[\boldsymbol{U}^{(1)^{\mathrm{T}}}\left(\boldsymbol{D} - \boldsymbol{E} + \frac{\boldsymbol{\lambda}_1}{\rho}\right)_{\langle 1\rangle}\boldsymbol{U}^{(3)} \bigotimes \boldsymbol{U}^{(2)}\boldsymbol{L}_{\langle 1\rangle}^{\mathrm{T}}\right] \end{aligned} \tag{2.5}$$

Let $P = (\boldsymbol{D} - \boldsymbol{E} + \frac{\boldsymbol{\lambda}_1}{\rho})_{\langle 1\rangle}\boldsymbol{U}^{(3)} \bigotimes \boldsymbol{U}^{(2)}\boldsymbol{L}_{\langle 1\rangle}^{\mathrm{T}}$ and $[\boldsymbol{U}_P, \boldsymbol{S}_P, \boldsymbol{V}_P] = \mathrm{SVD}(P)$. From the solution of the well-known orthogonal Procrustes problem [39]:

$$\boldsymbol{U}^{(1)} = \boldsymbol{U}_{P(1:r_1)}\boldsymbol{V}_{P(1:r_1)}^{\mathrm{T}}. \tag{2.6}$$

Similarly, the solution of

$$\boldsymbol{U}^{(2)}, \boldsymbol{U}^{(3)},$$

can be obtained as follows:

$$\left.\begin{aligned} P &= (\boldsymbol{D} - \boldsymbol{E} + \frac{\boldsymbol{\lambda}_1}{\rho})_{\langle 2\rangle} \boldsymbol{U}^{(3)} \bigotimes \boldsymbol{U}^{(1)} \boldsymbol{L}^{\mathrm{T}}_{\langle 2\rangle} \\ \boldsymbol{U}^{(2)} &= \boldsymbol{U}_{P(1:r_2)} \boldsymbol{V}^{\mathrm{T}}_{P(1:r_2)} \\ P &= \left(\boldsymbol{D} - \boldsymbol{E} + \frac{\boldsymbol{\lambda}_1}{\rho}\right)_{\langle 3\rangle} \boldsymbol{U}^{(2)} \bigotimes \boldsymbol{U}^{(1)} \boldsymbol{L}^{\mathrm{T}}_{\langle 3\rangle} \\ \text{and} \quad \boldsymbol{U}^{(3)} &= \boldsymbol{U}_{P(1:r_3)} \boldsymbol{V}^{\mathrm{T}}_{P(1:r_3)} \end{aligned}\right\} \tag{2.7}$$

(2) update $\mathcal{G}$

$$\mathcal{G} = \arg\min \|\mathcal{G}\|_* + \frac{\rho}{2} \left\| \mathcal{L} - \mathcal{G} - \mathcal{W} + \frac{\boldsymbol{\lambda}_2}{\rho} \right\|^2. \tag{2.8}$$

The above minimization problem is to compute the proximal operator of TNN [21]. It can be solved by tensor singular value thresholding (t-SVT) operators which is extension of matrix SVT and has closed-form solution. Based on T-SVD, Algorithm 1 give efficient T-SVT computing method. Now

$$\mathcal{G} = t - \mathrm{SVD}\left(\mathcal{L} - \mathcal{W} + \frac{\boldsymbol{\lambda}_2}{\rho}, \frac{1}{\rho}\right). \tag{2.9}$$

(3) Update $\mathcal{W}$

$$\mathcal{W} = \arg\min \lambda \|\mathcal{W}\|_1 + \frac{\rho}{2} \left\| \mathcal{L} - \mathcal{G} - \mathcal{W} + \frac{\boldsymbol{\lambda}_2}{\rho} \right\|^2. \tag{2.10}$$

From the well-known soft-thresholding algorithm

$$\mathcal{W} = S_{\lambda/\rho}\left(\mathcal{L} - \mathcal{G} + \frac{\boldsymbol{\lambda}_2}{\rho}\right), \tag{2.11}$$

where $S_{\lambda/\rho}(X) = \mathrm{sgn}(X)[|X| - (\lambda/\rho)]_+$.

(4) Update $\mathcal{L}$

$$\mathcal{L} = \arg\min \frac{\rho}{2} \left[\left\| \mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E} + \frac{\boldsymbol{\lambda}_1}{\rho} \right\|^2 + \left\| \mathcal{L} - \mathcal{G} - \mathcal{W} + \frac{\boldsymbol{\lambda}_2}{\rho} \right\|^2 \right]. \tag{2.12}$$

Above is a smooth convex optimization problem, thus we can obtain a closed-form solution:

$$\mathcal{L} = \frac{(\mathcal{D} - \mathcal{E} + (\boldsymbol{\lambda}_1/\rho)) \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} + \mathcal{G} + \mathcal{W} - (\boldsymbol{\lambda}_2/\rho)}{2}. \tag{2.13}$$

(5) Update $\mathcal{E}$

$$\mathcal{E} = \arg\min \|\mathcal{E}\|_1 + \frac{\rho}{2} \left\| \mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E} + \frac{\boldsymbol{\lambda}_1}{\rho} \right\|^2. \tag{2.14}$$

Similar with $\mathcal{W}$ by soft-thresholding under the proximal operator:

$$\mathcal{E} = S_{\frac{1}{\rho}}\left(\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} + \frac{\boldsymbol{\lambda}_1}{\rho}\right). \tag{2.15}$$

(6) Update Lagrange multipliers $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$

$$\left.\begin{aligned} \boldsymbol{\lambda}_1 &= \boldsymbol{\lambda}_1 + \rho(\mathcal{D} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}) \\ \text{and} \quad \boldsymbol{\lambda}_2 &= \boldsymbol{\lambda}_2 + \rho(\mathcal{L} - \mathcal{G} - \mathcal{W}). \end{aligned}\right\} \tag{2.16}$$

Summarize all the above AL-ADM optimization processes to get algorithm 2.

Algorithm 1. Tensor singular value thresholding: t-SVT($\mathcal{Y}, \tau$).

Input: $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times N}$; constant $\tau > 0$;
Output: $\mathcal{X}$
$\overline{\mathcal{Y}} = fft(\mathcal{Y}, [\,], 3)$.
Perform matrix SVT on each frontal slice of $\overline{\mathcal{Y}}$ by
for $i = 1, \ldots, \lceil \frac{N+1}{2} \rceil$ **do**
$[\boldsymbol{U}, \boldsymbol{S}, \boldsymbol{V}] = \mathrm{SVD}(\overline{\boldsymbol{Y}}^{(i)})$;
$\overline{\boldsymbol{X}}^{(i)} = \boldsymbol{U}(\boldsymbol{S} - \tau)_+ \boldsymbol{V}^*$;
end for
for $i = \lceil \frac{N+1}{2} \rceil + 1, \ldots, N$ **do**
$\overline{\boldsymbol{X}}^{(i)} = [\overline{\boldsymbol{X}}^{(N-i+2)}]^*$;
end for
$\mathcal{X} = ifft(\overline{\mathcal{X}}, [\,], 3)$

Algorithm 2. Hierarchical low-rank and sparse tensor decomposition (HLSTD) algorithm.

Input: Spatial–Temporal Thermography Tensor $\mathcal{D}^{W \times H \times T}$ rank r_1, r_2, r_3 .
1: Initialize: $\epsilon = 1e-5, \rho_0 = 1e-4, \rho_{\max} = 1e10, \eta = 1.5, \boldsymbol{U}_0{}^{(i)} = \mathcal{L}_0 = \mathcal{G}_0 = \mathcal{W}_0 = \mathcal{E}_0$.
2: **while** not convergence **do**
3: Update $\boldsymbol{U}^{(i)}$ by (6) and (7) ;
4: Update $\mathcal{G}$ by (9) ;
5: Update $\mathcal{W}$ by (11) ;
6: Update $\mathcal{L}$ by (13) ;
7: Update $\mathcal{E}$ by (15) ;
8: Update $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ by (16) ;
9: Update $\rho = \min(\rho_{\max}, \eta\rho)$;
10: Check the convergence conditions: $\|\mathcal{E}^{(k+1)} - \mathcal{E}^{(k)}\|_\infty < \epsilon$, $\|\mathcal{W}^{(k+1)} - \mathcal{W}^{(k)}\|_\infty < \epsilon$, $\|\mathcal{G}^{(k+1)} - \mathcal{G}^{(k)}\|_\infty < \epsilon$,$\|\mathcal{L} - \mathcal{G} - \mathcal{W}\|_\infty < \epsilon$ and $\|\mathcal{D}^{(k+1)} - \mathcal{L} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} - \mathcal{E}^{(k+1)}\|_\infty < \epsilon$
11: **end while**
Output: Defect Tensor $\mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)} + \mathcal{E}$.

3. Experimental

(a) Experimental set-up

The experimental schematic diagram of the proposed thermal imaging diagnostic system is shown in figure 2. Figure 3 shows the corresponding real-time practical experimental equipment and facilities. An Easyheat 224 from Cheltenham Induction Heating is used for coil excitation. The Easyheat has a maximum excitation power of 2.4 kW, a maximum current of 400 Arms and an excitation frequency range of 150–400 kHz (380 Arms and 256 kHz are used in this study). Water cooling of the coil is implemented to counteract direct heating of the coil. We use the IR camera (FLIR A655sc) and the frame rate is set to 100FPS to capture the thermal video sequences $\boldsymbol{D}^{W \times H \times T}$, which have two kind of resolution: $W \times H = 120 \times 640$ and $W \times H = 240 \times 640$. The excitation signal generated by the excitation module is a small period of high-frequency current. The current in the coil will induce the eddy currents and generate the resistive heat in the conductive material. The heat will diffuse in time until the heat reaches equilibrium in the material. Eddy current is forced to divert due to meeting abnormal gullies at the defects, which lead to areas of increased

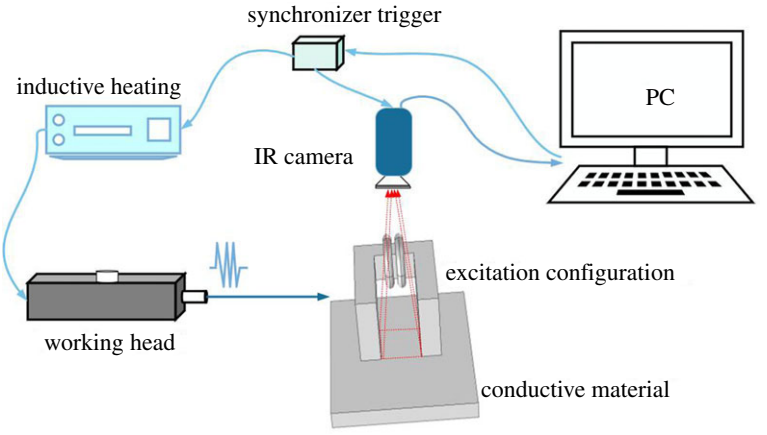

Figure 2. Inductive thermography schematic diagram. (Online version in colour.)

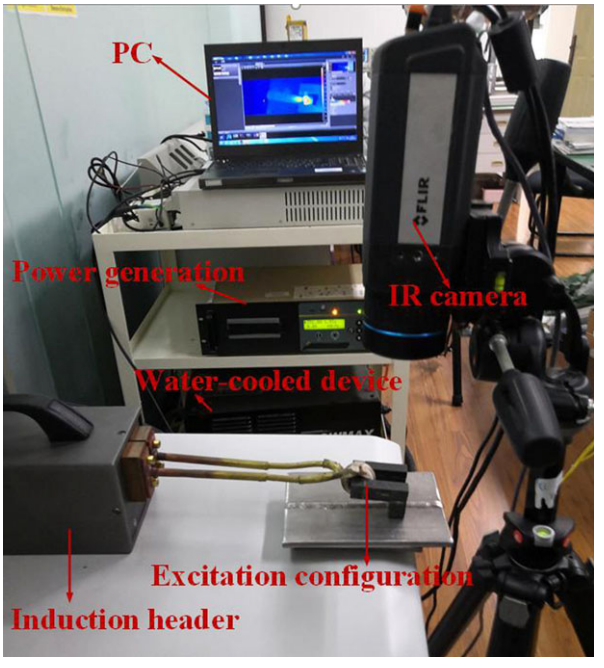

Figure 3. Practical IT experimental platform system. (Online version in colour.)

and decreased eddy current density. Therefore, in the heating phase, different areas have different heat generation rates which subsequently lead to temperature spatial-transient variation.

(b) Defective specimens

In order to illustrate the challenges faced by existing machine learning algorithms applied in inductive thermography non-destructive testing, we test them against the specimens which contain defects of irregular shapes. Specific test specimens are shown in figure 4. In particular, the complex characteristic of the rough surface of the specimens can be observed objectively but the micro cracks are invisible to the naked eye, which leads to enormous difficulty in defect detection.

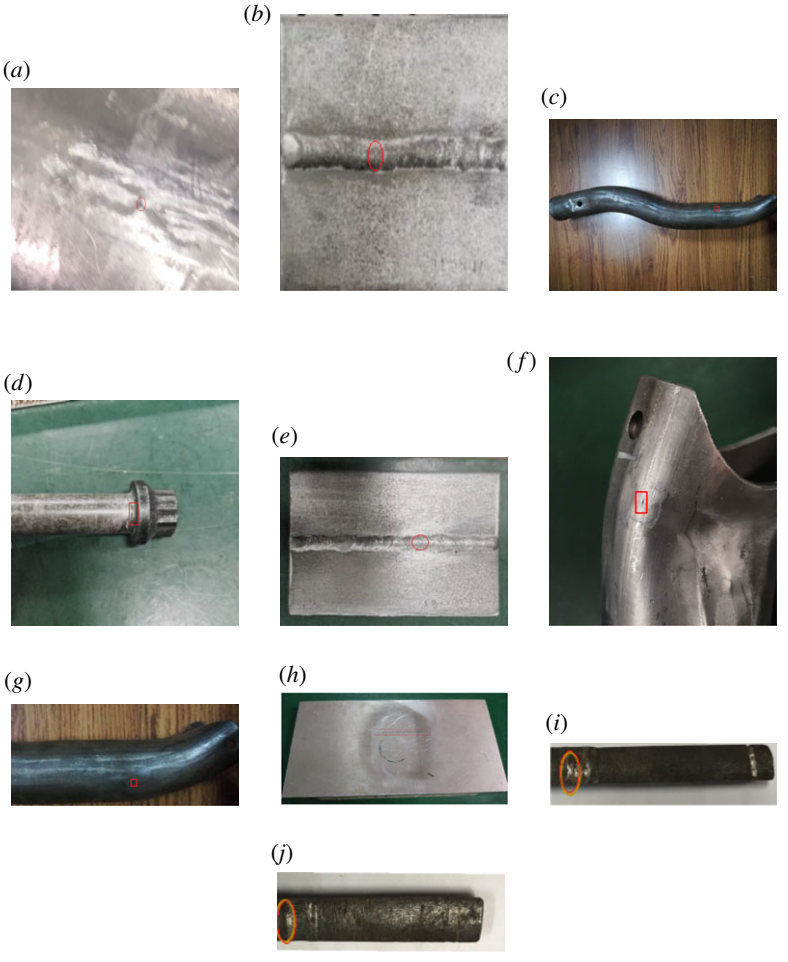

Figure 4. Defective specimen objects corresponding to figure 5. (*a*) pipeline1, (*b*) weld jointl, (*c*) pipeline2, (*d*) Axel, (*e*) weld joint2, (*f*) low-carbon steel, (*g*) pipeline3, (*h*), stainless steel, (*i*) weld joint3, (*j*) weld joint4. (Online version in colour.)

(c) Algorithm evaluation criteria

In order to validate the detection effectiveness of the proposed algorithm, we use the total computation time (as measured by the CPU) and signal-to-noise ratio (SNR) value as the numerical evaluation criteria. The SNR can evaluate the thermal contrast ratio between the defective and non-defective regions as calculated according to $\mathrm{SNR} = 20lg(T_d/T_{\mathrm{non}})$, where T_d is the temperature of all pixels in the local thermal image 1–1 in figure 5, T_{non} is the temperature of all pixels in the local thermal image 1–2 within the valid heating area near the defect. The size of area 1–2 is the same as area 1–1. The SNR of the entire thermal image is the average of the SNR calculated for all cracks. From the definition of SNR criteria, we can conclude intuitively that the more obvious the contrast between the defect and the surrounding of non-defect, the higher the SNR value. In other words, the higher the SNR value, the better the algorithm performance.

4. Results analysis

Since the paper focuses on the important application of a back-end processing algorithm in IT NDT, we select Tucker decomposition [40], GMRTF [41], BRTF [27] and TRPCA [21] for comparative verification, which are the latest algorithms in foreground background separation

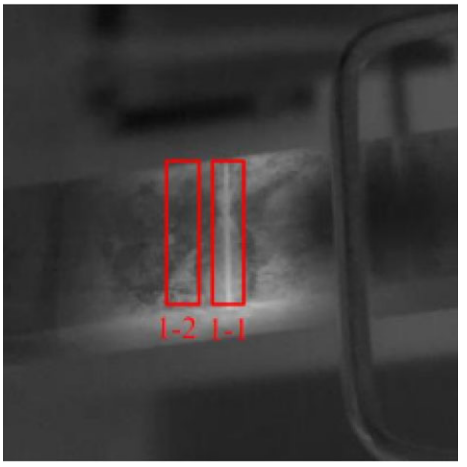

Figure 5. Illustration of SNR. (Online version in colour.)

and sparse anomaly pattern detection. Subsequent experiments show that our algorithm can obtain more ideal and efficient results than other algorithms in an induction thermography detection system. In order to conduct fair comparison of running time, all image recovery experiments were run in the same platform, with Matlab 2018a under windows10 on a PC of a 3.30 GHz i7-4590CPU and 16G RAM. The experimental parameters of each comparison algorithm either follow the default recommended settings in the paper or we adjust it to allow it to perform well in most cases. Visual results are listed in figure 6; SNR values and total CPU running time are shown table 1.

(a) Comparison

In order to illustrate the difficulty and challenge of our special defect task detection in the field of non-destructive testing, we first show the processing results of the traditional feature extraction algorithm PCA [42], sparse PCA (SPCA) [43], the latest online sparse matrix algorithm SGSM-BS [24] and incremental tensor decomposition IMTSL [32] algorithm in figure 7. More detailed experimental results are shown in the electronic supplementary material. As observed from the above results, not only classical methods, PCA and SPCA but also IMTSL and SGSB-BS fail in the specific detection task as shown in figure 4*a–d*. On the other hand, the proposed method has successfully detected the defects accurately where the defects' positions have been marked within the red boxes. In terms of energy concentration capabilities, the resulting images of the proposed algorithm show that the defects region (annotated with red boxes) are most centrally identified. According to the property of infrared thermography, the brighter the pixel, the higher the temperature, which shows that these positions gather more energy. Combined with the description of energy concentration above, it can be observed that the proposed algorithm obtains a higher T_d value and lower T_{non} value than other algorithms. In other words, the proposed method achieves higher SNR from the visual image results. Both the proposed algorithm and IMTSL (IHOSVD) are able to retain the details of the background of the test specimens. However, comparing with the IMTSL, the proposed algorithm could detect the defects with considerably higher SNR and improved visual inspection. In addition, the IMTSL is very time-consuming, which cannot be tolerated in practical industrial applications as measured by the 'run time', which is provided in electronic supplementary material.

It is worth mentioning that various advanced pulse compression favourable non-periodic thermal excitation techniques are in use for improving sub-surface defect detection and resolution

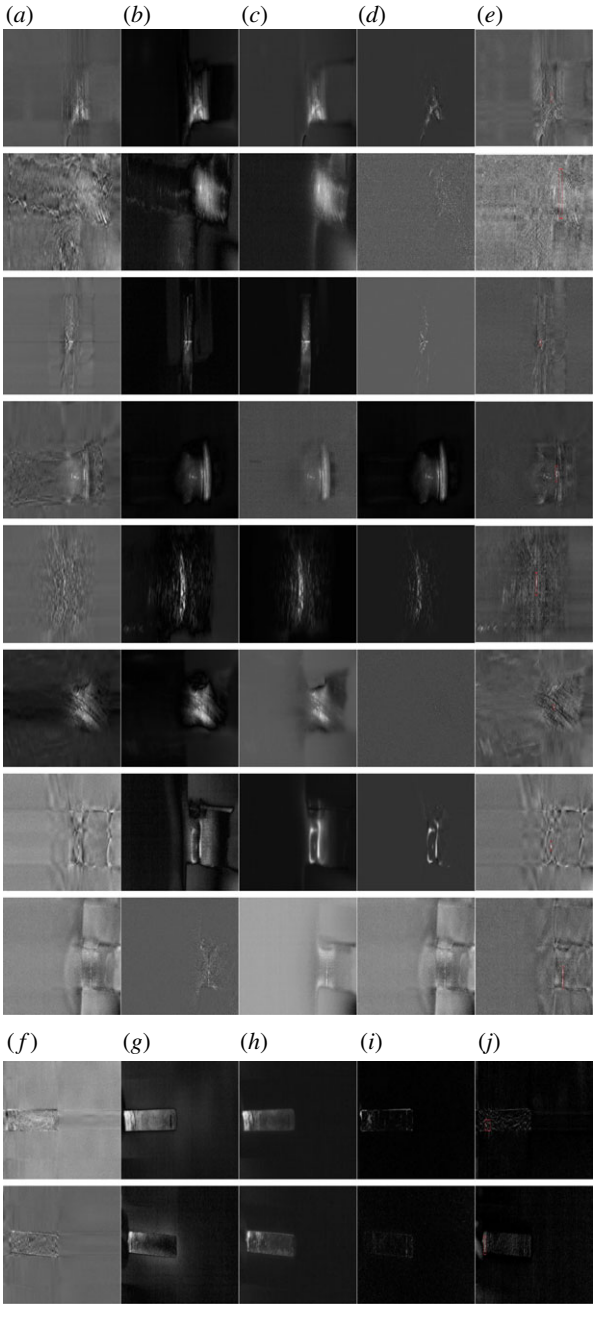

Figure 6. Contrast of image visual effect in defect detection for IT (zoom in for better visual effect). Each column from left to right presents results of GMRTF (*a*) and (*f*), TUCKER (*b*) and (*g*), BRTF (*b*) and (*f*), TRPCA (*d*) and (*i*), proposed (*e*) and (*j*), respectively. The order of each row from top to bottom corresponds to (*a–j*) sample specimens of figure 4. (Online version in colour.)

in the field of infrared thermal wave imaging which also makes full use of heat map post-processing algorithms to detect defects. In addition, there are different forms of excitation schemes such as eddy current, ultrasound and laser excited thermographic techniques. Kaur *et al.* [44] proposed PCA and SPCA based post-processing schemes for improving spatial contrast over the defective regions and enhancing the SNR, which was applied in frequency-modulated thermal

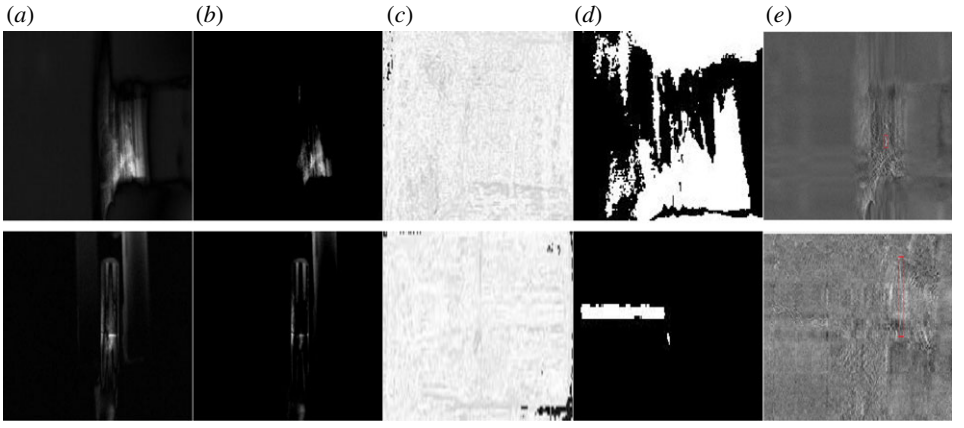

Figure 7. Comparison results (*a*) PCA (average SNR = −9.2, time = 14.4 s). (*b*) SPCA (average SNR = −11.5, time = 15.2 s) (*c*) IMTSL (average SNR = −4.75, time = 122.8 s) (*d*) SGSM-BS (average SNR = NaN, time = 10.1 s) (*e*) Proposed (average SNR = 8.35, time = 5.2 s). The top is Pipeline1 specimen, and at the bottom is Pipeline2 specimen. (Online version in colour.)

wave imaging for inspection of steel material. As shown in figure 7*a*,*b*, traditional methods such as PCA, SPCA and ICA perform inadequately in defect detection, especially with moving test specimens. By contrast, the proposed algorithm yields considerable improvement over the traditional methods and has demonstrated high accuracy of defect detection using ECPT under the effect of moving speed. This is shown in figures 7 and 6 using samples 9 and 10, respectively.

From figure 6 and table 1, it is not difficult to judge whether visual effect, SNR or running time value. The proposed algorithm shows the best performance among all methods. For the GMRTF method, it assumes that the detection videos are corrupted by noise with unknown distribution, in which they integrated the low-rank tensor decomposition with a mixture of Gaussian noise. To some extent, this is in line with the characteristics of IT data with strong noise and does not match with certain statistical laws. From the experiment results, high heat energy and strong noise could be eliminated to enhance defects. However, in the test of sample 1, sample 3, sample 4 and sample 6, the results of background heat elimination of GMRTF cannot be compared with the performance of the proposed algorithm. The SNR values are −1.411 and −14.588 for sample 1 and sample 3, respectively. Negative SNR means the heat energy of the non-defect region is higher than the defect region in which one may mistakenly recognize the high heat non-defect as a defect from the image. On the other hand, these optimizations are based on the EM algorithm, which is used to update parameters under the framework of probability. This leads to GMRTF not being efficient compared with the proposed method. This prevents the application of this algorithm in large-scale real-time industrial detection.

With regard to Tucker decomposition, it only uses low-rank decomposition, and it cannot get rid of the strong thermal background of the defect. It is difficult to distinguish the defects with the naked eye from the visual results. At the same time, it failed to enhance the contrast effect of defects. Tucker decomposition runs in a short time and saves 1–2 s in the experimental dataset. In addition, the visual results of BRTF are similar to Tucker decomposition. But more importantly, the parameters are updated under a full Bayesian inference framework, with the expansion of real data capacity as it is not scalable.

Tucker and BRTF visual results in sparse decomposition, because sparse defections are still embedded in high thermal background. However, for TRPCA, it usually gets over-sparse results. From visuals of samples 1, 2, 3, 6, 9, 10, there exist too few pixels to judge whether they are defects or not as it lacks of sufficient background information from specimens. What is worse,

Table 1. Comparative results SNR (left) and total running time (right in seconds) in figure 4.

	GMRTF		TUCKER		BRTF		TRPCA		proposed	
video name	SNR	time	SNR	time	SNR	time	SNR	time	SNR	time
pipeline1 (240 × 640 × 200)	−1.4	387.1	3.5	7.8	1.8	208.3	9.6	601.3	**12.9**	**5.1**
weld jointl (120 × 640 × 150)	0.8	128.1	1.9	3.8	1.4	105.9	0.8	227.5	**2.3**	**2.7**
pipeline2 (120 × 640 × 200)	−14.6	528.9	2.1	8.6	−6.4	205.2	−1.2	597.7	**4.8**	**5.2**
axel (240 × 640 × 200)	2.4	207.3	−1.7	7.0	−2.1	158.7	−1.2	449.6	**11.7**	**4.1**
weld joint2 (240 × 640 × 200)	6.1	127.9	5.3	3.7	6.5	98.2	8.0	220.1	**12.6**	**3.5**
low-carbon steel (240 × 640 × 150)	8.3	466.7	−11.5	8.3	−13.1	204.2	−11.5	608.6	**9.4**	**5.1**
pipeline3 (120 × 640 × 200)	5.4	355.1	5.1	8.4	2.3	209.7	1.6	601.1	**6.4**	**5.0**
stainless steel (120 × 640 × 200)	2.9	191.6	3.5	5.7	2.7	147.5	2.9	418.4	**4.2**	**3.8**
weld joint3 (240 × 640 × 150)	1.4	361.2	−9.0	7.3	−12.8	210.3	1.9	609.6	**7.1**	**5.2**
weld joint4 (240 × 640 × 200)	2.0	453.0	2.2	10.1	−1.3	209.0	8.5	614.1	**13.4**	**5.4**

royalsocietypublishing.org/journal/rsta *Phil. Trans. R. Soc. A* **378**: 20190584

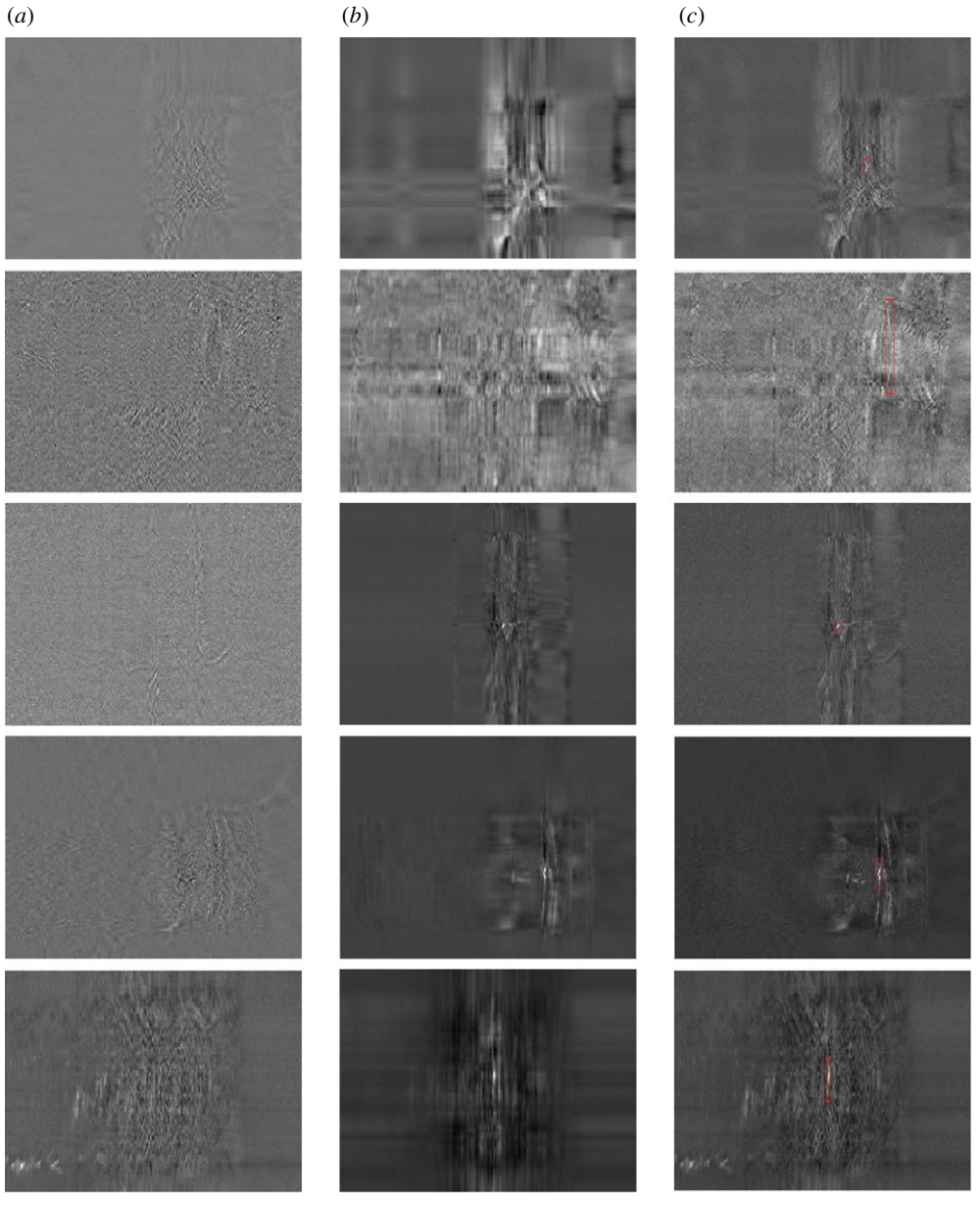

Figure 8. Robustness of HLSTD. (*a*) $\mathcal{C}_1$, (*b*) $\mathcal{C}_2$, (*c*) $\mathcal{C}_1$+$\mathcal{C}_2$. (Online version in colour.)

the algorithm requires a lot of matrix SVD operations, which consumes a lot of memory and computing resources and makes the speed of processing super slow.

In terms of robustness, the comparison algorithm can detect defects on parts of specimens. For natural weld cracks, this corresponds to test samples from 8 to 10. A natural weld crack is a challenging task in practical inspection, because an uneven surface is full of bumps and holes, as well as extremely irregular shape. In 9–10 samples, all the comparison algorithms fail to detect the weld crack, where the proposed HLSTD detects it with high-resolution and obvious enhancement. In terms of running time, the proposed algorithm gets the fastest speed with average 4.5 s.

(b) Robustness of the proposed HLSTD

The robustness of the proposed HLSTD algorithm is measured as the effectiveness of hierarchical decomposition for defect detection in the presence of complex background and strong noise. Specifically, defect information may be extracted in $\mathcal{E}$ of the first layer or in $\mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}$ of the second layer. The algorithm cannot predict in advance which part of the two components the defect will fall into. We integrate two parts to ensure that the defect can be extracted accurately and effectively, and the interference of noise and high-energy background is discarded. By relying only on a certain individual component, there is no guarantee that the algorithm can accurately extract the target defects. Hence the need for the proposed algorithm to incorporate the hierarchical nature of the decomposition in order to obtain robust defect detection. The following comparison of the intermediate experimental results shows the robustness of the HLSTD algorithm. For the sake of brevity, we mark $\mathcal{E}$ and $\mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}$ with $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. In figure 8, images from top to bottom correspond to the first five test pieces in figure 5, respectively.

Rank r1, r2, r3 are all set as 20, 20, 10 and $\lambda = 1/\sqrt{r_3 \max(r_1, r_2)}$ for all specimens in the proposed method. Defects lie in different component $\mathcal{C}_i$ for different specimens, $i = 1, 2$. It is observed that the first two defects lie in $\mathcal{C}_1$, and the last three lie in $\mathcal{C}_2$. The first two defects cannot be judged by combining the background information well because the image only contains the bright spot information in the state of over sparse. Because of the low-rank representation, the image showing the defect in the $\mathcal{C}_2$ component does not retain the background information well and hence it resulted in a blurred visual effect. The integration of $\mathcal{C}_1$ and $\mathcal{C}_2$ not only robustly extracts the defect with the background information preserved, but also produces clearer visual results.

5. Conclusion

In this paper, we have proposed a novel HLSTD algorithm. We have also developed it for IT for non-destructive defect detection. Natural crack defects in a variety of specimens with irregular shape have been used and validated by different advanced unsupervised decomposition methods and subsequently compared with the proposed algorithm. The proposed algorithm can accurately and efficiently suppress high heat background and strong noise while separating and enhancing the visual defects embedded in the thermal video stream. Both visual effects and numerical results have verified the robustness and efficiency of the proposed algorithm for defects detection.

Data accessibility. This article has no additional data.

Authors' contributions. T.W. contribute the idea of algorithm of the paper; B.G. and T.W. contribute the writing of the paper and experiment validation. W.l.W. contribute the refine of the paper.

Competing interests. We declare we have no competing interest.

Funding. No funding has been received for this article.

Acknowledgements. The work was supported by National Natural Science Foundation of China (grant nos. 61971093, 61527803 and 61960206010), by Science and Technology Department of Sichuan, China (grant nos. 2019YJ0208, 2018JY0655 and 2018GZ0047) and by Fundamental Research Funds for the Central Universities (grant no. ZYGX2019J067).

References

1. Goumeidane AB, Khamadja M, Odet C. 2007 Parametric active contour for boundary estimation of weld defects in radiographic testing. In *Int. Symp. on Signal Processing Its Applications*. Piscataway, NJ: IEEE.
2. Krautkrämer J, Krautkrämer H. 1990 *Ultrasonic testing of materials*. Berlin, Germany: Springer.
3. Singh R. 2020 Magnetic particle testing. In *Applied welding engineering*, 3rd edn, pp. 331–338. Amsterdam, The Netherlands: Elsevier.
4. Mix PE. 2004 Liquid penetrant tests. In *Introduction to nondestructive testing: a training guide* (2nd edn), chapter 6. Hoboken, NJ: Wiley.

5. Gao B, Woo WL, He Y, Tian GY. 2016 Unsupervised sparse pattern diagnostic of defects with inductive thermography imaging system. *IEEE Trans. Ind. Inf.* **12**, 371–383. (doi:10.1109/TII.2015.2492925)
6. Gao B, Bai L, Tian GY, Woo WL, Cheng Y. 2013 Spatial and time patterns extraction of eddy current pulsed thermography using blind source separation. *IEEE Sensors J.* **13**, 2094–2101. (doi:10.1109/JSEN.2013.2248931)
7. Cheng L, Gao B, Tian GY, Woo WL, Berthiau G. 2014 Impact damage detection and identification using eddy current pulsed thermography through integration of PCA and ICA. *IEEE Sensors J.* **14**, 1655–1663. (doi:10.1109/JSEN.2014.2301168)
8. Genest M, Dudzinski DC, Dawag L, Kersey RK. 2013 Crack detection using induction thermography during high temperature testing. *Proc. SPIE* **8705**, 87050P. (doi:10.1117/12.2015488)
9. JAckel P, Netzelmann U. 2013 The influence of external magnetic fields on crack contrast in magnetic steel detected by induction thermography. *Quant. Infrared Thermogr. J.* **10**, 237–247. (doi:10.1080/17686733.2013.852414)
10. He Y, Tian GY, Pan M, Chen D, Hong Z. 2014 An investigation into eddy current pulsed thermography for detection of corrosion blister. *Corros. Sci.* **78**, 1–6. (doi:10.1016/j.corsci.2013.09.001)
11. Candes E, Li X, Ma Y, Wright J. 2011 Robust principal component analysis. *J. ACM.* **58**, 11. (doi:10.1145/1970392.1970395)
12. Vaswani N, Bouwmans T, Javed S, Narayanamurthy P. 2018 Robust subspace learning: robust PCA, robust subspace tracking, and robust subspace recovery. *IEEE Signal Process. Mag.* **35**, 32–55. (doi:10.1109/MSP.2018.2826566)
13. Vaswani N, Chi Y, Bouwmans T. 2018 Rethinking PCA for modern data sets: theory, algorithms, and applications [scanning the issue]. *Proc. IEEE* **106**, 1274–1276. (doi:10.1109/JPROC.2018.2853498)
14. Bouwmans T, Sobral A, Javed S, Jung SK, Zahzah EH. 2017 Decomposition into low-rank plus additive matrices for background/foreground separation: a review for a comparative evaluation with a large-scale dataset. *Computer Rev.* **23**, 1–71. (doi:10.1016/j.cosrev.2016.11.001)
15. Han N, Song Y, Song Z. 2017 Bayesian robust principal component analysis with structured sparse component. *Comput. Stat. Data Anal.* **109**, 144–158. (doi:10.1016/j.csda.2016.12.005)
16. Netrapalli P, Niranjan UN, Sanghavi S, Anandkumar, Animashree. 2014 Provable non-convex robust PCA. *Adv. Neural Inf. Process. Syst.* **2**, 1107–1115.
17. Lois B, Vaswani N. 2015 Online matrix completion and online robust PCA. In *IEEE Int. Symp. Information Theory (ISIT)*, pp. 1826–1830. (doi:10.1109/ISIT.2015.7282771)
18. Gao B, Woo W, Dlay S. 2012 Variational regularized 2-d nonnegative matrix factorization. *IEEE Trans. Neural Netw. Learning Syst.* **23**, 703–716. (doi:10.1109/TNNLS.2012.2187925)
19. Gao B, Peng L, Woo WL, Tian GY, Johnston M. 2018 Variational Bayesian sub-group adaptive sparse component extraction for diagnostic imaging system. *IEEE Trans. Ind. Electron.* **65**, 8142–8152. (doi:10.1109/TIE.2018.2801809)
20. Wang Y, Gao B, Woo WL, Tian GY, Zhu Y. 2018 Thermal pattern contrast diagnostic of micro cracks with induction thermography for aircraft braking components. *IEEE Trans. Ind. Informatics* **14**, 5563–5574. (doi:10.1109/TII.2018.2802046)
21. Lu C, Feng J, Chen Y, Liu W, Lin Z, Yan S. 2018 Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE Trans. Pattern Analysis Machine Intell.* **42**, 925–938. (doi:10.1109/TPAMI.2019.2891760)
22. Gu Q, Wang Z, Liu H. 2016 Low-rank and sparse structure pursuit via alternating minimization. International conference on artificial intelligence and statistics, AISTATS. *AISTATS* **2016**, 600–609.
23. Ma M, Hu R, Chen S, Xiao J, Wang Z. 2018 Robust background subtraction method via low-rank and structured sparse decomposition. *China Commun.* **15**, 156–167. (doi:10.1109/CC.2018.8424611)
24. Shi G, Huang T, Dong D, Wu J, Xie X. 2018 Robust foreground estimation via structured Gaussian scale mixture modeling. *IEEE Trans. Image Process.* **27**, 4810–4824. (doi:10.1109/TIP.2018.2845123)

25. Javed S, Mahmood A, Al-Maadeed S, Bouwmans T, Jung SK. 2019 Moving object detection in complex scene using spatiotemporal structured-sparse RPCA. *IEEE Trans. Image Process.* **28**, 1007–1022. (doi:10.1109/TIP.2018.2874289)
26. Cichocki A, Mørup M, Smaragdis P, Wang W, Zdunek R. 2008 Advances in nonnegative matrix and tensor factorization. *Comput. Intell. Neurosci.* **2008**, 852187. (doi:10.1155/2008/852187)
27. Zhao Q, Zhou G, Zhang L, Cichocki A, Amari S. 2017 Bayesian robust tensor factorization for incomplete multiway data. *IEEE Trans. Neural Networks Learn. Syst.* **27**, 736–748. (doi:10.1109/TNNLS.2015.2423694)
28. Zhou P, Feng J. 2017 Outlier-robust tensor PCA. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI*, pp. 3938–3946. (doi:10.1109/CVPR.2017.419)
29. Xue J, Zhao Y, Liao W, Chan J. 2019 Nonconvex tensor rank minimization and its applications to tensor recovery. *Inf. Sci.* **503**, 109–128. (doi:10.1016/j.ins.2019.06.061)
30. Chen W, Gong X, Song N. 2019 Nonconvex robust low-rank tensor reconstruction via an empirical Bayes methods. *IEEE Trans. Signal Process.* **67**, 5785–5797. (doi:10.1109/TSP.2019.2946022)
31. Javed S, Bouwmans T, Jung SK. 2015 Stochastic decomposition into low rank and sparse tensor for robust background subtraction. In *6th Int. Conf. on Imaging for Crime Prevention and Detection (ICDP-15), London*, pp. 1–6. (doi:10.1049/ic.2015.0105)
32. Campilho A, Kamel M, Sobral A, Baker C, Bouwmans T, Zahzah E. 2014 Incremental and multi-feature tensor subspace learning applied for background modeling and subtraction. In *Image Analysis and Recognition*. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-11758-4_11)
33. Qiu C, Wu X, Xu H. 2014 Recursive Projected Sparse Matrix Recovery (ReProCSMR) With Application In Real-Time Video Layer Separation. In *IEEE Int. Conf. on Image Processing*, pp. 1332–1336. (doi:10.1109/ICIP.2014.7025266)
34. Gao B, Yin A, Wang Y, Tian GY, Woo WL, Liu H. 2014 Thermography spatial-transient-stage tensor model and material property characterization. In *IEEE Far East Forum on Nondestructive Evaluation/Testing, Chengdu, 2014*, pp. 199–203. (doi:10.1109/FENDT.2014.6928262)
35. Gao B, He Y, Lok Woo W, Tian GY, Liu J, Hu Y. 2016 Multidimensional tensor-based inductive thermography with multiple physical fields for offshore wind turbine gear inspection. *IEEE Trans. Industrial Elect.* **63**, 6305–6315. (doi:10.1109/TIE.2016.2574987)
36. Lu P, Gao B, Feng Q, Yang Y, Tian GY. 2017 Ensemble variational Bayes tensor factorization for super resolution of CFRP debond detection. *Infrared Phys. Technol.* **85**, 335–346. (doi:10.1016/j.infrared.2017.07.012)
37. Zhang C, Fu H, Si L, Liu G, Cao X. 2015 Low-rank tensor constrained multiview subspace clustering. In *IEEE Int.Conf. on Computer Vision. Santiago, 2015*, pp. 1582–1590. (doi:10.1109/ICCV.2015.185)
38. Lin Z, Chen M, Ma Y. 2010 The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. eprint arxiv, 9.
39. Gower JC, Dijksterhuis GB. 2004 Orthogonal Procrustes problems. In *Procrustes problems*. Oxford, UK: Oxford University Press. (doi:10.1093/acprof:oso/9780198510581.003.0004)
40. Kolda TG, Bader BW. 2009 Tensor decompositions and applications. *Siam Rev.* **51**, 455–500. (doi:10.1137/07070111X)
41. Chen X, Han Z, Wang Y, Zhao Q, Meng D, Tang Y. 2016 Robust tensor factorization with unknown noise. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV*, pp. 5213–5221. (doi:10.1109/CVPR.2016.563)
42. Jolliffe IT. 2002 *Principal component analysis*, 2nd edn. Berlin, Germany: Springer.
43. Cai T, Ma Z, Wu Y. 2013 Recent results on sparse principle component analysis. 2013, *5th IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP 2013*, pp. 181–183. (doi:10.1109/CAMSAP.2013.6714037)
44. Kaur K, Mulaveesala A. 2019 An efficient data processing approach for frequency modulated thermal wave imaging for inspection of steel material. *Infrared Phys. Technol.* **103**, 103083. (doi:10.1016/j.infrared.2019.103083)