REVISED Manuscript (Clean version in word or latex Format) Click here to view linked References

Mechanical Systems and Signal Processing, Volume 204, December 2023, Article Number 110761
DOI: 10.1016/j.ymssp.2023.110761

Tensor robust principal component analysis based on Bayesian Tucker decomposition for thermographic inspection

Yue Hu[a,b], Fangsen Cui[c], Yifan Zhao[d], Fucai Li[e,*], Shuai Cao[c], and Fu-zhen Xuan[a,b,*]

[a] School of Mechanical and Power Engineering, East China University of Science and Technology, Shanghai, China
[b] Shanghai Key Laboratory of Intelligent Sensing and Detection Technology, East China University of Science and Technology, Shanghai, China
[c] Institute of High Performance Computing, A*STAR, Singapore, Singapore
[d] School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK
[e] State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China

* Corresponding authors.
E-mail addresses: fcli@sjtu.edu.cn (F. Li), fzxuan@ecust.edu.cn (F. Xuan).

Published by Elsevier. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 4.0).
The final published version (version of record) is available online at DOI:10.1016/j.ymssp.2023.110761. Please refer to any applicable publisher terms of use.

Abstract: Thermographic inspection is considered an effective and promising nondestructive testing tool because of its intuitiveness, wide range and noncontact property. Despite this, the detection of weak defects and the recovery of their shape remain difficult, particularly when the surface being inspected is the opposite of the surface being drilled. This study proposes a new tensor robust principal component analysis method based on Bayesian Tucker decomposition to improve the spatial resolution of thermography. A hierarchical form of a generalized Student-t prior is imposed on the model parameters in the Bayesian framework so as to approximate the low-rank component related to the defect feature. Through variational Bayesian inference, all model parameters are adaptively estimated. Based on two experimental data, it appears that the proposed method is capable of improving the spatial resolution and detection accuracy of the thermographic inspection system.

Keywords: Tensor robust principal component analysis, Bayesian Tucker decomposition, thermographic inspection, nondestructive testing, generalized Student-t distribution

1. Introduction

Nondestructive testing (NDT) is a powerful tool to guarantee structural integrity, reduce system downtime and lower unexpected repair costs. Due to its intuitive, wide-range, and noncontact characteristics, thermography technique has gained considerable attention among NDT methods[1–3]. During thermographic inspection, thermal images reveal temperature variations within a material or object. These variations can indicate the presence of defects. However, background noise and complex heat diffusion may cause defect feature weak and unclear in the thermal image [4]. Hence, signal processing algorithms are necessary to extract the defect features from thermal images.

The defect feature in thermal images exhibits a low-rank property that has been discovered in recent studies, and researchers have developed robust principal component analysis (RPCA) to identify such low-rank components [5]. The RPCA decomposes a signal into two components, one of which is a low-rank component and the other sparse. The RPCA has been investigated by Momeni [9] as one of the high-dimensional data analytics methods in structural health monitoring and destructive evaluation. Gao *et al.* [10] proposed a variational Bayesian method that used sub-sparse grouping and adaptive fine-tuning strategies to improve thermographic imaging. However, this method failed to detect the subsurface defect. To remove the strong noise in the thermal data, Ahmed [8] proposed a sparse ensemble matrix factorization (SEMF) method for the detection of debonding in composite. Ahmed *et al.* [5] developed an ensemble joint sparse low-rank matrix decomposition (EJSLRMD) approach based on an expectation-maximization framework. In the EJSLRMD method, the low-rank and sparse components are jointly modelled in a concatenated feature space, and are estimated in an iterative manner. The EJSLRMD could detect the inner debonding on carbon fiber reinforced plastics (CFRP) composites at the cost of a high computation complexity. Moreover, a new formulation of the RPCA, Orthogonal Inexact Augmented Lagrangian Multiplier (OIALM), was established to detect defects and discontinuities in CFRP structures [11].

The above matrix-based methods required reshaping the thermographic sequences, which are three-dimensional data, into a matrix before determining the defect features. However, such a data-reshaping process may lead to performance degradation and information loss [12]. For the multi-dimensional data, it makes sense to employ the counterpart of the RPCA, tensor RPCA (TRPCA). Some new TRPCA methods have emerged which include improved robust tensor principal component analysis (IRTPCA), tensor singular value thresholding (t-SVT) [12], tensor-tensor product [14]. Motivated by the above variants, some TRPCA-based methods have been developed for the thermographic inspection. Gao *et al.* [15] constructed a physics-based multidimensional tensor model for offshore wind turbine gear inspection by using thermography. Although this method could successfully detect the contact fatigue defect for the gear, it costs extensive computational time and requires high data storage space. To analyze the thermal data acquired from the irregular shape specimens, Ahmed *et al.* [16] proposed a low-rank tensor with a sparse mixture of Gaussian (LRTSMoG) decomposition algorithm. The LRTSMoG algorithm extracts the low-rank and sparse components using the t-SVT algorithm [12] and the MoG matrix factorization [17], respectively. To separate defect features from complex and strong noise, Wu *et al.* [18] proposed a hierarchical low-rank and sparse tensor decomposition (HLSTD) algorithm that is based on higher-order singular value decomposition. The HLSTD method uses Augmented Lagrangian Alternating Direction Minimization to adaptively estimate the model parameters and extract the weak and small defects from the complex thermal data. The HLSTD method was validated on a variety of metallic specimens with irregular shapes. However, the HLSTD method required tuning hyperparameters to obtain a good performance. To lower the computational cost of tensor-based algorithms, Bai *et al.* [19] introduced the global Moran's index in the tensor decomposition for thermal data processing. Most of the above methods need repeatedly tuning hyperparameters or model parameters to obtain a good performance, which may improve the computational time and reduce the adaptivity.

In this regard, some researchers considered to develop Bayesian TRPCA methods to automatedly determine the model parameters and improve the adaptivity for the TRPCA-based methods. The Tucker [23] and Candecomp/Parafac (CP) [24] decompositions are two widely used frameworks in the Bayesian TRPCA methods. Based on CP decomposition, Zhao *et al.* [25] proposed the Bayesian robust tensor factorization (BRTF) method for processing the image sequence. In the BRTF method, the low-rank tensor denotes the global information, and the sparse tensor denotes the local information like outliers. The BRTF method constructs a variational Bayesian framework then avoids repeatedly tuning model parameters. Motivated by the BRTF method [25], Lu *et al.* [26] proposed an ensemble variational Bayes tensor factorization (EVBTF) method for detection of debonding in composites. Compared with the CP decomposition, the Tucker decomposition allows for more control over the rank of the decomposition and can handle tensors with different modes of variation. However, few people consider using the Tucker decomposition in the Bayesian TRPCA.

In this paper, a new tensor robust principal component analysis method is proposed to improve the spatial resolution and detection accuracy of the thermography system. Contributions of this work are listed as follows:

(1) A new Bayesian Tucker decomposition model is constructed for the TRPCA to extract weak features in the thermal images.

(2) The generalized student-t distribution, a much sparser prior, and its hierarchical model are developed to fully exploit the low-rank property.

(3) By using sparse Bayesian learning, all model parameters and latent variables are assigned with proper priors and adaptively determined. Furthermore, some intractable variables' computational equations are derived by using Bessel functions and confluent hypergeometric functions.

(4) The proposed method is able to extract the weak defect feature from the thermal image sequences, and its effectiveness is verified by two experimental tests.

The rest of this paper is organized as follows. Some notations and preliminaries are introduced in Section II. The proposed

method is described in Section III, and its performance is investigated by an experimental test in Section IV. Finally, the work is concluded in Section V.

2. Notations and preliminaries

In this study, a tensor is denoted by boldface calligraphic letters like $\boldsymbol{\mathcal{X}}$. For a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its $(i_1, i_2, \ldots, i_N)$th element is denoted by $\mathcal{X}_{i_1 i_2 \ldots i_N}$, where $i_n = 1,2, \cdots, I_n$ and $n = 1,2, \cdots, N$. The mode-n matricization $\boldsymbol{\mathcal{X}}$ of a tensor is denoted by $\mathbf{X}_{(n)}$. The tensor n-mode product, multiplying a tensor $\boldsymbol{\mathcal{X}}$ by a matrix $\mathbf{U}$ in mode n, is denoted by $\boldsymbol{\mathcal{X}} \times {}_n\mathbf{U}$ and the tensor n-mode product can be calculated in a matrix form:

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} \times {}_n\mathbf{U} \Leftrightarrow \mathbf{Z}_{(n)} = \mathbf{U}\mathbf{X}_{(n)} \tag{1}$$

The Tucker decomposition is defined as:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{G}} \times {}_1\mathbf{U}^{(1)} \times {}_2\mathbf{U}^{(2)} \times \cdots \times {}_N\mathbf{U}^{(N)} \tag{2}$$

where $\boldsymbol{\mathcal{G}} \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_N}$ is the core tensor and $\left\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}\right\}$ is a set of factor matrices. The Tucker decomposition can also be represented in the following forms:

$$\begin{aligned} \mathbf{X}_{(n)} &= \mathbf{U}^{(n)}\mathbf{G}_{(n)}\left(\bigotimes_{k \neq n} \mathbf{U}^{(k)T}\right) \\ \mathrm{vec}(\boldsymbol{\mathcal{X}}) &= \left(\bigotimes_{n} \mathbf{U}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}}) \\ \mathcal{X}_{i_1 \ldots i_N} &= \left(\bigotimes_{n} \mathbf{u}_{i_n}^{(n)T}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}}) \end{aligned} \tag{3}$$

where $\mathrm{vec}(\cdot)$ converts a matrix or a tensor into a column vector and $\mathbf{u}_{i_n}^{(n)}$ denotes the row vector of the matrix $\mathbf{U}^{(n)}$. $\bigotimes_{n} \mathbf{U}^{(n)}$ and $\bigotimes_{k \neq n} \mathbf{U}^{(k)}$ denote the sequential Kronecker products and are defined as:

$$\begin{aligned} \bigotimes_{n} \mathbf{U}^{(n)} &= \mathbf{U}^{(N)} \otimes \mathbf{U}^{(N-1)} \otimes \cdots \otimes \mathbf{U}^{(1)} \\ \bigotimes_{k \neq n} \mathbf{U}^{(k)} &= \mathbf{U}^{(N)} \cdots \otimes \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n-1)} \otimes \cdots \otimes \mathbf{U}^{(1)} \end{aligned} \tag{4}$$

where $\otimes$ denotes the Kronecker product.

3. Proposed method

A. Generalized Student-t distribution

In sparse Bayesian learning, the Laplace distribution is usually employed to impose the sparsity constraint. A much sparser alternative, the generalized Student-t distribution [27] is used in this study, and the generalized Student-t distribution is defined as:

$$GSt(x|\eta,\gamma) = \frac{1}{\sqrt{2\pi\gamma}} \frac{\Gamma\left(\eta + \frac{1}{2}\right)}{\Gamma(\eta)} \left(1 + \frac{x^2}{2\gamma}\right)^{-\left(\eta + \frac{1}{2}\right)} \tag{5}$$

where η and γ are parameters. $\Gamma(\cdot)$ is the Gamma function.

Moreover, the generalized Student-t prior can be written as a hierarchical model consisting of a normal distribution and a Gamma distribution:

$$\begin{aligned} GSt(x|\eta,\gamma) &= \int p(x|\lambda)p(\lambda|\eta,\gamma)\mathrm{d}\lambda \\ p(x|\lambda) &= N(x;0,1/\lambda) \\ p(\lambda|\eta,\gamma) &= Ga(\lambda;\eta,\gamma) \end{aligned} \tag{6}$$

The normal distribution and the Gamma distribution are defined as:

$$N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{7}$$

$$Ga(x;\alpha,\beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \tag{8}$$

where μ and σ are the mean and the standard deviation of the normal distribution respectively, α and β are the shape and scale parameters of the Gamma distribution.

Fig. 1 gives the comparison among the generalized Student-t distribution with different parameters and the Laplace distribution. Compared with the Laplace distribution, the generalized Student-t distribution has better sparse enforcement by changing its parameters η and γ. For the generalized Student-t distribution, if the parameter η is fixed, the parameter γ can be lowered to generate a sparser distribution. By comparing different generalized Student-t distributions, the constraints for the parameters η and γ are empirically set in this study:

$$\begin{aligned} \eta &= 1 \\ 0 < \gamma &\leq 0.2 \end{aligned} \tag{9}$$

Then, the hierarchical model of the generalized Student-t distribution is rewritten as:

$$\begin{aligned} p(x|\lambda) &= N(0,1/\lambda) \\ p(\lambda|\eta,\gamma) &= Ga(\lambda;1,\gamma) \end{aligned} \tag{10}$$

The parameter γ is assigned with a beta distribution, which is supported by a bounded interval:

$$p(\gamma) = p(\gamma|\alpha_\gamma, \beta_\gamma) = \mathcal{B}e(5\gamma|\alpha_\gamma, \beta_\gamma) \tag{11}$$

where α_γ and β_γ are parameters of the beta distribution. The beta distribution is defined as:

$$\mathcal{B}e(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \tag{12}$$

where $B(\alpha, \beta)$ is the beta function, α and β are its shape parameters.

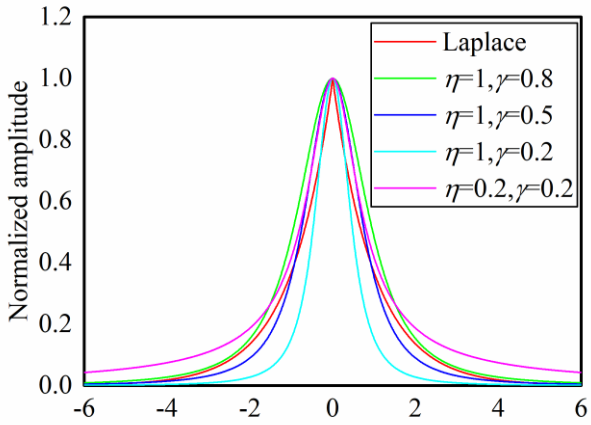


Fig. 1 Comparisons among the Laplace distribution and different generalized Student-t distributions.

B. Probabilistic model

The thermographic sequence is denoted by a three-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ and $\boldsymbol{\mathcal{X}}$ can be decomposed into a low-rank tensor $\boldsymbol{\mathcal{L}}$, a sparse tensor $\boldsymbol{\mathcal{S}}$, and the noise $\boldsymbol{e}$:

$$\begin{aligned} \boldsymbol{\mathcal{X}} &= \boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{S}} + \boldsymbol{e} \\ \boldsymbol{\mathcal{L}} &= \boldsymbol{\mathcal{G}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \end{aligned} \tag{13}$$

where $\boldsymbol{\mathcal{G}} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ and $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}$. In the subsequent analysis, the tensor order N is equal to 3.

Based on the observation model (Eq. (13)), the likelihood function of $\boldsymbol{\mathcal{X}}$ is expressed as:

$$\begin{aligned} &p\left(\boldsymbol{\mathcal{X}}|\{\mathbf{U}^{(n)}\}, \boldsymbol{\mathcal{G}}, \boldsymbol{\mathcal{S}}, \rho\right) = p\left(\mathrm{vec}(\boldsymbol{\mathcal{X}})|\{\mathbf{U}^{(n)}\}, \boldsymbol{\mathcal{G}}, \boldsymbol{\mathcal{S}}, \rho\right) \\ &= N\left(\mathrm{vec}(\boldsymbol{\mathcal{X}}); \left(\bigotimes_n \mathbf{U}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}}) + \mathrm{vec}(\boldsymbol{\mathcal{S}}), \rho^{-1}\mathbf{I}\right) \end{aligned} \tag{14}$$

where ρ is the noise precision and is assigned with a Gamma prior:

$$p(\rho) = p(\rho|\alpha_\rho, \beta_\rho) = Ga(\rho; \alpha_\rho, \beta_\rho) \tag{15}$$

where α_ρ and β_ρ are the shape and scale parameters used in the Gamma distribution.

R_1, R_2, and R_3 are dimensions of the core tensor $\boldsymbol{\mathcal{G}}$, and also column numbers of $\{\mathbf{U}^{(n)}\}$. To minimize the tensor rank, the sparsity is enforced over columns of $\mathbf{U}^{(n)}$. Moreover, a generalized Student-t prior is placed over the matrix $\mathbf{U}^{(n)}$:

$$\begin{aligned} p\left(\mathbf{U}^{(n)}|\boldsymbol{\lambda}^{(n)}\right) &= \prod_{r_n} p\left(\mathbf{u}_{\cdot r_n}^{(n)}|\lambda_{r_n}^{(n)}\right) = \prod_{r_n} N\left(0, \left(\lambda_{r_n}^{(n)}\right)^{-1}\mathbf{I}\right) p\left(\boldsymbol{\lambda}^{(n)}|\boldsymbol{\gamma}^{(n)}\right) \\ &= \prod_{r_n} Ga\left(\lambda_{r_n}^{(n)}|1, \gamma_{r_n}^{(n)}\right) p\left(\boldsymbol{\gamma}^{(n)}\right) = p\left(\boldsymbol{\gamma}^{(n)}|\alpha_\gamma, \beta_\gamma\right) = \prod_{r_n} \mathcal{B}e\left(5\gamma_{r_n}^{(n)}|\alpha_\gamma, \beta_\gamma\right) \end{aligned} \tag{16}$$

where $\mathbf{u}_{\cdot r_n}^{(n)}$ denotes the column vector of the matrix $\mathbf{U}^{(n)}$. $\lambda_{r_n}^{(n)}$ and $\gamma_{r_n}^{(n)}$ are the r_nth element of $\boldsymbol{\lambda}^{(n)}$ and $\boldsymbol{\gamma}^{(n)}$.

In the Tucker decomposition, $\mathcal{G}_{r_1 r_2 r_3}$ is regarded as the weight of the vectors, $\mathbf{u}_{\cdot r_1}^{(1)}$, $\mathbf{u}_{\cdot r_2}^{(2)}$, and $\mathbf{u}_{\cdot r_3}^{(3)}$. Then the sparsity constraint, a generalized Student-t prior, is enforced over $\boldsymbol{\mathcal{G}}$:

$$\begin{aligned} p\left(\boldsymbol{\mathcal{G}}|\left\{\{\boldsymbol{\lambda}^{(n)}\}, \eta\right\}\right) &= \prod_{r_1, r_2, r_3} p\left(\mathcal{G}_{r_1 r_2 r_3}|\left\{\lambda_{r_1}^{(1)}, \lambda_{r_2}^{(2)}, \lambda_{r_3}^{(3)}, \eta\right\}\right) \\ &= \prod_{r_1, r_2, r_3} N\left(0, \left(\eta \prod_n \lambda_{r_n}^{(n)}\right)^{-1}\right) p\left(\boldsymbol{\lambda}^{(n)}|\boldsymbol{\gamma}^{(n)}\right) = \prod_{r_n} Ga\left(\lambda_{r_n}^{(n)}|1, \gamma_{r_n}^{(n)}\right) \end{aligned} \tag{17}$$

where η is the weight parameter and is assigned with a Gamma prior:

$$p(\eta) = p(\eta|\alpha_\eta, \beta_\eta) = Ga(\eta; \alpha_\eta, \beta_\eta) \tag{18}$$

where α_η and β_η are parameters of the Gamma distribution.

The sparse tensor $\boldsymbol{\mathcal{S}}$ is also placed over a hierarchical model of a generalized Student-t prior:

$$\begin{aligned} p(\boldsymbol{\mathcal{S}}|\boldsymbol{\xi}) &= \prod_m p(\mathrm{vec}(\boldsymbol{\mathcal{S}})_m|\xi_m) = \prod_m N(0, \xi_m^{-1})\, p(\boldsymbol{\xi}|\boldsymbol{\tau}) = \prod_m p(\xi_m|\tau_m) \\ &= \prod_m Ga(\xi_m|1, \tau_m)\, p(\boldsymbol{\tau}) = p(\boldsymbol{\tau}|\alpha_\tau, \beta_\tau) = \prod_m \mathcal{B}e(5\tau_m|\alpha_\tau, \beta_\tau) \end{aligned} \tag{19}$$

where $\mathrm{vec}(\boldsymbol{\mathcal{S}})_m$, ξ_m, and τ_m are the mth element of $\mathrm{vec}(\boldsymbol{\mathcal{S}})$, $\boldsymbol{\xi}$, and $\boldsymbol{\tau}$. α_τ and β_τ are the parameters of the beta distribution.

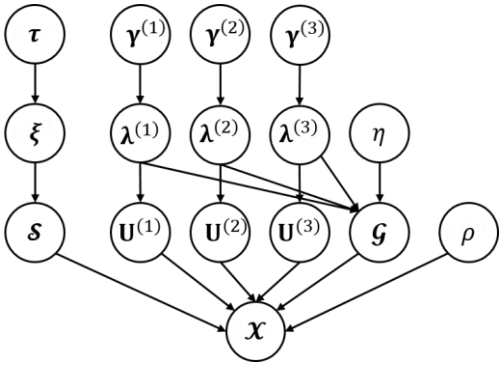


Fig. 2 Probabilistic graph model of the proposed method.

Finally, **Fig. 2** gives the probabilistic graph model of the proposed method and the joint distribution is expressed as:

$$p(\mathcal{X},\boldsymbol{\theta}) = p\left(\mathcal{X}|\{\mathbf{U}^{(n)}\},\mathcal{G},\mathcal{S},\rho\right)\left(\prod_n p\left(\mathbf{U}^{(n)}|\boldsymbol{\lambda}^{(n)}\right)\right)p\left(\mathcal{G}|\{\boldsymbol{\lambda}^{(n)},\eta\}\right)$$
$$\left(\prod_n p\left(\boldsymbol{\lambda}^{(n)}|\boldsymbol{\gamma}^{(n)}\right)\right)p(\mathcal{S}|\xi)p(\xi|\boldsymbol{\tau})p(\boldsymbol{\gamma})p(\boldsymbol{\tau})p(\eta)p(\rho) \tag{20}$$

The set $\boldsymbol{\theta}$ consists of all unknown variables and is defined as $\boldsymbol{\theta} = \left[\mathcal{G},\{\mathbf{U}^{(n)}\},\{\boldsymbol{\lambda}^{(n)}\},\{\boldsymbol{\gamma}^{(n)}\},\eta,\mathcal{S},\xi,\boldsymbol{\tau},\rho\right]$. Then the variational inference is used to estimate all unknown variables.

C. Variational inference

In the variational inference, a distribution $q(\boldsymbol{\theta})$ is used to approximate the posterior distribution $p(\boldsymbol{\theta}|\mathbf{x})$. The distribution is defined as a factorized form:

$$q(\boldsymbol{\theta}) = q(\mathcal{G})\left(\prod_n q\left(\mathbf{U}^{(n)}\right)\right)\left(\prod_n q\left(\boldsymbol{\lambda}^{(n)}\right)\right)q(\mathcal{S})$$
$$q(\xi)q(\eta)\left(\prod_n q\left(\boldsymbol{\gamma}^{(n)}\right)\right)q(\boldsymbol{\tau})q(\rho) \tag{21}$$

The unknown variable can be estimated by using the mean-field approximation:

$$\mathrm{In}q(\boldsymbol{\theta}_k) = \mathrm{E}_{q(\boldsymbol{\theta}\backslash\boldsymbol{\theta}_k)}\left[\mathrm{In}\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right] \tag{22}$$

where $\boldsymbol{\theta}\backslash\boldsymbol{\theta}_k$ denotes the set $\boldsymbol{\theta}$ with $\boldsymbol{\theta}_k$ removed. In each iteration, the posterior $q(\boldsymbol{\theta}_k)$ is updated and other posteriors $q(\boldsymbol{\theta}\backslash\boldsymbol{\theta}_k)$ are fixed.

1) Estimation of $\mathcal{G}$

Based on Eq. (22), the posterior of the vectorized $\mathcal{G}$, $\mathrm{vec}(\mathcal{G})$, is identified as a normal distribution (details are given in Section 1 of Appendix):

$$q\left(\mathrm{vec}(\mathcal{G})\right) = N\left(\mathrm{vec}(\mathcal{G})|\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{G})},\boldsymbol{\Sigma}_{\mathrm{vec}(\mathcal{G})}\right) \tag{23}$$

where $\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{G})}$ and $\boldsymbol{\Sigma}_{\mathrm{vec}(\mathcal{G})}$ denote the means vector and the covariance matrix and are expressed as:

$$\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{G})} = \mathrm{E}(\rho)\boldsymbol{\Sigma}_{\mathrm{vec}(\mathcal{G})}\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)T}\right)\right)\left(\mathrm{vec}(\mathcal{X}) - \boldsymbol{\mu}_{\mathrm{vec}(\mathcal{S})}\right) \tag{24}$$

$$\boldsymbol{\Sigma}_{\mathrm{vec}(\mathcal{G})} = \left(\mathrm{E}(\rho)\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)T}\mathbf{U}^{(n)}\right) + \mathrm{E}(\eta)\bigotimes_n \mathrm{E}\left(\boldsymbol{\Lambda}^{(n)}\right)\right)^{-1} \tag{25}$$

Then $\mathrm{vec}(\mathcal{G})$ is estimated as:

$$\mathrm{E}\left(\mathrm{vec}(\mathcal{G})\right) = \boldsymbol{\mu}_{\mathrm{vec}(\mathcal{G})} \tag{26}$$

2) Estimation of $\mathbf{U}^{(n)}$

The logarithmic approximate posterior of $\mathbf{U}^{(n)}$ is

$$\mathrm{In}q\left(\mathbf{U}^{(n)}\right) = \mathrm{E}_{q(\boldsymbol{\theta}\backslash\mathbf{U}^{(n)})}\left[\mathrm{In}\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right]$$
$$= -\frac{1}{2}\mathrm{tr}\left(\begin{array}{c}\mathbf{U}^{(n)}\left(\mathrm{E}(\rho)\mathrm{E}\left(\mathbf{G}_{(n)}\left(\bigotimes_{k\neq n}\mathbf{U}^{(k)T}\mathbf{U}^{(k)}\right)\mathbf{G}_{(n)}^T\right) + \mathrm{E}\left(\boldsymbol{\Lambda}^{(n)}\right)\right)\mathbf{U}^{(n)T} \\ -2\rho\left(\mathbf{X}_{(n)} - \mathrm{E}\left(\mathbf{S}_{(n)}\right)\right)\left(\bigotimes_{k\neq n}\mathrm{E}\left(\mathbf{U}^{(k)}\right)\right)\mathrm{E}\left(\mathbf{G}_{(n)}^T\right)\mathbf{U}^{(n)T}\end{array}\right) + \mathrm{const} \tag{27}$$

Then the expected value for $\mathbf{U}^{(n)}$ is given as:

$$\mathrm{E}\left(\mathbf{U}^{(n)}\right) = \mathrm{E}(\rho)\left(\mathbf{X}_{(n)} - \mathrm{E}\left(\mathbf{S}_{(n)}\right)\right)\bigotimes_{k\neq n}\mathrm{E}\left(\mathbf{U}^{(k)}\right)\mathrm{E}\left(\mathbf{G}_{(n)}^T\right)\boldsymbol{\Phi}_{\mathbf{U}^{(n)}} \tag{28}$$

where $\boldsymbol{\Phi}_{\mathbf{U}^{(n)}}$ is defined as:

$$\boldsymbol{\Phi}_{\mathbf{U}^{(n)}} = \left(\mathrm{E}(\rho)\mathrm{E}\left(\mathbf{G}_{(n)}\left(\bigotimes_{k\neq n}\mathbf{U}^{(k)T}\mathbf{U}^{(k)}\right)\mathbf{G}_{(n)}^T\right) + \mathrm{E}\left(\boldsymbol{\Lambda}^{(n)}\right)\right)^{-1} \tag{29}$$

3) Estimation of $\boldsymbol{\lambda}^{(n)}$

For the vector $\boldsymbol{\lambda}^{(n)}$, its element $\lambda_{r_n}^{(n)}$ is independent. Based on Eq. (22), the posterior of $\boldsymbol{\lambda}^{(n)}$ is identified as a Gamma distribution (details are given in Section 2 of Appendix):

$$q\left(\boldsymbol{\lambda}^{(n)}\right)=\prod\prod Ga\left(\lambda_{r_n}^{(n)}|\alpha_{q\left(\lambda_{r_n}^{(n)}\right)},\beta_{q\left(\lambda_{r_n}^{(n)}\right)}\right) \tag{30}$$

where the parameters $\alpha_{q\left(\lambda_{r_n}^{(n)}\right)}$ and $\beta_{q\left(\lambda_{r_n}^{(n)}\right)}$ are given as:

$$\begin{aligned}\alpha_{q\left(\lambda_{r_n}^{(n)}\right)}&=\frac{I_n+\prod_{k\neq n}R_k}{2}+1\\ \beta_{q\left(\lambda_{r_n}^{(n)}\right)}&=\frac{1}{2}\mathrm{E}\left(\mathbf{u}_{:r_n}^{(n)T}\mathbf{u}_{:r_n}^{(n)}\right)+\frac{\mathrm{E}(\eta)}{2}\sum_{r_l,r_m}\mathrm{E}\left(\lambda_{r_l}^{(l)}\right)\mathrm{E}\left(\lambda_{r_m}^{(m)}\right)\mathrm{E}\left(\mathcal{G}_{r_l r_m r_n}^2\right)+\mathrm{E}\left(\gamma_{r_n}^{(n)}\right)\end{aligned} \tag{31}$$

Then $\lambda_{r_n}^{(n)}$ is estimated as:

$$\mathrm{E}\left(\lambda_{r_n}^{(n)}\right)=\frac{\alpha_{q\left(\lambda_{r_n}^{(n)}\right)}}{\beta_{q\left(\lambda_{r_n}^{(n)}\right)}} \tag{32}$$

4) Estimation of $\boldsymbol{\gamma}^{(n)}$

$\gamma_{r_n}^{(n)}$ is r_nth element of the vector $\boldsymbol{\gamma}^{(n)}$ and its logarithmic approximate posterior is given as:

$$\mathrm{In}q\left(\gamma_{r_n}^{(n)}\right)=\mathrm{E}_{q\left(\boldsymbol{\theta}\backslash\gamma_{r_n}^{(n)}\right)}\left[\mathrm{In}\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right]=-\mathrm{E}\left(\lambda_{r_n}^{(n)}\right)\gamma_{r_n}^{(n)}+\alpha_\gamma\mathrm{In}\left(\gamma_{r_n}^{(n)}\right)+\left(\beta_\gamma-1\right)\mathrm{In}\left(1-5\gamma_{r_n}^{(n)}\right)+\mathrm{const} \tag{33}$$

It is difficult to find a proper distribution to represent the posterior distribution of $\gamma_{r_n}^{(n)}$. In this case, $\gamma_{r_n}^{(n)}$ is estimated by directly calculating the integral $\int\gamma_{r_n}^{(n)}q\left(\gamma_{r_n}^{(n)}\right)\mathrm{d}\gamma_{r_n}^{(n)}$. Though this integral is intractable, the Bessel function and the confluent hypergeometric function of the first kind are used to calculate the estimate of $\gamma_{r_n}^{(n)}$:

$$\mathrm{E}\left(\gamma_{r_n}^{(n)}\right)=\int\gamma_{r_n}^{(n)}q\left(\gamma_{r_n}^{(n)}\right)\mathrm{d}\gamma_{r_n}^{(n)}=\frac{1}{5}\frac{B\left(\beta_\gamma,\alpha_\gamma+2\right)\,{}_1F_1\left(\alpha_\gamma+2;\beta_\gamma+\alpha_\gamma+2;-\frac{\mathrm{E}\left(\lambda_{r_n}^{(n)}\right)}{5}\right)}{B\left(\beta_\gamma,\alpha_\gamma+1\right)\,{}_1F_1\left(\alpha_\gamma+1;\beta_\gamma+\alpha_\gamma+1;-\frac{\mathrm{E}\left(\lambda_{r_n}^{(n)}\right)}{5}\right)} \tag{34}$$

where ${}_1F_1(a;b;c)$ denotes the confluent hypergeometric function of the first kind.

5) Estimation of η

Based on Eq. (22), the posterior of η is identified as a Gamma distribution (details are given in Section 3 of Appendix):

$$q(\eta)=Ga\left(\eta|\alpha_{q(\eta)},\beta_{q(\eta)}\right) \tag{35}$$

where the parameters $\alpha_{q(\eta)}$ and $\beta_{q(\eta)}$ are given as:

$$\begin{aligned}\alpha_{q(\eta)}&=\frac{R_1R_2R_3}{2}+\alpha_\eta\\ \beta_{q(\eta)}&=\frac{1}{2}\mathrm{E}\left(\mathrm{vec}^T(\boldsymbol{\mathcal{G}})\left(\bigotimes_n\boldsymbol{\Lambda}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right)+\beta_\eta\end{aligned} \tag{36}$$

where $\mathrm{E}\left(\mathrm{vec}^T(\boldsymbol{\mathcal{G}})\left(\bigotimes_n\boldsymbol{\Lambda}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right)$ is calculated as:

$$\mathrm{E}\left(\mathrm{vec}^T(\boldsymbol{\mathcal{G}})\left(\bigotimes_n\boldsymbol{\Lambda}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right)=\mathrm{tr}\left(\mathrm{E}\left(\bigotimes_n\boldsymbol{\Lambda}^{(n)}\right)\left(\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})}\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})}^T+\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})}\right)\right) \tag{37}$$

Then η is estimated as:

$$\mathrm{E}(\eta)=\frac{\alpha_{q(\eta)}}{\beta_{q(\eta)}} \tag{38}$$

6) Estimation of $\boldsymbol{\mathcal{S}}$

Based on Eq. (22), the posterior of $\boldsymbol{\mathcal{S}}$ is identified as a Gamma distribution (details are given in Section 4 of Appendix):

$$q\left(\mathrm{vec}(\boldsymbol{\mathcal{S}})\right)=N\left(\mathrm{vec}(\boldsymbol{\mathcal{S}})|\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})},\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}\right) \tag{39}$$

where $\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}$ and $\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}$ denote the means vector and the covariance matrix and are expressed as:

$$\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}=\mathrm{E}(\rho)\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}\left(\mathrm{vec}(\boldsymbol{\mathcal{X}})-\bigotimes_n\mathrm{E}\left(\mathbf{U}^{(n)}\right)\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})}\right) \tag{40}$$

$$\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}=\left(\mathrm{E}(\rho)+\mathrm{E}(\Xi)\right)^{-1} \tag{41}$$

where $\Xi=\mathrm{diag}(\boldsymbol{\xi})$.

Then $\mathrm{vec}(\boldsymbol{\mathcal{S}})$ is estimated as:

$$\mathrm{E}\left(\mathrm{vec}(\boldsymbol{\mathcal{S}})\right)=\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})} \tag{42}$$

7) Estimation of $\boldsymbol{\xi}$

ξ_m is mth element of the vector $\boldsymbol{\xi}$. Based on Eq. (22), the posterior of $\boldsymbol{\xi}$ is identified as a Gamma distribution (details are given in Section 5 of Appendix):

$$q(\boldsymbol{\xi}) = \prod \prod Ga\left(\xi_m | \alpha_{q(\xi_m)}, \beta_{q(\xi_m)}\right) \tag{43}$$

where the parameters $\alpha_{q(\xi_m)}$ and $\beta_{q(\xi_m)}$ are given as:

$$\alpha_{q(\xi_m)} = 3/2$$
$$\beta_{q(\xi_m)} = \frac{\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})_m}^T \boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})_m} + \left(\boldsymbol{\Sigma}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}\right)_{m,m}}{2} + \mathrm{E}(\tau_m) \tag{44}$$

Then ξ_m is estimated as:

$$\mathrm{E}(\xi_m) = \frac{\alpha_{q(\xi_m)}}{\beta_{q(\xi_m)}} \tag{45}$$

8) Estimation of $\boldsymbol{\tau}$

τ_m is mth element of the vector $\boldsymbol{\tau}$ and its logarithmic approximate posterior is given as:

$$\mathrm{In}q(\tau_m) = \mathrm{E}_{q(\boldsymbol{\theta}\setminus\tau_m)}\left[\mathrm{In}\left(p(\mathbf{x}, \boldsymbol{\theta})\right)\right] = -\mathrm{E}(\xi_m)\tau_m + \alpha_\tau \mathrm{In}(\tau_m) + (\beta_\tau - 1)\mathrm{In}(1 - 5\tau_m) + \mathrm{const} \tag{46}$$

Then τ_m is estimated as:

$$\mathrm{E}(\tau_m) = \int \tau_m q(\tau_m) \mathrm{d}\tau_m = \frac{B(\beta_\tau, \alpha_\tau + 2)\ {}_1F_1\left(\alpha_\tau + 2; \beta_\tau + \alpha_\tau + 2; -\frac{\mathrm{E}(\xi_m)}{5}\right)}{5B(\beta_\tau, \alpha_\tau + 1)\ {}_1F_1\left(\alpha_\tau + 1; \beta_\tau + \alpha_\tau + 1; -\frac{\mathrm{E}(\xi_m)}{5}\right)} \tag{47}$$

9) Estimation of ρ

Based on Eq. (22), the posterior of ρ is identified as a Gamma distribution (details are given in Section 5 of Appendix):

$$q(\rho) = Ga\left(\rho | \alpha_{q(\rho)}, \beta_{q(\rho)}\right) \tag{48}$$

where the parameters $\alpha_{q(\rho)}$ and $\beta_{q(\rho)}$ are given as:

$$\alpha_{q(\rho)} = \alpha_\rho + \frac{I_1 I_2 I_3}{2} \beta_{q(\rho)} = \beta_\rho + \frac{1}{2} \mathrm{E}\left(\left\|\mathrm{vec}(\boldsymbol{\mathcal{X}}) - \mathrm{vec}(\boldsymbol{\mathcal{S}}) - \left(\bigotimes_n \mathbf{U}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right\|_F^2\right) \tag{49}$$

where $\mathrm{E}\left(\left\|\mathrm{vec}(\boldsymbol{\mathcal{X}}) - \mathrm{vec}(\boldsymbol{\mathcal{S}}) - \left(\bigotimes_n \mathbf{U}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right\|_F^2\right)$ is calculated as:

$$\begin{aligned}
&\mathrm{E}\left(\left\|\mathrm{vec}(\boldsymbol{\mathcal{X}}) - \mathrm{vec}(\boldsymbol{\mathcal{S}}) - \left(\bigotimes_n \mathbf{U}^{(n)}\right)\mathrm{vec}(\boldsymbol{\mathcal{G}})\right\|_F^2\right) \\
&= \mathrm{vec}^T(\boldsymbol{\mathcal{X}})\mathrm{vec}(\boldsymbol{\mathcal{X}}) + \mathrm{E}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}\left(\mathrm{vec}^T(\boldsymbol{\mathcal{S}})\mathrm{vec}(\boldsymbol{\mathcal{S}})\right) + \mathrm{tr}\left(\mathrm{E}\left(\mathrm{vec}(\boldsymbol{\mathcal{G}})\mathrm{vec}^T(\boldsymbol{\mathcal{G}})\right)\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)T}\mathbf{U}^{(n)}\right)\right)\right) \\
&-2\mathrm{vec}^T(\boldsymbol{\mathcal{X}})\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)}\right)\right)\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})} - 2\mathrm{vec}^T(\boldsymbol{\mathcal{X}})\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})} + 2\left(\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{S}})}\right)^T\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)}\right)\right)\boldsymbol{\mu}_{\mathrm{vec}(\boldsymbol{\mathcal{G}})}
\end{aligned} \tag{50}$$

Then ρ is estimated as:

$$\mathrm{E}(\rho) = \frac{\alpha_{q(\rho)}}{\beta_{q(\rho)}} \tag{51}$$

The whole steps of the proposed method are summarized in **Algorithm 1**.

Algorithm 1: Tensor robust principal component analysis based on Bayesian Tucker decomposition

Input: data tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, prior parameters α_η, β_η, α_γ, β_γ, α_τ, β_τ, α_ρ, and β_ρ.

Output: The low-rank tensor $\boldsymbol{\mathcal{L}}$.

Initialize: ρ, $\mathbf{U}^{(n)}$, $\boldsymbol{\mathcal{S}}$, $\boldsymbol{\gamma}^{(n)}$, η, and $\boldsymbol{\tau}$.

while *stopping criterion is not satisfied* **do**

  Update the posterior of $\boldsymbol{\mathcal{G}}$ by Eq. (24).

  for $n \leftarrow 1$ **to** 3

    Update the posterior of $\mathbf{U}^{(n)}$ by Eq.(28).

      for $r_n \leftarrow 1$ **to** R_n

        Update the posterior $\lambda_{r_n}^{(n)}$ by Eq. (32).

        Update the posterior $\gamma_{r_n}^{(n)}$ by Eq. (34).

      end for

  end for

  Update the posterior of η by Eq. (38).

  Update the posterior of $\boldsymbol{\mathcal{S}}$ by Eq. (42).

  Update the posterior of $\boldsymbol{\xi}$ by Eq. (45).

  Update the posterior of $\boldsymbol{\tau}$ by Eq. (47).

  Update the posterior of ρ by Eq. (51).

end while

The computational complexity of the proposed method is discussed. In each iteration, the computational complexity is $O(\prod_n R_n^3 + \prod_n I_n R_n)$ for estimating $\boldsymbol{\mathcal{G}}$, $O(R_n^3 + \prod_n I_n)$ for estimating $\mathbf{U}^{(n)}$, $O(I_n R_n + \prod_n R_n)$ for estimating $\boldsymbol{\lambda}^{(n)}$, $O(\prod_n R_n)$

for estimating $\boldsymbol{\gamma}^{(n)}$, $O(\prod_n R_n^2)$ for estimating η, $O(\prod_n I_n)$ for estimating $\boldsymbol{\delta}$, $O(\prod_n I_n)$ for estimating $\boldsymbol{\tau}$, and $O(\prod_n R_n + \prod_n I_n)$ for estimating ρ. Hence, the overall computational complexity of the proposed method is $O(K\prod_n R_n^3 + K\prod_n I_n R_n)$, where K is the maximum iteration.

4. Experimental verification

In order to evaluate the performance of the proposed method, two experiments are conducted.

Case 1: an aluminum plate

The first experiment is carried out on an aluminum plate and its set-up is shown in **Fig. 3** (a). The thermography system consists of two flash lamps, an infrared camera (SC7000), a control unit, and a desktop computer. The infrared camera is used to acquire the thermal data with a frame rate of 50 Hz. The specimen, a 5mm thick aluminum plate, has three drilled holes with a depth of 1mm and diameter of 32 mm, 16 mm, and 8 mm from left to right. That is, the holes are drilled at the bottom of the specimen as shown in **Fig. 3** (b). In the thermographic inspection, the specimen was placed in front of the IR camera at a distance of around 25 cm, and the surface of the specimen was kept perpendicular to the central line of the camera lens. The specimen is examined from the top surface rather than the drilled surface. These results in weak defect features and increases the difficulty of defect detection using thermal data.

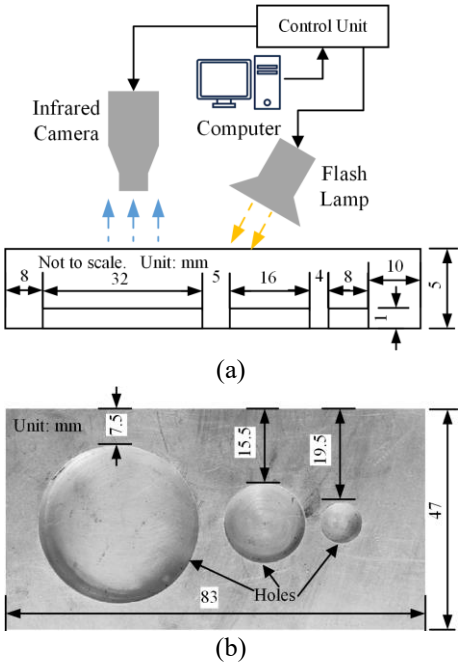


Fig. 3 (a) The experimental set-up of the thermographic inspection and (b) the bottom view of the specimen.

The proposed method is compared with the principal component thermography (PCT) [28], EVBTF, variational Bayesian sub-group adaptive sparse component extraction (VBAS) [10], IRTPCA [13], ensemble joint sparse low rank matrix decomposition (EJSLRMD) [5], TRPCA [12], and online mixture of Gaussians low rank matrix factorization (OMoGMF) [17] methods on defect detection. As the output of these methods is a sequence of maps, a quantitative indicator, signal-to-noise ratio (SNR), is used to intuitively choose the optimal image. The SNR is defined as:

$$\mathrm{SNR} = \frac{\left|\overline{P(n_x,n_y)_{\text{defective}}} - \overline{P(n_x,n_y)_{\text{intact}}}\right|}{\mathrm{Std}\left[P(n_x,n_y)_{\text{intact}}\right]} \tag{52}$$

where $\overline{P(n_x,n_y)_{\text{defective}}}$ and $\overline{P(n_x,n_y)_{\text{intact}}}$ are the mean pixel values in the defective and intact regions. $\mathrm{Std}\left[P(n_x,n_y)_{\text{intact}}\right]$ denotes the standard deviation of the pixel values in the intact region.

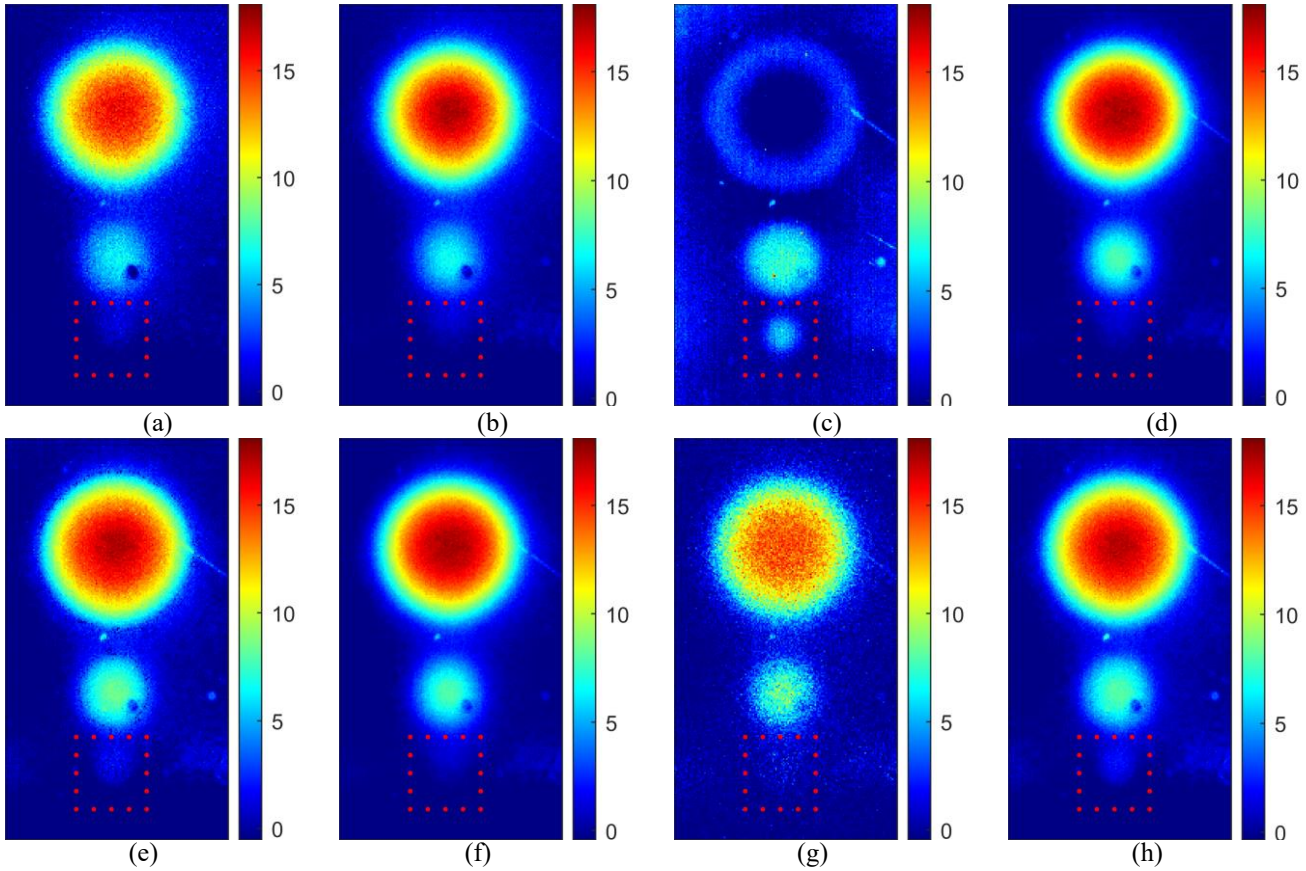


Fig. 4 presents the thermographic imaging results. The red dashed rectangles are drawn to highlight the smallest hole

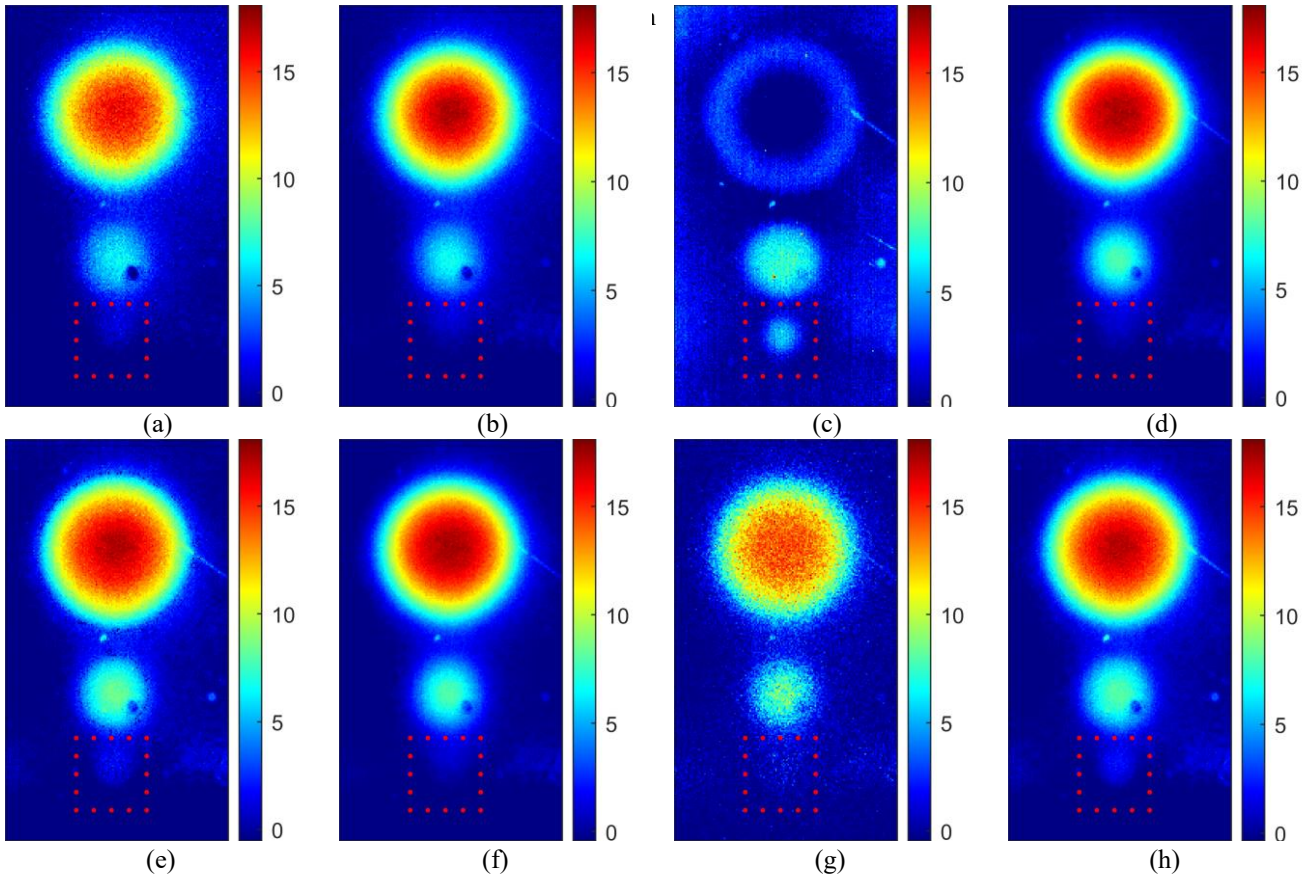


Fig. 4. Aside from the VBAS method, all methods can extract the first two largest holes. Despite this, detecting the smallest hole is very challenging since the inspected surface is opposed to the drilled surface. It is indicated from

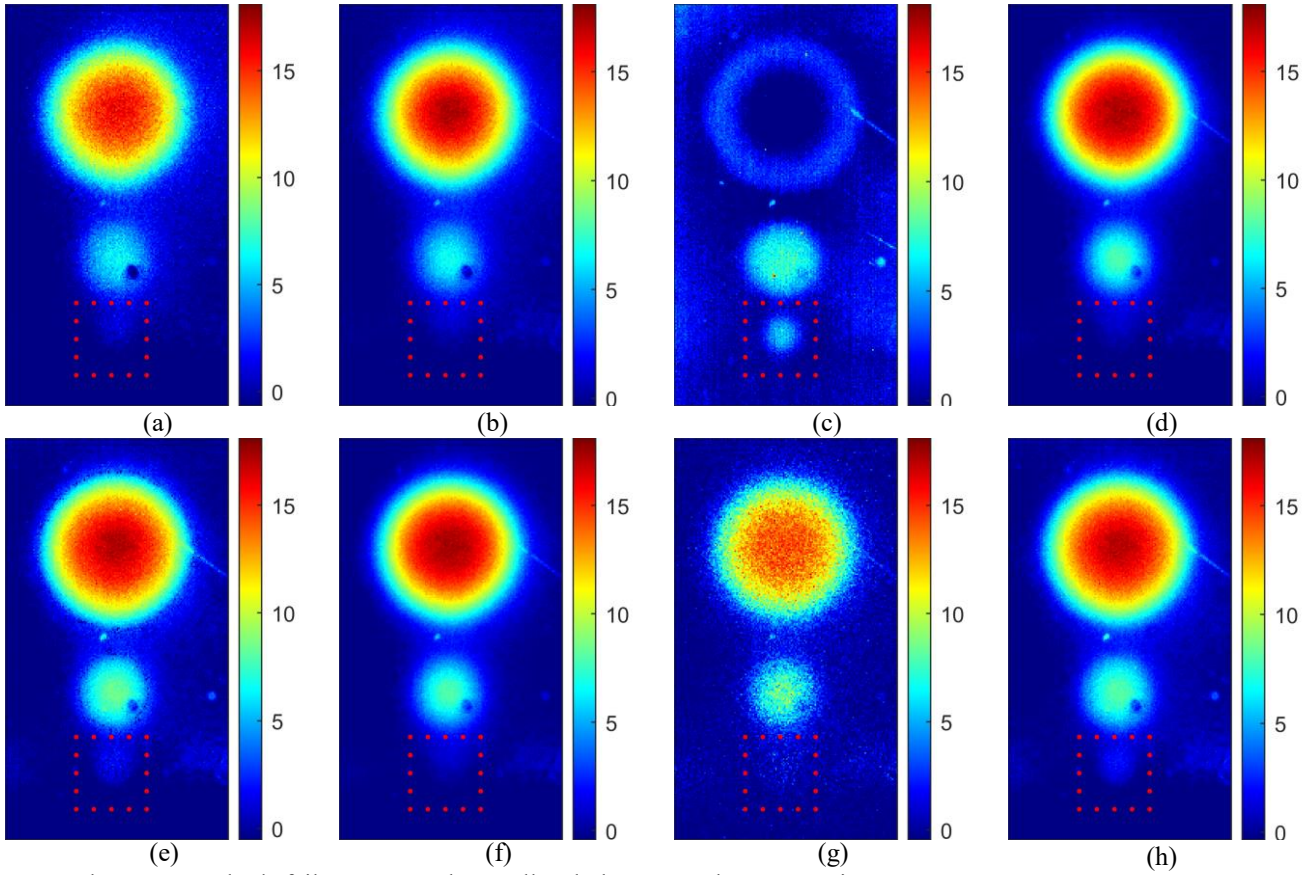


Fig. 4 that most methods fail to recover the smallest hole accurately. As seen in

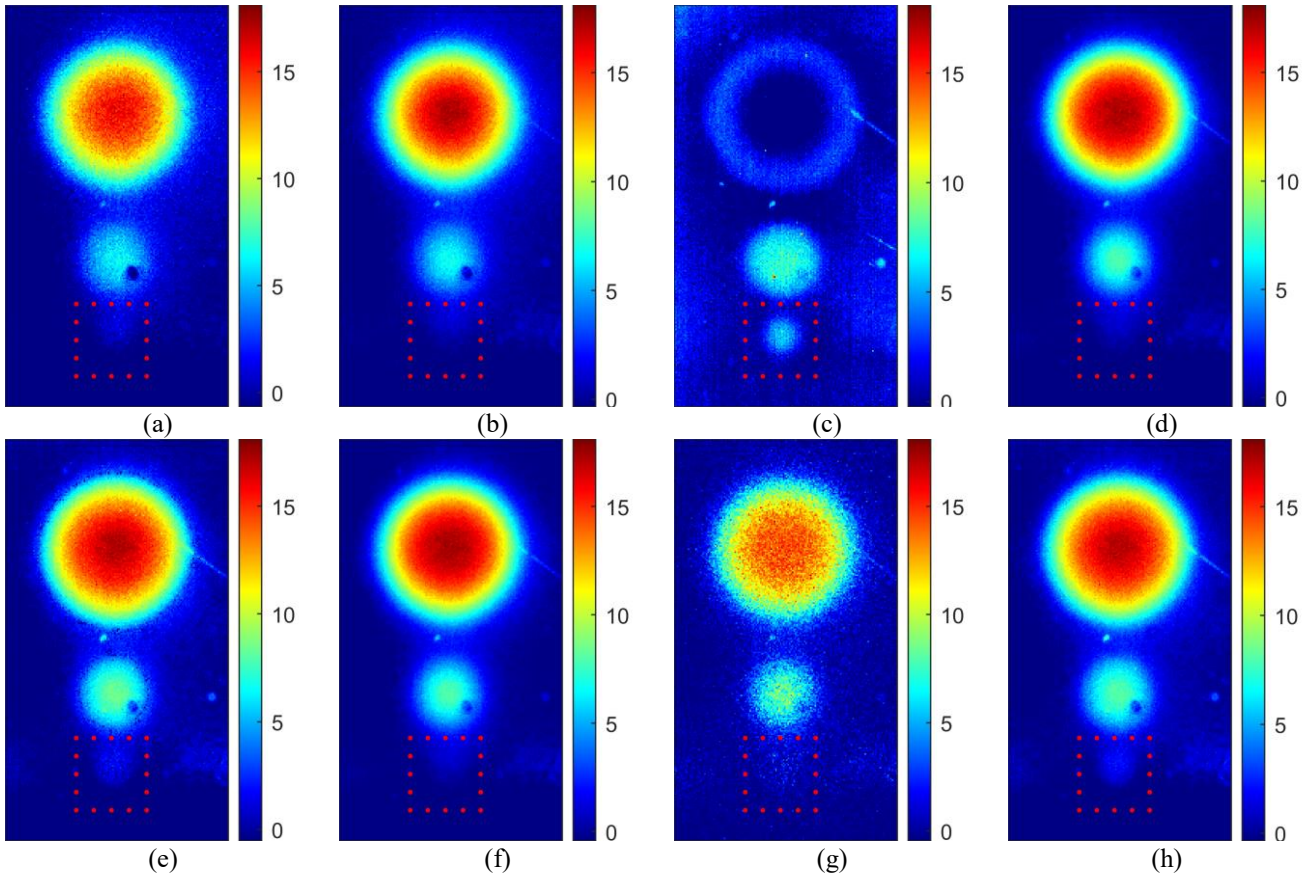


Fig. 4 (c) of the VBAS result, the estimated profiles match those of the ground truth, but the recovered internal regions around the largest hole diverge significantly from those of the ground truth.

In the result of the proposed method (

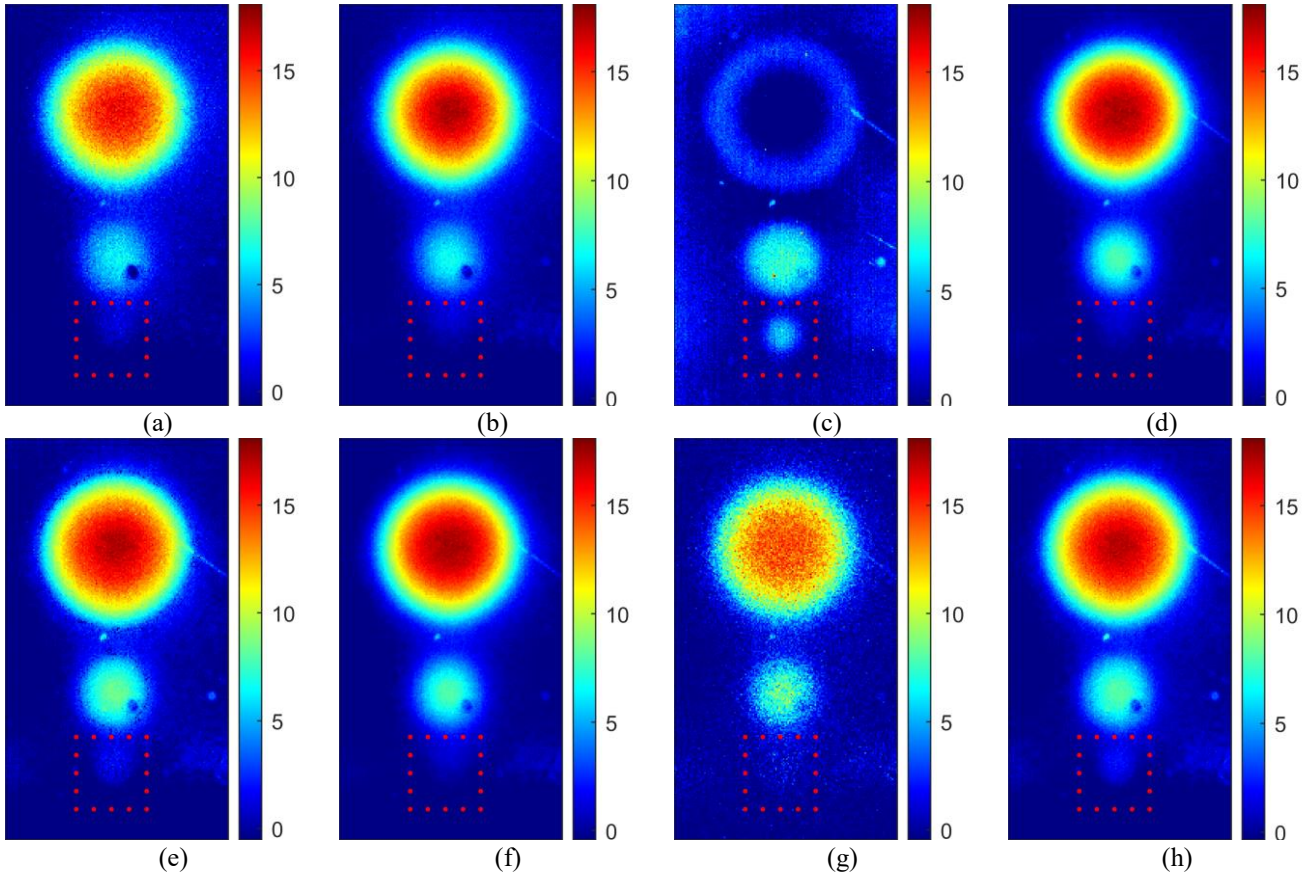


Fig. 4 (h)), all three holes are accurately and simultaneously recovered, and the recovered defective and intact regions are in good agreement with the ground truth. It is also demonstrated that the proposed method, as a variant of Bayesian TRPCA, can extract low-rank features and removing sparse outliers. The outcome of the proposed method is superior to that of other methods.

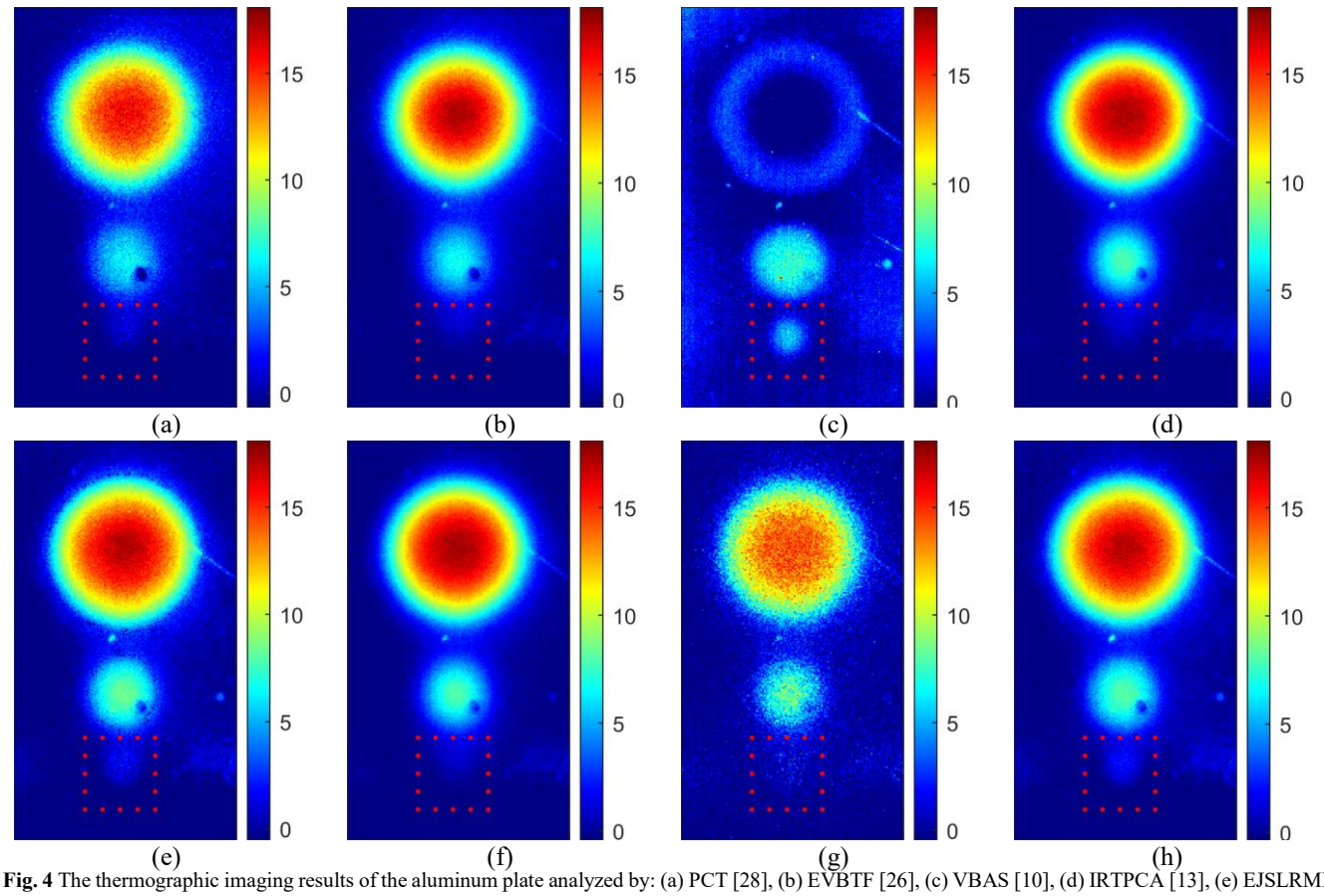


Fig. 4 The thermographic imaging results of the aluminum plate analyzed by: (a) PCT [28], (b) EVBTF [26], (c) VBAS [10], (d) IRTPCA [13], (e) EJSLRMD [5], (f) TRPCA [12], (g) OMoGMF [17], (h) the proposed method.

Meanwhile, a quantitative indicator, F-score [10], is used to evaluate the performances of different methods. The F-score is defined as:

$$\mathrm{F-score} = \frac{(\eta^2 + 1) \times \mathrm{Precision} \times \mathrm{Recall}}{\eta^2 \times (\mathrm{Precision} + \mathrm{Recall})} \tag{53}$$

where η is a parameter that determines the contribution of the Precision and the Recall. This parameter is recommended as 1 in thermographic diagnosis [29]. The terms Precision and Recall are expressed as:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$
$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{54}$$

where true positive (TP) denotes the case where a defect exists and is detected, false positive (FP) denotes the case where no defect exists but is detected, false negative (FN) denotes a defect exists but is not detected.

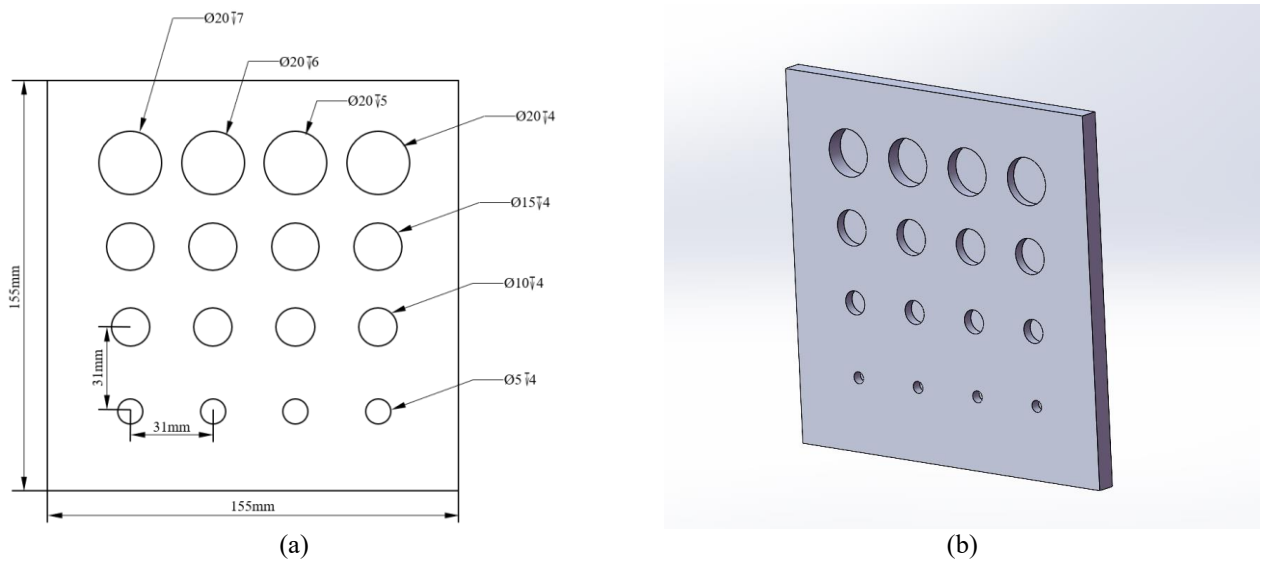
Table 1 lists F-score, SNR values, and the computational time of different methods. These methods are running on a workstation with a 2.90GHz Intel Xeon Gold 6226R CPU and 256.0 GB RAM. It can be observed that the F-score and SNR results of the proposed method are the largest ones among these methods, which demonstrates the superiority of the proposed method for accurately recovering the defect shape from the thermographic sequence. Though the PCT's computational time is shorter than that of other methods, its F-score is smaller than that of most methods. Compared with the VBAS and TRPCA methods, the proposed method reduces the computational cost. Compared with the rest methods, the proposed method improves the defect detection performance with the increase of the computational cost.

Table 1
F-score, SNR, and computational time results of different methods analyzing the first experimental signal.

	PCT [28]	EVBTF [26]	VBAS [10]	IRTPCA [13]	EJSLRMD [5]	TRPCA [12]	OMoGMF [17]	Proposed method
F-score	0.77	0.83	0.54	0.85	0.88	0.88	0.87	**0.93**
SNR	31.85	40.24	24.99	40.33	36.93	40.25	36.32	**41.31**
Computational time	**0.056**	13.92	106.13	3.35	3.22	55.51	0.91	19.03

The values in bold show the best results.

Case 2: a CFRP laminate



(a) (b)

Fig. 5 (a) The frond side and (b) the back side of the specimen in the second experiment. The circles in (a) give the dimensions and locations of holes.

The second experiment is carried out on a unidirectional Toray 800 CFRP laminate with a dimension of 155 mm × 155 mm × 8 mm. As shown in **Fig. 5**, the specimen contains 16 drilled holes arranged in a 4 by 4 array, with a distance between two adjacent hole centers of 31mm. These holes range in diameter from 5 mm to 20 mm, whereas the thicknesses vary between 4 mm and7 mm, indicating defect depths between 1 mm and 4 mm from the top inspection surface. The FLIR SC7600 IR camera is used to acquire the thermal data with a frame rate of 10 Hz. The specimen is also examined from the front surface rather than the drilled surface. Not all defects can be detected by the above methods due to the complex material properties and anisotropic nature of composite materials, the weak defect feature, and limitations of the used methods. Holes 1-4 can be easily detected with defect depths of 1 mm, but Holes 13-16 are difficult to be detected with defect depths of 4 mm. Therefore, the aim of analyzing the thermal data is to accurately identify and recover as many holes as possible.

The PCT, EVTBF, VBAS, IRTPCA, EJSLRMD, TRPCA, OMoGMF methods, and the proposed method are used to evaluate their performances on defect detection.

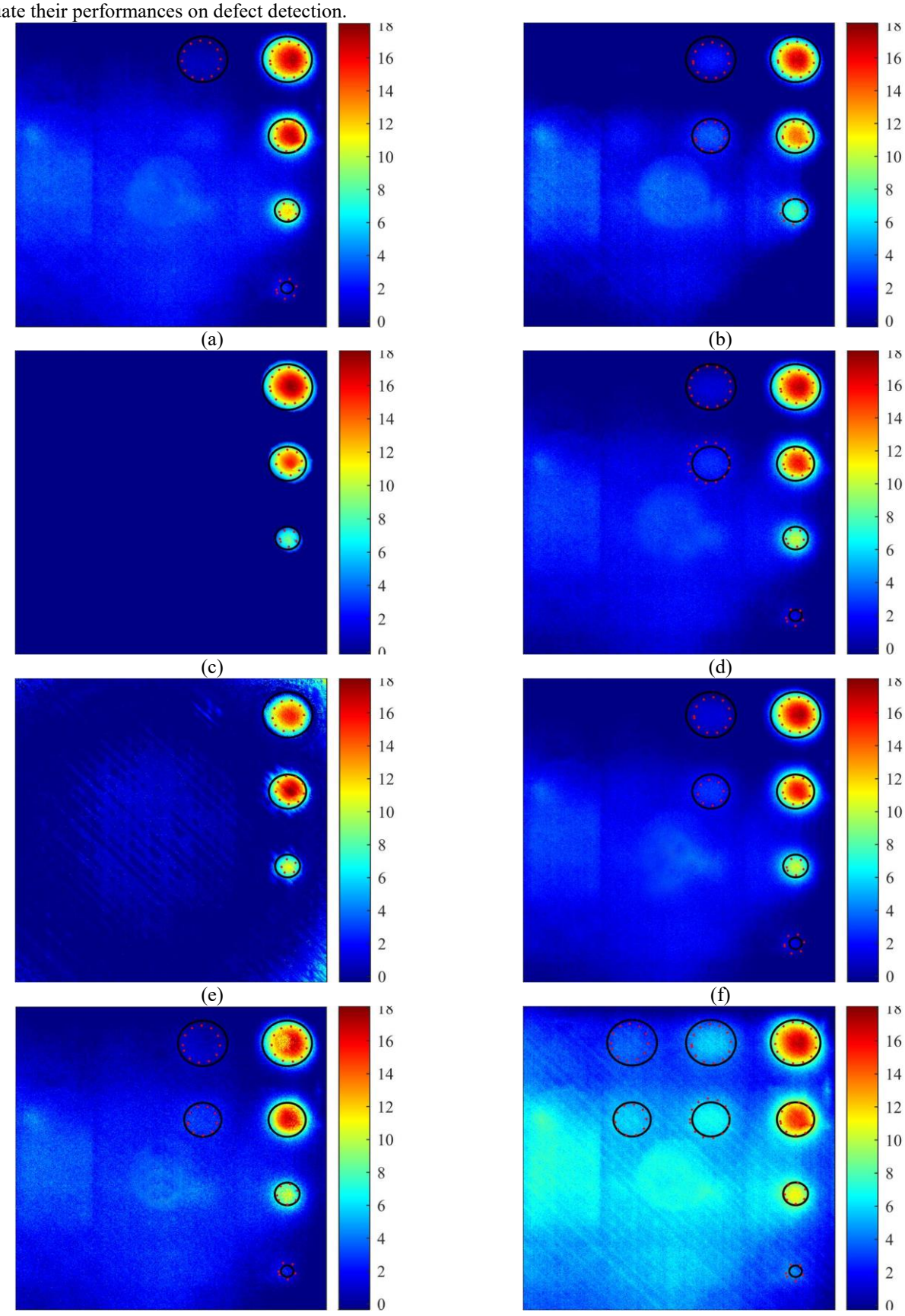
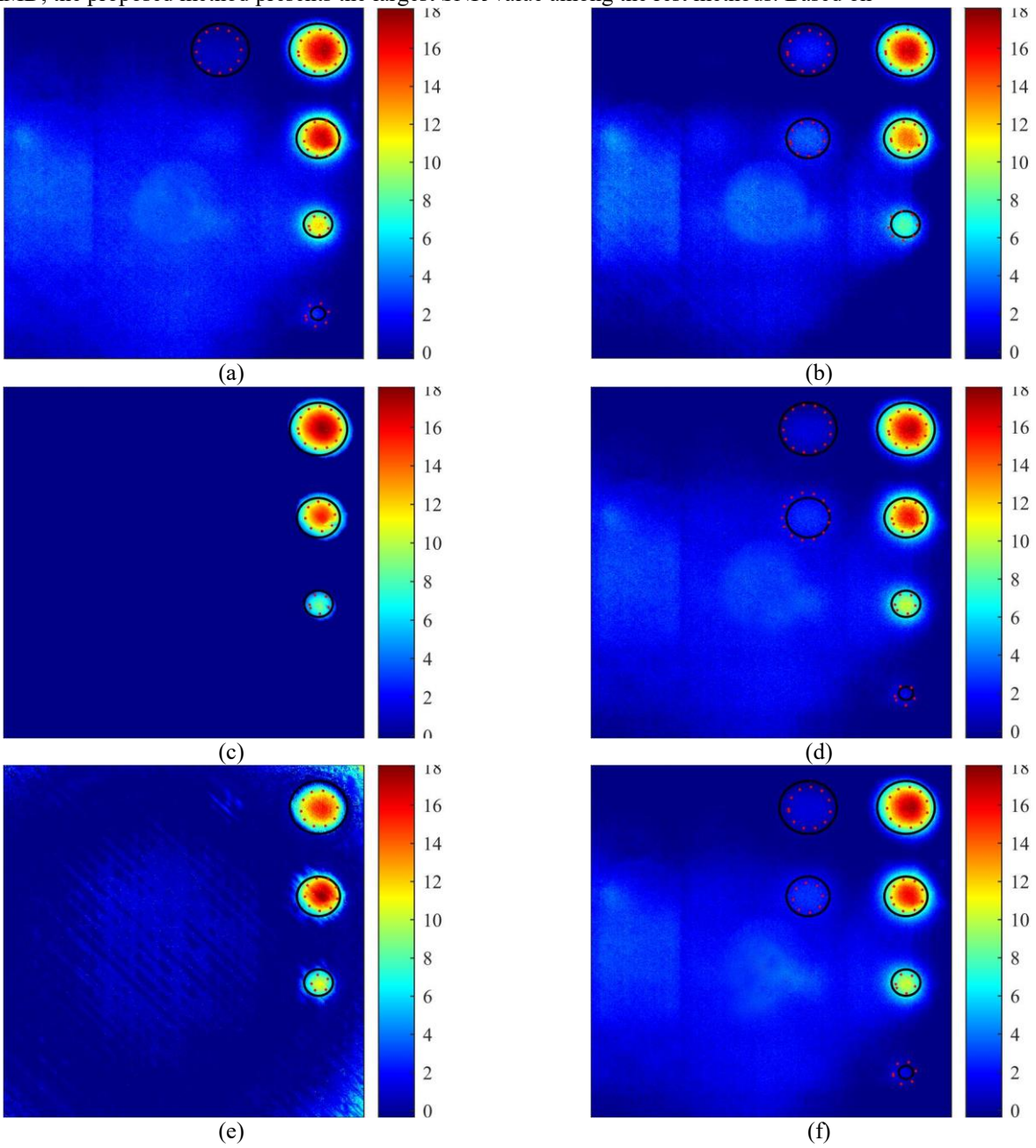


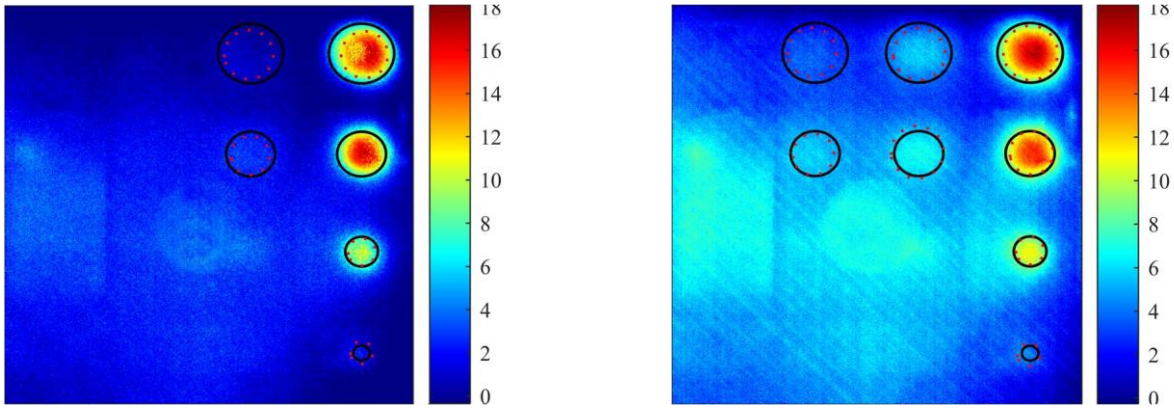
Fig. 6 shows the thermographic imaging results of the CFRP laminate. The black circles represent the actual profiles of the holes, and the red dash circles denote the profiles of the recovered holes. The VBAS and EJSLRMD methods are only capable of recovering the dominant features, Holes 1-3, while failing to identify the smallest hole in this column, Hole 4. Other than the EVTBF, VBAS, and EJSLRMD methods, other methods can recover Holes 1-4 (the rightmost column on the front side). In the IRTPCA, TRPCA, and OMoGMF methods' results, six holes can be recovered, including Holes 1, 2, 3, 4, 5, 6. The proposed method recovers eight holes of Holes 1, 2, 3, 4, 5, 6, 9, 10. The recover hole numbers of different methods are

summarized in **Table 2**. Apparently, only the proposed method could recover Holes 9-10 (the second column from the left defects in the front side). Compared to these commonly used methods, the proposed method accurately recovers the most holes. At the same time, the holes' profiles that are recovered by the proposed method have a good agreement with the actual profiles of the holes.

Furthermore, the SNR values and computational time are used to quantitatively evaluate the performances of different methods, and the results are given in **Table 2**. The F-score is not used to compare the performance of different methods since these methods cannot recover all holes. In terms of SNR, the VBAS method has the largest value. It should be noted that SNR is not a measure of the accuracy of the recovered results, but rather a measure of the thermal contrast ratio between defective and intact regions. As discussed before, the VBAS and EJSLRMD methods fail to recover Hole 4. Except the VBAS and EJSLRMD, the proposed method presents the largest SNR value among the rest methods. Based on

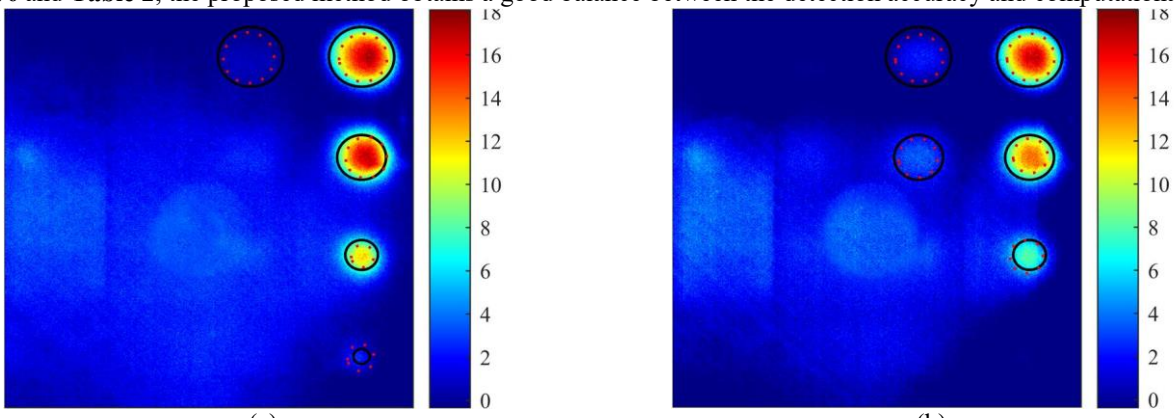


(a) (b) (c) (d) (e) (f)

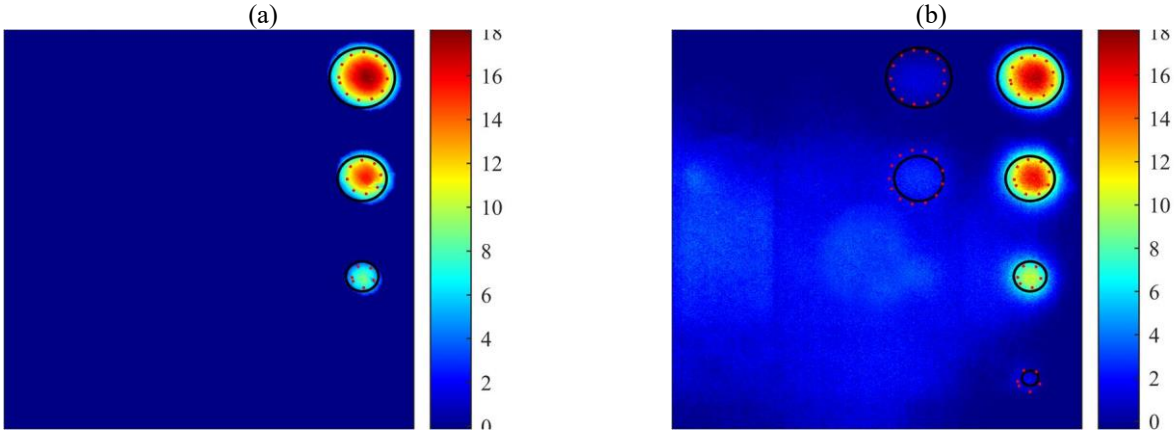


(g) (h)

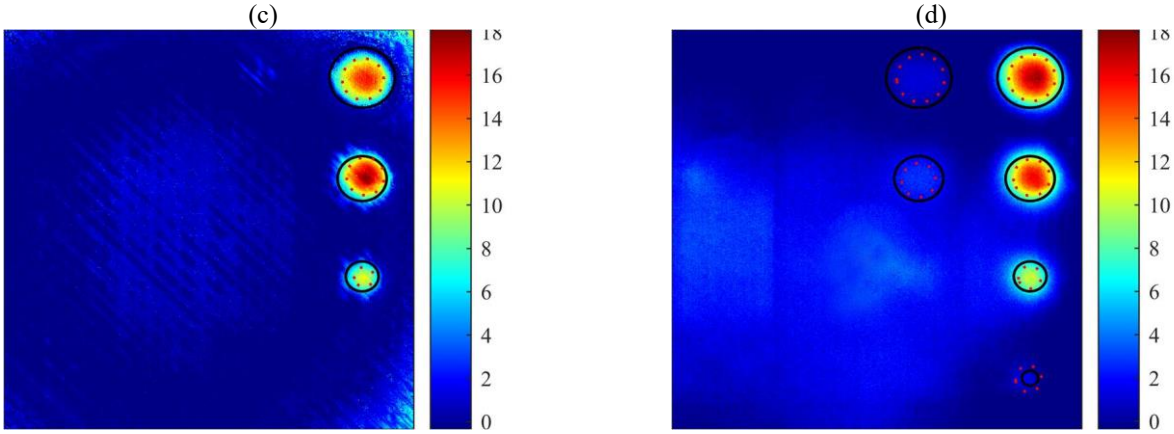
Fig. 6 and **Table 2**, the proposed method obtains a good balance between the detection accuracy and computational cost.



(a) (b)



(c) (d)



(e) (f)

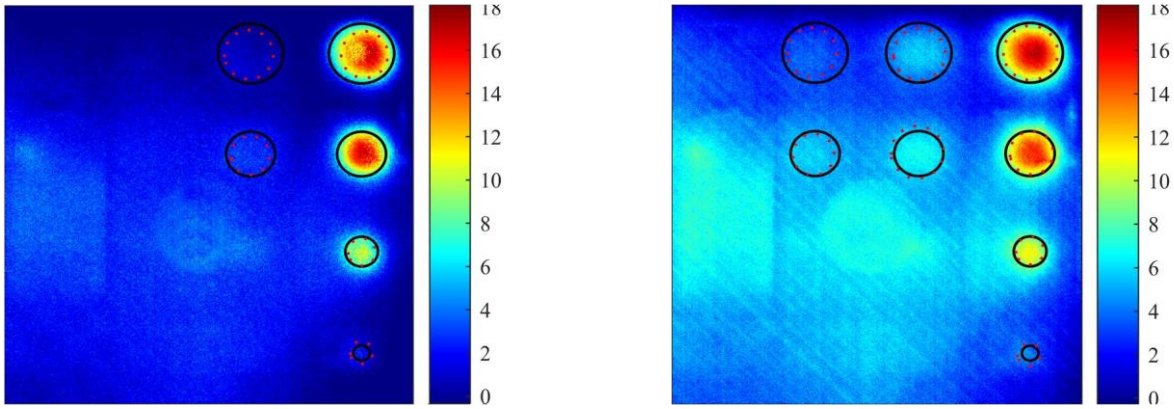


Fig. 6 The thermographic imaging results of the CFRP laminate analyzed by: (a) PCT [28], (b) EVBTF [26], (c) VBAS [10], (d) IRTPCA [13], (e) EJSLRMD [5], (f) TRPCA [12], (g) OMoGMF [17], (h) the proposed method.

Table 2
Detected holes, SNR, and computational time results of different methods analyzing the second experimental signal.

	PCT [28]	EVBTF [26]	VBAS [10]	IRTPCA [13]	EJSLRMD [5]	TRPCA [12]	OMoGMF [17]	Proposed method
Nos of detected holes	1, 2, 3, 4, 5	1, 2, 3, 5, 6	1, 2, 3	1, 2, 3, 4, 5, 6	1, 2, 3	1, 2, 3, 4, 5, 6	1, 2, 3, 4, 5, 6	1, 2, 3, 4, 5, 6, 9, 10
SNR	23.12	30.02	45.98	25.32	41.89	24.95	27.91	32.49
Computational time	2.76	179.69	271.17	58.05	52.46	848.94	28.93	114.45

5. Conclusion

This paper focuses on development of a robust and adaptive low-rank component recovery method for the thermographic inspection. The proposed method aims to extract the low-rank component from the thermographic sequence by constructing a Bayesian Tucker model. In this method, a Bayesian Tucker decomposition model is constructed to capture complex patterns and extract weak features. To exploit the low-rank property of thermal data, a generalized student-t distribution along with its hierarchical model is developed. To achieve optimal performance without the need for manual parameter tuning, the SBL is employed.

Two experimental signals obtained from a thermography system are used to evaluate the proposed method's performance. In both experiments, the surface being inspected is the opposite of the surface being drilled, which leads to weak defects and makes defect detection more difficult. As indicated by the thermographic imaging results as well as the quantitative results based on the F-score, the SNR, the computational time for the thermographic inspection system, the proposed method has proven to be the most effective method for improving the spatial resolution and accuracy for such a system. The proposed method improves the defect detection performance with the increase of the computational cost. In the future, we want to improve its computational efficiency to meet the demand online analysis applications. We also want to use the proposed method on defect detection for specimens with the irregular shape.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants Nos. 52105113 and 52175104) and the Fundamental Research Funds for the Central Universities. The authors thank Prof. Bin Gao with the University of Electronic Science and Technology, China, for his valuable help in the EJSLRMD algorithm.

References

[1] P. Zhu, C. Yin, Y. Cheng, X. Huang, J. Cao, C.M. Vong, P.K. Wong, An improved feature extraction algorithm for automatic defect identification based on eddy current pulsed thermography, Mech. Syst. Signal Process. 113 (2018) 5–21. https://doi.org/10.1016/j.ymssp.2017.02.045.

[2] F. Wang, J. Liu, P. Song, J. Gong, W. Peng, G. Liu, M. Chen, Y. Wang, Multimodal optical excitation pulsed thermography: Enhanced recognize debonding defects of the solid propellant rocket motor cladding layer, Mech. Syst. Signal Process. 163 (2021). https://doi.org/10.1016/j.ymssp.2021.108164.

[3] Z. Tong, S. Xie, H. en Chen, J. Qiu, W. Cai, C. Pei, Z. Chen, T. Uchimoto, T. Takagi, Quantitative mapping of depth profile of fatigue cracks using eddy current pulsed thermography assisted by PCA and 2D wavelet transformation, Mech. Syst. Signal Process. 175 (2022). https://doi.org/10.1016/j.ymssp.2022.109139.

[4] B. Azizinasab, R.P.R. Hasanzadeh, S. Hedayatrasa, M. Kersemans, Defect Detection and Depth Estimation in CFRP through Phase of Transient Response of Flash Thermography, IEEE Trans. Ind. Informatics. 3203 (2021) 2364–2373. https://doi.org/10.1109/tii.2021.3101492.

[5] J. Ahmed, B. Gao, W.L. Woo, Y. Zhu, Ensemble Joint Sparse Low-Rank Matrix Decomposition for Thermography Diagnosis System, IEEE Trans. Ind. Electron. 68 (2021) 2648–2658. https://doi.org/10.1109/TIE.2020.2975484.

[6] B. Yousefi, S. Sfarra, F. Sarasini, X.P.V. Maldague, IRNDT Inspection Via Sparse Principal Component Thermography, Can. Conf. Electr. Comput. Eng. 2018-May (2018) 1–4. https://doi.org/10.1109/CCECE.2018.8447887.

[7] J.Y. Wu, S. Sfarra, Y. Yao, Sparse principal component thermography for subsurface defect detection in composite products, IEEE Trans. Ind. Informatics. 14 (2018) 5594–5600. https://doi.org/10.1109/TII.2018.2817520.

[8] J. Ahmed, B. Gao, G.Y. Tian, Y. Yang, Y.C. Fan, Sparse ensemble matrix factorization for debond detection in CFRP composites using optical thermography, Infrared Phys. Technol. 92 (2018) 392–401. https://doi.org/10.1016/j.infrared.2018.06.018.

[9] H. Momeni, A. Ebrahimkhanlou, High-dimensional data analytics in structural health monitoring and non-destructive evaluation: A review paper, Smart Mater. Struct. 31 (2022). https://doi.org/10.1088/1361-665X/ac50f4.
[10] B. Gao, P. Lu, W.L. Woo, G.Y. Tian, Y. Zhu, M. Johnston, Variational Bayesian Subgroup Adaptive Sparse Component Extraction for Diagnostic Imaging System, IEEE Trans. Ind. Electron. 65 (2018) 8142–8152. https://doi.org/10.1109/TIE.2018.2801809.
[11] S. Ebrahimi, J. Fleuret, M. Klein, L.D. Théroux, M. Georges, C. Ibarra-Castanedo, X. Maldague, Robust principal component thermography for defect detection in composites, Sensors. 21 (2021) 1–19. https://doi.org/10.3390/s21082682.
[12] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, S. Yan, Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm, IEEE Trans. Pattern Anal. Mach. Intell. 42 (2020) 925–938. https://doi.org/10.1109/TPAMI.2019.2891760.
[13] Y. Liu, L. Chen, C. Zhu, Improved robust tensor principal component analysis via low-rank core matrix, IEEE J. Sel. Top. Signal Process. 12 (2018) 1378–1389.
[14] M.E. Kilmer, C.D. Martin, Factorization strategies for third-order tensors, Linear Algebra Appl. 435 (2011) 641–658. https://doi.org/10.1016/j.laa.2010.09.020.
[15] B. Gao, Y. He, W.L. Woo, G.Y. Tian, J. Liu, Y. Hu, Multidimensional Tensor-Based Inductive Thermography with Multiple Physical Fields for Offshore Wind Turbine Gear Inspection, IEEE Trans. Ind. Electron. 63 (2016) 6305–6315. https://doi.org/10.1109/TIE.2016.2574987.
[16] J. Ahmed, B. Gao, W.L. Woo, Sparse Low-Rank Tensor Decomposition for Metal Defect Detection Using Thermographic Imaging Diagnostics, IEEE Trans. Ind. Informatics. 17 (2021) 1810–1820. https://doi.org/10.1109/TII.2020.2994227.
[17] H. Yong, D. Meng, W. Zuo, L. Zhang, Robust Online Matrix Factorization for Dynamic Background Subtraction, IEEE Trans. Pattern Anal. Mach. Intell. 40 (2018) 1726–1740. https://doi.org/10.1109/TPAMI.2017.2732350.
[18] T. Wu, B. Gao, W.L. Woo, Hierarchical low-rank and sparse tensor micro defects decomposition by electromagnetic thermography imaging system, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 378 (2020). https://doi.org/10.1098/rsta.2019.0584.
[19] L. Bai, Y. Liang, J. Shao, Y. Cheng, Q. Zhou, J. Zhang, Moran's Index-Based Tensor Decomposition for Eddy Current Pulsed Thermography Sequence Processing, IEEE Trans. Instrum. Meas. 70 (2021) 1-12. https://doi.org/10.1109/TIM.2021.3096277.
[20] Z. Luo, H. Luo, S. Wang, F. Chen, Z. Su, P. Shen, H. Zhang, Enhanced CFRP Defect Detection from Highly Undersampled Thermographic Data via Low-Rank Tensor Completion-Based Thermography, IEEE Trans. Ind. Informatics. 18 (2022) 8641–8653. https://doi.org/10.1109/TII.2022.3154786.
[21] X. Ding, L. He, L. Carin, Bayesian robust principal component analysis, IEEE Trans. Image Process. 20 (2011) 3419–3430. https://doi.org/10.1109/TIP.2011.2156801.
[22] J. Liu, B.D. Rao, Sparse Bayesian learning for robust PCA: Algorithms and analyses, IEEE Trans. Signal Process. 67 (2019) 5837–5849. https://doi.org/10.1109/TSP.2019.2946026.
[23] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika. 31 (1966) 279–311.
[24] R. Bro, PARAFAC. Tutorial and applications, Chemom. Intell. Lab. Syst. 38 (1997) 149–171.
[25] Q. Zhao, G. Zhou, L. Zhang, A. Cichocki, S.I. Amari, Bayesian Robust Tensor Factorization for Incomplete Multiway Data, IEEE Trans. Neural Networks Learn. Syst. 27 (2016) 736–748. https://doi.org/10.1109/TNNLS.2015.2423694.
[26] P. Lu, B. Gao, Q. Feng, Y. Yang, W.L. Woo, G.Y. Tian, Ensemble variational Bayes tensor factorization for super resolution of CFRP debond detection, Infrared Phys. Technol. 85 (2017) 335–346. https://doi.org/10.1016/j.infrared.2017.07.012.
[27] A. Mohammad-Djafari, M. Dumitru, Bayesian sparse solutions to linear inverse problems with non-stationary noise with Student-t priors, Digit. Signal Process. A Rev. J. 47 (2015) 128–156. https://doi.org/10.1016/j.dsp.2015.08.005.
[28] N. Rajic, Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures, Compos. Struct. 58 (2002) 521–528.
[29] Y. Wang, B. Gao, W.L. Woo, G. Tian, X. Maldague, L. Zheng, Z. Guo, Y. Zhu, Thermal pattern contrast diagnostic of microcracks with induction thermography for aircraft braking components, IEEE Trans. Ind. Informatics. 14 (2018) 5563–5574. https://doi.org/10.1109/TII.2018.2802046.

Appendix

1. The variational posterior distribution of $\mathcal{G}$

The logarithmic approximate posterior of $\mathcal{G}$ is given as:

$$\begin{aligned}&\ln q(\mathcal{G}) = \ln q\left(\mathrm{vec}(\mathcal{G})\right) = \mathrm{E}_{q(\boldsymbol{\theta})}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right]\\&= -\frac{1}{2}\mathrm{vec}^T(\mathcal{G})\left(\mathrm{E}(\rho)\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)T}\mathbf{U}^{(n)}\right)\right) + \mathrm{E}(\eta)\bigotimes_n \mathrm{E}\left(\boldsymbol{\Lambda}^{(n)}\right)\right)\mathrm{vec}(\mathcal{G})\\&+\mathrm{E}(\rho)\mathrm{vec}^T(\mathcal{G})\left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)T}\right)\right)\left(\mathrm{vec}(\mathcal{X}) - \boldsymbol{\mu}_{\mathrm{vec}(\mathcal{S})}\right) + \mathrm{const}\end{aligned} \tag{A-1}$$

where $\boldsymbol{\Lambda}^{(n)} = \mathrm{diag}\left(\boldsymbol{\lambda}^{(n)}\right)$. $\mathrm{E}(\boldsymbol{\theta}_k)$ is the expected value of $\boldsymbol{\theta}_k$. Then it can be seen that $q(\mathcal{G})$ follows a normal distribution.

2. The variational posterior distribution of $\boldsymbol{\lambda}^{(n)}$

For the vector $\boldsymbol{\lambda}^{(n)}$, its element $\lambda_{r_n}^{(n)}$ is independent. Then the logarithmic approximate posterior of $\lambda_{r_n}^{(n)}$ is given as:

$$\begin{aligned}&\ln q\left(\lambda_{r_n}^{(n)}\right) = \mathrm{E}_{q\left(\boldsymbol{\theta}\setminus\lambda_{r_n}^{(n)}\right)}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right] = \ln\left(\lambda_{r_n}^{(n)}\right)\left(\frac{I_n + \prod_{k\neq n} R_k}{2}\right)\\&-\lambda_{r_n}^{(n)}\frac{\mathrm{E}\left(\mathbf{u}_{r_n}^{(n)T}\mathbf{u}_{r_n}^{(n)}\right)}{2} - \lambda_{r_n}^{(n)}\left(\frac{\mathrm{E}(\eta)}{2}\sum_{r_l,r_m}\mathrm{E}\left(\lambda_{r_l}^{(l)}\right)\mathrm{E}\left(\lambda_{r_m}^{(m)}\right)\mathrm{E}\left(\mathcal{G}_{r_l r_m r_n}^2\right) + \mathrm{E}\left(\gamma_{r_n}^{(n)}\right)\right) + \mathrm{const}\end{aligned} \tag{A-2}$$

Then $q\left(\lambda_{r_n}^{(n)}\right)$ follows a Gamma distribution.

3. The variational posterior distribution of η

The logarithmic approximate posterior of η is expressed as:

$$\begin{aligned}&\ln q(\eta) = \mathrm{E}_{q(\boldsymbol{\theta}\setminus\eta)}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right]\\&= \left(\frac{R_1R_2R_3}{2} + \left(\alpha_\eta - 1\right)\right)\ln\eta - \left(\frac{1}{2}\mathrm{E}\left(\mathrm{vec}^T(\mathcal{G})\left(\bigotimes_n \boldsymbol{\Lambda}^{(n)}\right)\mathrm{vec}(\mathcal{G})\right) + \beta_\eta\right)\eta + \mathrm{const}\end{aligned} \tag{A-3}$$

Hence, $q(\eta)$ follows a Gamma distribution.

4. The variational posterior distribution of $\mathcal{S}$

The logarithmic approximate posterior of $\mathcal{S}$ is:

$$\begin{aligned}&\ln q(\mathcal{S}) = \ln q\left(\mathrm{vec}(\mathcal{S})\right) = \mathrm{E}_{q(\boldsymbol{\theta}\setminus\mathcal{S})}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right] = -\frac{1}{2}\mathrm{vec}^T(\mathcal{S})\left(\mathrm{E}(\rho) + \mathrm{E}(\Xi)\right)\mathrm{vec}(\mathcal{S})\\&+\mathrm{E}(\rho)\left(\mathrm{vec}(\mathcal{X}) - \left(\bigotimes_n \mathrm{E}\left(\mathbf{U}^{(n)}\right)\right)\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{G})}\right)^T\mathrm{vec}(\mathcal{S}) + \mathrm{const}\end{aligned} \tag{A-4}$$

where $\Xi = \mathrm{diag}(\boldsymbol{\xi})$. $q(\mathcal{S})$ follows a normal distribution.

5. The variational posterior distribution of $\boldsymbol{\xi}$

ξ_m is mth element of the vector $\boldsymbol{\xi}$ and its logarithmic approximate posterior is given as:

$$\begin{aligned}&\ln q(\xi_m) = \mathrm{E}_{q(\boldsymbol{\theta}\setminus\xi_m)}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right]\\&= \frac{1}{2}\ln(\xi_m) - \left(\frac{\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{S})_m}^T\boldsymbol{\mu}_{\mathrm{vec}(\mathcal{S})_m} + \boldsymbol{\Sigma}_{\mathrm{vec}(\mathcal{S})_{m,m}}}{2} + \mathrm{E}(\tau_m)\right)\xi_m + \mathrm{const}\end{aligned} \tag{A-5}$$

Then $q(\xi_m)$ follows a Gamma distribution.

6. The variational posterior distribution of ρ

The logarithmic approximate posterior of ρ is:

$$\begin{aligned}&\ln q(\rho) = \mathrm{E}_{q(\boldsymbol{\theta}\setminus\rho)}\left[\ln\left(p(\mathbf{x},\boldsymbol{\theta})\right)\right] = \left(\frac{I_1I_2I_3}{2} + \alpha_\rho - 1\right)\ln\rho - \rho\beta_\rho\\&-\frac{\rho}{2}\mathrm{E}\left(\left\|\mathrm{vec}(\mathcal{X}) - \mathrm{vec}(\mathcal{S}) - \left(\bigotimes_n \mathbf{U}^{(n)}\right)\mathrm{vec}(\mathcal{G})\right\|_F^2\right) + \mathrm{const}\end{aligned} \tag{A-6}$$

Therefore, $q(\rho)$ follows a Gamma distribution.

https://dspace.org/

Tensor robust principal component analysis based on Bayesian Tucker decomposition for thermographic inspection

Hu, Yue; Cui, Fangsen; Zhao, Yifan; Li, Fucai; Cao, Shuai ...

2023-12-01
Elsevier
https://doi.org/10.1016/j.ymssp.2023.110761
Hu Y, Cui F, Zhao Y, et al., (2023) Tensor robust principal component analysis based on Bayesian Tucker decomposition for thermographic inspection. Mechanical Systems and Signal Processing, Volume 204, December 2023, Article Number 110761

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