



Regular article

Sparse ensemble matrix factorization for debond detection in CFRP composites using optical thermography

Junaid Ahmed^a, Bin Gao^{a,*}, Gui Yun Tian^{a,b}, Yang Yang^c, Yun Chen Fan^c^a School of Automation, University of Electronic Science and Technology of China, China^b School of Electrical and Electronic Engineering, Newcastle University, England, UK^c Chengdu Aircraft Industry Co., Ltd, China

ARTICLE INFO

Keywords:

Optical pulsed thermography

Debond defects

CFRP

Sparse MoG low-rank matrix factorization

Non-destructive testing

ABSTRACT

With the increasing use of CFRP in aircraft and automotive industry, it has become necessary to monitor the health of the composite structure during the manufacture and service process from flaws such as debond or delamination by using optical pulsed thermography non-destructive testing (OPTNDT) techniques. However, current OPTNDT methods cannot efficiently tackle the influence from the strong noise, and the resolution enhancement of the defects detection remains as a critical challenge. To alleviate this problem, this paper proposes the sparse ensemble matrix factorization approach to remove the noise and enhance the resolution for the defects detection. Specifically, the algorithm is based on the sparse representation and noise is modeled as a mixture of Gaussian (MoG) distribution. It provides a projection from the raw data onto the sparse and low dimensional sub-space while the defects information is significantly enhanced with the layer decomposed approach in the low dimensional space. Not withstanding above, a Gaussian low pass filtering and non-linear enhancement is conducted for further enhancement. The proposed method is coupled with several comparison studies to verify the efficacy.

1. Introduction

The use of CFRP as the composite material is increasing sharply in the aerospace industry [1]. This is due to its unique advantage over other materials such as high strength, high fatigue resistance, low thermal expansion and good resistance to corrosive damages [2]. Thus, it becomes necessary for the industry manufacturers to inspect and check the health and quality to avoid any unwanted disaster. General types of defects in the CFRP are delamination and debond. These type of defects usually occur during the manufacturing and acquisition process. Apart from the manual inspection nowadays non-destructive testing (NDT) techniques have been used for its inspection.

In [3,4], the researchers have emphasized the importance of NDT technology to evaluate and quantize the quality and structural health of the composite material after production, repair and in monitoring phase. In the literature, different NDT methods have been applied to evaluate the defects in the composite materials. These methods include eddy current method, ultrasonic method, magnetic resonance imaging method and so on [5]. With the recent advances in optical technology, infrared thermography has been widely used for the inspection of the specimen. It has the advantage of non-contact, wide area and fast inspection. In [6], Cramer et al. proposed flash infrared thermography to

analyze and evaluate the structural health of the composites at NASA. In [7], R. Yang et al. gave a concise survey of the existing NDT techniques and methods for composite structures. In [8], Ryu et al. have demonstrated various external heat sources that can be used in active IRT. The optical source for IRT has been successfully used for the inspection of the specimen at hand. In [9], Chulkov et al. have demonstrated the use of various optical sources such as xenon and halogen lamps. In general, the process of infrared thermography involves the application of external heating on the specimen under test. If there are defects available on the specimen, non-uniform thermal distribution will occur. These distributions can then be recorded by infrared camera [10,11]. In [12], Maierhofer et al. have put some light on the use of flash thermography for CFRP materials i.e. limitations and accuracy. Currently, available OPT based feature extraction and defect detection methods are subject to poor resolution of the thermal images. In addition, the background and thermal noise are two components that are affecting the performance of the algorithms. Many authors in the literature have tried different methods and techniques to resolve this problem. But still, current methods and techniques are far from being implemented in the real scenario. In [13], Zhang et al. have used the optical pulse thermography for analyzing the defect in the dry carbon pre-forms. The comparative analysis was carried out by applying

* Corresponding author.

E-mail address: bin_gao@uestc.edu.cn (B. Gao).

different image processing techniques and use the optical thermography for the CFRP composites along with its hybrid composite called basalt-carbon fiber polymer for impact loading [14]. The evaluations were carried out by using different pulse thermography techniques.

In [15] Zhang et al., propose a combined spectral-spatial domain feature extraction algorithm. The proposed algorithm is evaluated on a set of publically available hyperspectral images. In [16] Zhang et al., proposed an ensemble manifold regularized sparse low rank approximation for multi-view feature embedding. To directly remove the noise component a group sparsity approach is introduced. In [17], Yuanlin et al. proposed a novel fitting the time derivative coefficient algorithms (FCA) for defect detection. Further, they employed ant colony algorithm (ACA) to detect the defect edges. In [18], Yousefi et al. proposed a new Eigen decomposition approach for principal component thermography (PCT) called candid-covariance free incremental principal component thermography (CCIPCT). They developed a new approach to reduce the computational cost. In [19], Maldague et al. have employed the discrete Fourier transform (DFT) for the estimation and defect depths by using the frequency analysis. In [20], Marinetti et al. have used the pulse phase thermography (PPT) for defect detection. They used the amplitude and frequency information from the Fourier transform and mapped the temperature changes onto the frequency domain for defect detection. The principal component analysis (PCA) [21,22] is used for defect detection. It is based on singular value decomposition (SVD) for feature extraction. In [23], Liang et al. have employed wavelet processing along with the PCA to detect defects that are shallow and deep. In [24,25], thermal signal reconstruction (TSR) is proposed. It is based on polynomial representation and fitting the data. In [26], Sripragash et al. employed TSR for defect depth estimation and characterization. In [27], Lopez et al. proposed the use of partial least square thermography for defect detection. In [28], Peng et al. proposed a seeded region growing method with TSR for debond detection in CFRP. In [29], Zauner et al. used wavelet processing for NDT of wood plastic. The divergence feature was used by Ren et al. [30] to solve the NDT problem in CFRP. In [31], Junyan et al. used the simulation annealing method for defect detection in the CFRP specimen. In [32], independent component analysis (ICA) algorithm has been used for detecting defects in specimen under test. In [33], Zhao et al. proposed a variational bayes tensor factorization (VBTF) to extract useful information from the image sequence. In [34], Lu et al. provided extended variational bayes approach by layering the algorithms in [33] for the problem of debond detection in CFRP. Their algorithm provides reasonable results under comparison using the F-score parameter. In [35], Yong et al. provided a low-rank matrix factorization approach by modeling the noise using mixture of Gaussian distribution and considering a separate distribution for each frame under consideration. The model developed is robust and provides good results.

However, the algorithms proposed so far are still limited in detecting defects at deeper depths as well as a challenge issue of the poor resolution of the defects. In addition, the noise component is still inherent and needs to be tackled. Due to these reasons, current techniques are still not completely applicable. To contribute in this area, we propose the sparse representation based approach. Orthogonal Matching Pursuit (OMP) [36] is used as the sparse representation technique. As OMP projects the data onto a low dimensional space, the noise component present in the higher dimension is inherently removed. Furthermore, this sparse data is processed using MoG low-rank matrix factorization algorithm. As the defect information is mostly present in the low-rank matrix [34], we can extract this low-frequency information from the thermal sequence. The algorithm of [34] assumes the noise has Gaussian distribution, however, this assumption is not always true and in some cases deviates from real cases where noise has more complex distributions. To tackle this problem, the MoG distribution is proposed [35]. For general noise distribution, the MoG has good approximation properties [37]. Along with the sparse representation and MoG noise distribution, we propose a multi-layered structure of the

sparse matrix decomposition algorithm. The layer decomposition is evaluated by running several tests on different datasets. Particular, we propose an alternate mechanism which avoids the layering structure by using the Gaussian low pass filtering as well as the non-linear enhancement. As the layering structure is avoided, the overall computational cost of the algorithm is reduced in comparison to [34] which use a 4 layer structure. In [34], the authors process the raw data for their experiment and in our experiment we process the sparse data. This step accounts for overall computational cost reduction as the data processed by the MoG algorithm. The inclusion of the sparse representation and Gaussian filtering completely removes the noise and quantifies the defects information.

The proposed algorithm emphasis the sparse and deep layer decomposition for matrix factorization which helps remove the noise but in turn induces slight increase the additional computational cost. The proposed algorithm is coupled with Gaussian filtering and non-linear enhancement approach to further enhance the defect resolution which can be accounted as the advantage for defect detection. The main weak point with respect to the algorithm of [35] can be the additional computational cost when a three layered structure is constructed.

The proposed algorithm is tested for subsurface debond detection in the CFRP material. Initially, the CFRP sheets are processed by the OPT system to get the time series thermographic images. Later defect information is extracted by applying the proposed algorithm. The proposed algorithm is compared with general OPTNDT algorithms and the results are evaluated on the basis of event F-score [34]. Results indicate our algorithm performs better in comparison with other algorithms.

The rest of the paper is organized as; Section 2 gives the details of the proposed algorithms. Section 3 gives the details about the experimental setup and specimen under test. Section 4 gives the experimental results and discussion. Section 5 gives the conclusion of the paper.

2. The proposed methodology

2.1. Optical pulse thermography

Generally, this technology is applied in NDT because of its usefulness and owing to these properties such as fast inspection of a large area of specimen under test and it works on the heating and cooling principle of the pulse heating. This type of thermographic technique falls under the surface heating thermography (SHT) class [38]. The excitation sources are the halogen lamps and the reflection mode configuration is used where the infrared camera and the excitation source are kept facing the same direction. Suitable duration of the pulse excitation is chosen based on the specimen thickness and thermal behavior (conductivity and diffusivity of the specimen). In the heating and cooling phase the temperature of the surface of the specimen under test changes. However, if there are defects present in the specimen such as debonds, they change the thermal diffusion rate of the specimen. If any defect found inside the structure it can be identified by the difference of this temperature from the normal case (see Fig. 1).

The thermal sequence obtained by the infrared camera is usually processed by some image processing algorithms. Most of these signal processing algorithms work on the pixels of the images. Generally, these algorithms process the time series of each pixel separately. Here, series represents the general temperature variation of surface with respect to time. The thermal diffusion of the surface which is the change in temperature for the lateral direction can be discarded as compared to the normal direction. The single dimensional heat transfer equation can be expressed as;

$$\frac{\partial^2 Temp}{\partial y^2} = \frac{1}{\beta} \frac{dTemp}{dt} \quad (1)$$

where *Temp* represents the temperature of the surface under test and β represents the thermal diffusivity, here *t* and *y* represent the time and

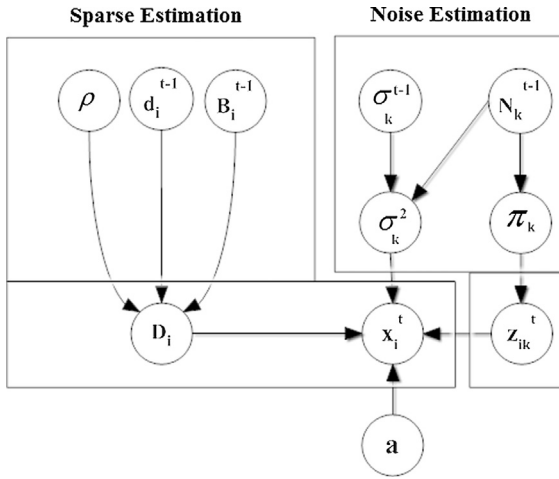


Fig. 1. Probabilistic model of the MoG-LRMF.

depth of the specimen under test. For an ideal impulsive heat flux, the response for a semi-infinite surface is given by;

$$T(y, t) = \frac{E}{e\sqrt{\pi t}} e^{-\frac{y^2}{4\pi t}} \quad (2)$$

where e is thermal effusivity of the material determined by the thermal conductivity, mass density, and specific heat E is the quantity of the energy absorbed by the surface. Since thermal imaging is only applicable to the surface temperature we take $y = 0$ in the above equation and the equation for surface temperature becomes;

$$\Delta T(t) = \frac{E}{e\sqrt{\pi t}} = \frac{E}{e\sqrt{\pi t}} t^{-0.5} \quad (3)$$

Here ΔT represents the change in temperature after the pulse heating is applied. Taking the logarithmic transform the time-temperature curve produces a decreasing linear slope of 0.5 and this can be used for defect detection and characterization.

$$\ln(T) = \ln\left(\frac{E}{e\sqrt{\pi t}}\right) - 0.5\ln(t) \quad (4)$$

The heating style is the surface heating thermography with the reflection mode configuration. The pulse thermography analysis in the time domain is made [38].

2.2. Sparse MoG low-rank matrix factorization

The given thermographic sequences are converted into the matrix form representing their time series. Let $X = [x_1, x_2, \dots, x_n] \in R^{m \times n}$ be the data matrix containing noise and defect information, where m, n are the dimension and number of the given data. Each column x_i has m dimension. Matrix X contains the time series of the thermal images that are vectorized in the matrix form. Now a general LRMF (Low Rank Matrix Factorization) problem can be written as

$$\min_{D, A} \|M \cdot X - DA\|_T \quad (5)$$

where $D \in R^{m \times p}$ and $A \in R^{n \times p}$ are the bases function and the coefficient matrix with $p \ll \min(m, n)$. M is the indicator matrix and T denotes the L_2 or L_1 norm. Eq. (4) is first converted into a sparse representation problem as

$$\arg \min_A \|M \cdot X - DA\|_2 \quad s. t. \|A\|_0 < s \quad (6)$$

here s represents the sparsity parameter. The problem in Eq. (5) can be solved by the sparse representation algorithm of [36] to get the sparse data. This can be re-written under the MLE (Maximum Likelihood Estimation) as

$$x_{ij} = (d_i)^T a_j + e_{ij} \quad (7)$$

where $d_i, a_j \in R^p$ are the i th and j th row vectors of D and A . Where e_{ij} represents the noise present in the data. Under the conventional MLE setup, it is assumed that this noise term is supposed to have a Gaussian or Laplacian distribution. However, in real case this is always not true and this noise has been found to have more complicated configurations [35]. To alleviate this problem in literature authors use MoG (Mixture of Gaussian) to estimate this noise by using the parametric probability distribution [35]. Under this assumption, it can be written as

$$x_{ij} \sim \sum_{k=1}^K \pi_k \mathcal{N}(x_{ij} | (d_i)^T a_j, \delta_k^2) \quad (8)$$

under the i.i.d distribution, the log-likelihood function can be written as

$$L(D, A, \Gamma, \Phi | X) = \sum_{i,j \in \phi} \ln(x_{ij} | \Gamma, \Phi, d_i, a_j) \quad (9)$$

where $\Gamma = [\pi_k]_{k=1}^K$ and $\Phi = [\delta_k^2]_{k=1}^K$ are the mixture rates and variances. The model in Eq. (8) is solved by EM Algorithm. This is a two step algorithm. In (E step) the responsibility parameter is calculated as

$$\gamma_{ijk} = \frac{\pi_k \mathcal{N}(x_{ij} | (d_i)^T a_j, \delta_k^2)}{\sum_{k=1}^K \pi_k \mathcal{N}(x_{ij} | (d_i)^T a_j, \delta_k^2)} \quad (10)$$

Thus, the MoG parameters are $\pi_k = \frac{\mathcal{N}_k}{\sum_{k=1}^K \mathcal{N}_k}$ and $\delta_k^2 = \frac{1}{\mathcal{N}_k} \sum_{i,j} \gamma_{ijk} (x_{ij} | (d_i)^T a_j)^2$, where $\mathcal{N}_k = \sum_{i,j} \gamma_{ijk}$ and the parameter D, A are obtained by solving the L_2 -norm problem in [36], namely

$$\min_{D, A} \|M \cdot X - DA\|_F \quad (11)$$

Instead of using a single MoG model for the entire thermal sequence, separate MoG model for each thermal frame is used and adaptively learning the model from the previous frame gets good results. The model assuming separate MoG for each thermal frame can be given as;

$$x_i^t \sim \prod_{k=1}^K \mathcal{N}(x_i^t | (d_i)^T a, \delta_k^2)^{z_{ik}^t} \text{Multi}(z_i^t | \Gamma) \quad (12)$$

where x_i^t is the i th pixel of the x^t and Multi is the multinomial distribution. Here the natural conjugate priors δ_k^2 and Γ have the *Inverse-Gamma* and *Dirichlet* distributions as follows;

$$\delta_k^2 \sim \text{Inv-Gamma}\left(\delta_k^2 | \frac{\mathcal{N}_k^{t-1}}{2} - 1, \frac{\mathcal{N}_k^{t-1} \delta_k^{t-1,2}}{2}\right) \quad (13)$$

$$\text{Gamma Dir}(\Gamma | a), a = (\mathcal{N}_1^{t-1} \pi_1^{t-1} + 1, \dots, \mathcal{N}_K^{t-1} \pi_K^{t-1} + 1) \quad (14)$$

where $\mathcal{N}^{t-1} = \sum_{k=1}^K \mathcal{N}_k^{t-1}$, $\pi_k^{t-1} = \frac{\mathcal{N}_k^{t-1}}{\mathcal{N}^{t-1}}$. For the conjugate prior in the above equations, there maximums are Φ^{t-1} and Γ^{t-1} . This proves the fact that these prior learn the information from the previous data and encode it into the current model which in turn helps to correct the current MoG and other parameters. The subspace D , assumes to follow a Gaussian distribution prior and it can be given as

$$D_i \sim \mathcal{N}\left(D_i | D_i^{t-1}, \frac{1}{\rho} B_i^{t-1}\right) \quad (15)$$

where $\frac{1}{\rho} B_i^{t-1}$ represents positive semi-definite matrix. This facilitates the fact that current subspace variable D follows the previously learned D^{t-1} one. The full graphical model is shown in Fig. 2. The hyper parameters are given by Θ^{t-1} and after the latent variable z^t is marginalized then the posterior distribution of (Γ, Φ, A, D) can be given as

$$p(\Gamma, \Phi, A, D | x^t, \Theta^{t-1}) \propto p(x^t | \Gamma, \Phi, A, D) p(\Phi | \Theta^{t-1}) p(\Gamma | \Theta^{t-1}) p(D | \Theta^{t-1}) p(A) \quad (16)$$

According to the MAP principle, we can write the minimization problem for $(\Gamma^t, \Phi^t, A^t, D^t)$ as follows

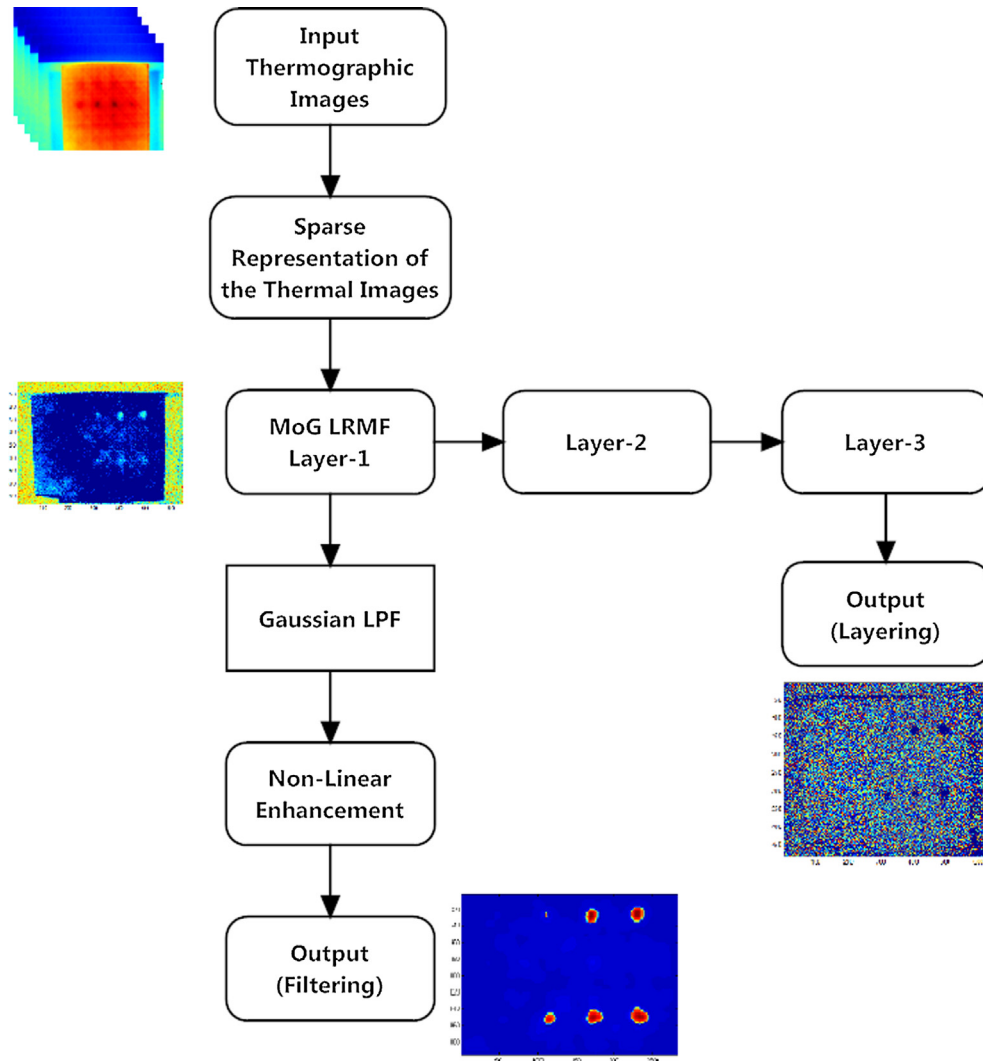


Fig. 2. Flow graph of the proposed algorithm.

$$L^t(\Gamma, \Phi, A, D) = -\ln p(x^t | \Gamma, \Phi, A, D) + \mathfrak{R}_F^t(\Gamma, \Phi) + \mathfrak{R}_B^t(D) \quad (17)$$

where

$$\ln p(x^t | \Gamma, \Phi, A, D) = \sum_i \ln \left(\sum_{k=1}^K \pi_k \mathcal{N}(x_i^t | (d_i)^T a, \delta_k^2) \right) \quad (18)$$

$$\mathfrak{R}_F^t(\Gamma, \Phi) = \sum_{k=1}^K \mathcal{N}_k^{t-1} \left(\frac{1}{2} \frac{\delta_k^{t-2}}{\delta_k^2} + \ln \delta_k \right) - \mathcal{N}^{t-1} \sum_{k=1}^K \pi_k^{t-1} \ln \pi_k \quad (19)$$

$$\mathfrak{R}_B^t(D) = \rho \sum_{i=1}^l (d_i - d_i^{t-1}) (B_i^{t-1})^{-1} (d_i - d_i^{t-1}) \quad (20)$$

The first term in the above equation represents the likelihood term, it helps enforce the parameters under learning mechanism to adapt to the current frame. The second term represents regularization term for the noise distribution. The last term is the Mahalanobis distance between the D and D^{t-1} , which helps update the current subspace based on the previous learned one. To solve the problem in Eq. (16), (EM) Algorithm is utilized. This is achieved by applying the E and M step on each frame altogether.

E-Step: In this step the expectation of the posterior probability for the latent variable z_{ik}^t is estimated. This is also called as responsibility term γ_{ik}^t . The equation can be given as

$$E(z_{ik}^t) = \gamma_{ik}^t = \frac{\pi_k \mathcal{N}(x_{ij}^t | (d_i)^T a_j, \delta_k^2)}{\sum_{k=1}^K \pi_k \mathcal{N}(x_{ij}^t | (d_i)^T a_j, \delta_k^2)} \quad (21)$$

M-Step: On updating the MoG parameters Γ and Φ . The following sub-problem is solved

$$L^t(\Gamma, \Phi) = -E_z^t \ln p(x^t | \Gamma, \Phi, a, d) + \mathfrak{R}_F^t(\Gamma, \Phi) \quad (22)$$

The closed form solution is

$$\pi_k = \pi_k^{t-1} \frac{\mathcal{N}_k}{\mathcal{N}} (\pi_k^{t-1} - \pi_k^t); \delta_k^2 = \delta_k^{t-2} \frac{\mathcal{N}_k}{\mathcal{N}} (\delta_k^{t-2} - \delta_k^2) \quad (23)$$

$$\begin{aligned} \mathcal{N} &= l; \mathcal{N}_k = \sum_i \gamma_{ik}^t; \pi_k^t = \frac{\mathcal{N}_k}{\mathcal{N}}; \delta_k^2 = \frac{1}{\mathcal{N}_k} \sum_i \gamma_{ik}^t (x_i^t - (d_i)^T a)^2 \mathcal{N} \\ &= \mathcal{N}^{t-1} + \mathcal{N}_k; \mathcal{N}_k = \mathcal{N}_k^{t-1} + \mathcal{N}_k \end{aligned} \quad (24)$$

For solving the coefficient a , the following sub-problem of Eq. (16) is solved with respect to a as

$$\min_a \|m^t \cdot (x^t - Da)\|_F^2 \quad (25)$$

This is called a weighted least square problem and its closed form solution can be given as

$$a = (d^T \text{diag}(m^t)^2 d)^{-1} d^T \text{diag}(m^t)^2 x^t \quad (26)$$

For the subspace parameter D , the following sub-problem from Eq. (16) is solved, namely

$$L^t(D) = -E_z^t \ln p(x^t, z^t | \Gamma, \Phi, a, d) + \mathfrak{R}_D^t(D) = \|m^t \cdot (x^t - Da^t)\|_F^2 + \mathfrak{R}_B^t(D) \quad (27)$$

and it has closed form solution for each of the row vector as

$$d_i^t = (\rho(B_i^{t-1})^{-1} + (m_i^t)^2 a^t (a^t)^T) (\rho(B_i^{t-1})^{-1} d_i^{t-1} + (m_i^t)^2 x_i^t (a^t)^T) \quad (28)$$

In order to get the updating rule, namely

$$(B_i^t)^{-1} = \rho(B_i^{t-1})^{-1} + (m_i^t)^2 a^t (a^t)^T; c_i^t = (\rho(B_i^{t-1})^{-1} d_i^{t-1} + (m_i^t)^2 x_i^t (a^t)^T) \quad (29)$$

We have $d_i^t = B_i^t c_i^t$. By using the matrix inverse equation, the update rules can be written as

$$B_i^t = \frac{1}{\rho} \left(B_i^{t-1} - \frac{(m_i^t)^2 B_i^{t-1} d^t (d^t)^T B_i^{t-1}}{\rho + (m_i^t)^2 (d^t)^T B_i^{t-1} d^t} \right); c_i^t = \rho c_i^{t-1} + (m_i^t)^2 x_i^t a^t \quad (30)$$

It can be noted here that for updating the D^t in each step, we only need to save the $(B_i^{t-1})_{i=1}^I$ and $(c_i^{t-1})_{i=1}^I$ evaluated in the last step. Due to this, only a fixed storage memory is used. It should also be noted that as the matrix inverse is avoided in the computations we can say that efficiency of the algorithm will be good.

The model in [35] is a single layer model. As the model is proposed for the case of natural images and provides good results and it can be further decomposed to obtain the enhancement visualization. Thus, the use of multilayer structure following the convention from Eq. (7) is proposed, namely

$$X_1 = D_1 A_1 + E_1 \quad (31)$$

Here X represents the data, D , A represents the sparse term and E represents the noise. For layer two we can write the equation as

$$X_2 = f^1(X_1) + D_2 A_2 + E_2 \quad (32)$$

and in general for n layers we can write

$$X_n = f^{n-1}(X_{n-1}) + D_n A_n + E_n \quad (33)$$

Here f^n represents the non-linear activation function used in the multilayer structure. By adopting this layering structure, we want to extract the deep defects information present in the infrared images. We have tested this layering structure for many layers and from experiments, the layer three provides good results as compared to the other layers. For the initialization of the algorithm, a warm start is needed by employing the PCA on a small sample from the data to get the algorithm initializing parameters, the MoG parameters, and subspace learning parameters [35]. The step by step description is provided in the Table 1 and Fig. 2.

2.3. Filtering approach

In the sparse representation and layering approach described in the previous section, the layering techniques serve as a computationally

Table 1

Sparse ensemble MoG low rank matrix factorization.

Input: the MoG parameters: $\{\Gamma^{t-1}, \Phi^{t-1}, N^{t-1}\}$, model variables: $\{[B_i^{t-1}]_{i=1}^I, [c_i^{t-1}]_{i=1}^I, [D^{t-1}]\}$ and data $[x^t]$.
Initialization: $\{\Gamma, \Phi\} = \{\Gamma^{t-1}, \Phi^{t-1}\}, a^t$.
do for all layers;
1. Given the data $[x^t]$; solve for sparse representation using Eq. (6).
2. Until convergence do;
3. Online E-Step: compute γ_{ik}^t by Eq. (21)
4. Online M-Step: compute $\{\Phi, \Gamma, N\}$ by Eq. (23) and $\{a\}$ by Eq. (26).
5. end while.
6. for each $\{D_i^t\}$, $i = 1, 2, \dots, I$. do,
7. compute $\{[B_i^{t-1}]_{i=1}^I, [c_i^{t-1}]_{i=1}^I\}$ by Eq. (30).
8. compute $\{D_i^t\}$ by $D_i^t = B_i^t c_i^t$
9. end for.
Output: MoG Parameters $\{\Gamma^{t-1}, \Phi^{t-1}, N^{t-1}\}$ and model variables $a^t, \{[B_i^{t-1}]_{i=1}^I, [c_i^{t-1}]_{i=1}^I, D^t\}$.

expensive step. As we are repeating the algorithm three times to get the desired results, the computational cost also increases three times, which is acceptable when compared with the algorithm in [34] which utilizes 4 layer mechanism to get the desired results. However, to make the algorithm fast and robust, we propose a Gaussian index low pass filtering and enhancement approach. Here, we avoid the layering mechanism and take the data processed up to only layer one.

Let $A(m, n)$ be the clean image obtained after the layer one of the proposed algorithm in Table 1. The Gaussian low pass filtering function can be given as

$$\tilde{A}(m, n) = A(m, n) * H(m, n) \quad (34)$$

Here, $\tilde{A}(m, n)$ is the filtered image and $H(m, n)$ is the filtering function. It should be noted here that the filtering is performed in the frequency domain. So, the image $A(m, n)$ is transformed into the frequency domain by the Fast Fourier Transform (FFT) and after filtering converted back into the time domain by the Inverse Fast Fourier Transform (IFFT). The filtering function $H(m, n)$ can be described as

$$H(m, n) = \exp \left[-\frac{D(m, n)^2}{2 \cdot D_0^2} \right] \quad (35)$$

where D_0 represents the cutoff frequency and $D(m, n)$ is the distance from the point (m, n) in the image to the centre and it can be described as

$$D(m, n) = \sqrt{\left[\left(m - \frac{M}{2} \right)^2 + \left(n - \frac{N}{2} \right)^2 \right]} \quad (36)$$

This Gaussian low pass filtering removes the unwanted noise in the image and preserves the defect information present in the image. As we know that the filtering is an average and add mechanism. In this mechanism by averaging most of the noise present in the background is removed and the image becomes smooth. However, a small portion of noise is still present. To completely remove this noise and improve the spatial contrast of the image we propose a simple non-linear enhancement approach and it can be given as

$$\tilde{\tilde{A}}(m, n) = \begin{cases} k_1 \tilde{A}(m, n) > T \\ k_2 \tilde{A}(m, n) < T \end{cases} \text{ for } m, n = 1, 2, 3, \dots, M, N. \quad (37)$$

Here, $\tilde{\tilde{A}}(m, n)$ is the enhanced image and T represents the threshold. Where k_1, k_2 are the weights used for the enhancement of the defects. In our experiments with the given CFRP specimen T is chosen as $\text{Max}[\tilde{A}(m, n)]$ with the deviation of 10, 20 and 30 in the experiments. Here k_1 is chosen as 3 for the enhancement of the defect and k_2 is chosen as 0.25 to remove the background noise from the image. The cutoff frequency is chosen based on the experimental analysis with different values between [0, 1]. In this way this approach searches the whole image for the defects in the filtered image and in the given pixel position if the defect is present, it gets enhanced and if there is noise, it gets removed. The flow chart of the proposed algorithm is shown in Fig. 2. Further, the results are shown in the results and discussion section.

3. Experimental setup

The OPT system is shown in Fig. 3. Generally, for the experiment, we require the specimen for testing, the external heating source, the infrared camera and a computer to store the thermal sequences. A control unit is also required along with computer. For our experiment, we have used halogen lamp as the source of excitation for the CFRP specimen under test. The IR camera is A655sc which comes with an uncooled maintenance free Vanadium Oxide (VoX) microbolometer detector. It produces very clear images. It has high quality and high accuracy. The frame rate used in our experiment is 50 Hz. The IR camera in our lab can distinguish temperature differences as low as 50mK

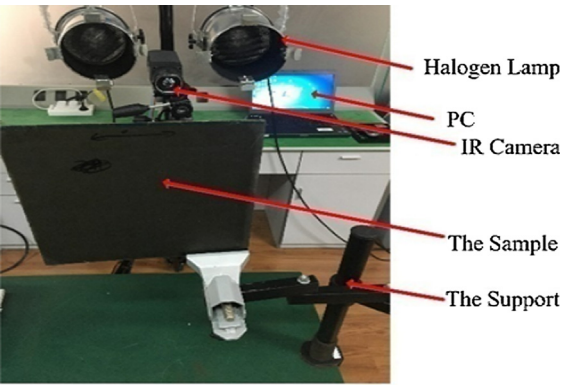


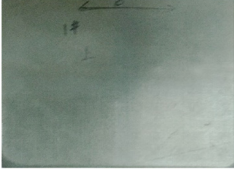
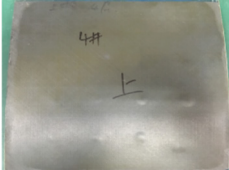
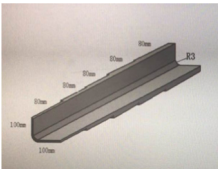
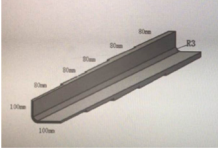


Fig. 3. The optical pulsed thermography system at our lab.

clearly visible. Four test samples are used in the experimental setup. The information about their defect size and dimension can be seen in Table 2. We have selected these CFRP sheets based on their different characteristics.

Table 2
Defect size and depth details.

Specimen number	Dimension (mm)	Defect information (mm)	Picture
1 CFRP/honeycomb core	250 × 300 × 22.2	Depth: 1 or 1.2 Diameter: 3, 6, 10, 14, 18	
2 CFRP/adhesive	250 × 250 × 24.2	Depth: 2 or 2.2 Diameter: 2, 4, 6, 8, 10, 12, 16, 20	
3 CFRP/adhesive	450 × 300 × 22	Depth: 0.375 or 1 Diameter: 6, 10, 15	
4 CFRP/metal	250 × 250 × 22.2	Depth: 2 or 2.2 Diameter: 2, 4, 6, 8	
5 CFRP/metal	100 × 100 × 80	Depth: 2 or 2.5 Diameter: 6, 8	
6 CFRP/metal	100 × 100 × 80	Depth: 2 or 2.5 Diameter: 9, 10	

4. Experimental results and discussion

The proposed algorithm is compared with other algorithms of PCA [1], ICA [32], PPT [20] and TSR [25], EVBTF [34] and [35]. Both the visual and quantitative comparisons are shown. For the quantitative analysis, we have the event based F-score parameter. The F-score is based on the precision and recall parameters and it's details can be found in [30]. Fig. 4 shows the comparative results of the specimen 1. In Fig. 4, the results obtained from PPT and TSR algorithm, only 8 out of the 16 defects are visible whereas they are still blurry and suffer from noise. The PCA and ICA algorithms at some extend clearly detect a few more defects. The results obtained from EVBTF [34,35] are blurry and incomplete. However, the proposed algorithm is able to detect most of the defects successfully and the background noise is removed as shown in Fig. 4(h), it can be seen that the proposed filtering mechanism provides sharp and clear results and almost all the noise is removed. The results of the specimen 2 are shown in Fig. 5. Looking at the results of the PPT, TSR, PCA and ICA, all these algorithms are only able to detect at most 6 out of the total 10 defects while the background noise is strong. However, the proposed algorithm removes the background noise completely and the defects are much clear. Fig. 6 shows the results of the specimen 3. Here, the defect diameters are large and also the

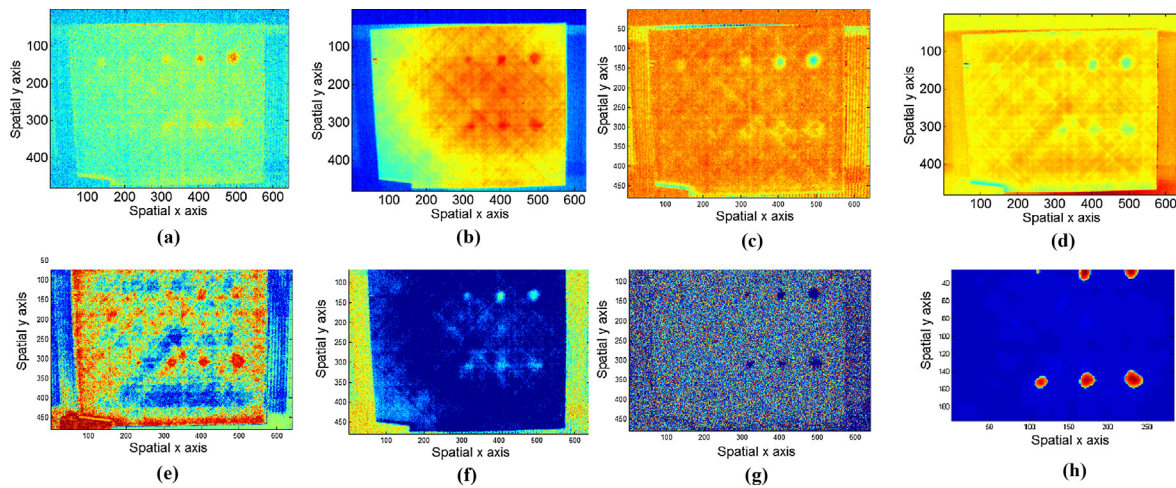


Fig. 4. Visual results for specimen 2 (a) PPT; (b) TSR; (c) PCA; (d) ICA; (e) [34]; (f) [35]; (g) Proposed Layering; (h) Proposed Filtering.

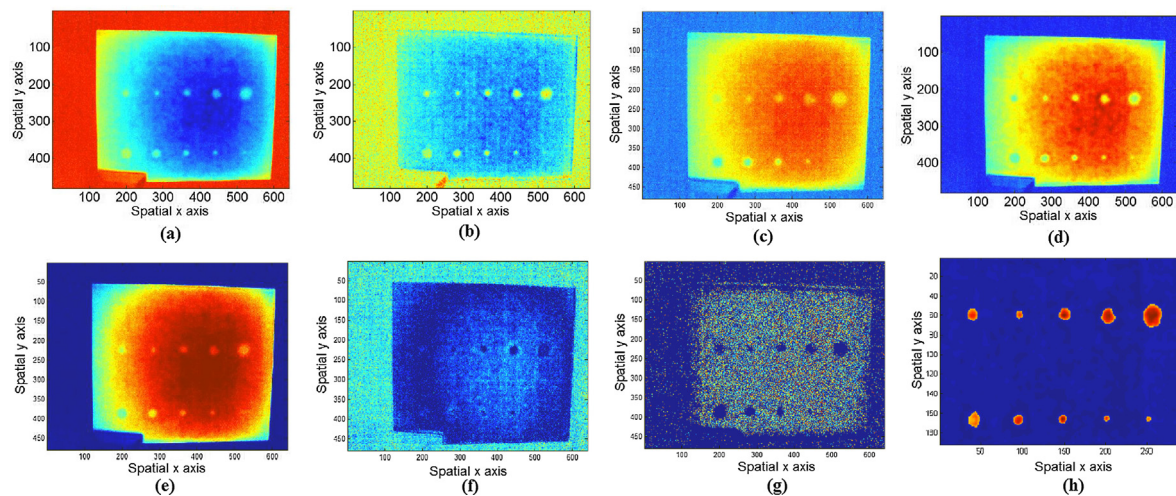


Fig. 5. Visual results for specimen 3 (a) PPT; (b) TSR; (c) PCA; (d) ICA; (e) [34]; (f) [35]; (g) Proposed Layering; (h) Proposed Filtering.

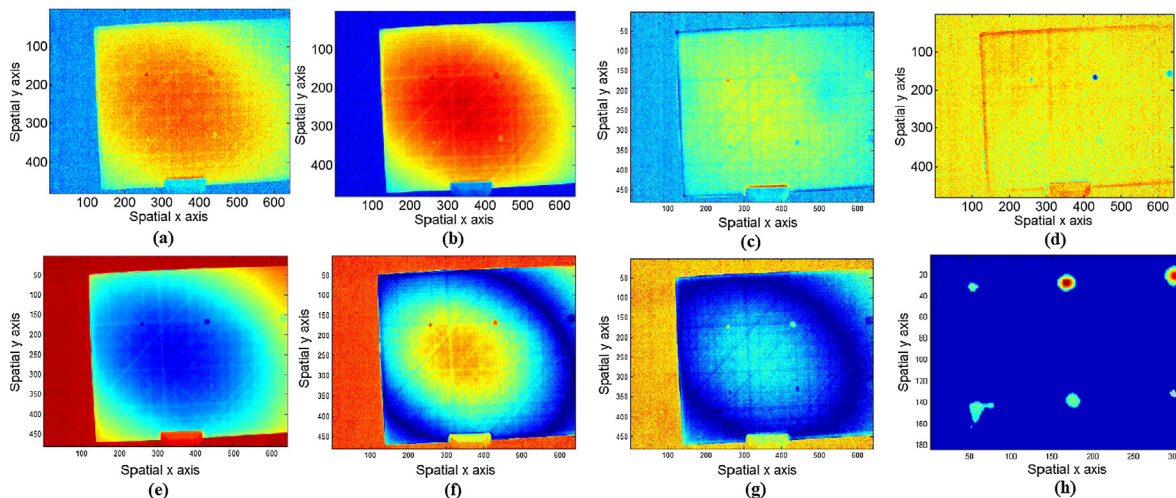


Fig. 6. Visual results for specimen 4 (a) PPT; (b) TSR; (c) PCA; (d) ICA; (e) [34]; (f) [35]; (g) Proposed Layering; (h) Proposed Filtering.

depth of the defects is below 1 mm. All proposed algorithms successfully detect almost all the defects. The results of specimen 4 are shown in Fig. 7. It can be clearly seen that the performance of the proposed algorithm is much better than the other algorithms in terms of detecting defects that are much deeper and smaller. This proves the fact that

sparse representation based low dimensional projection onto a linear and sparse subspace can be utilized efficiently in the debond detection of the CFRP using the OPT. Table 3 shows the F-score [30] results. We have shown the consumption time of different algorithms. The proposed algorithm provides on average 27% increase in detection of

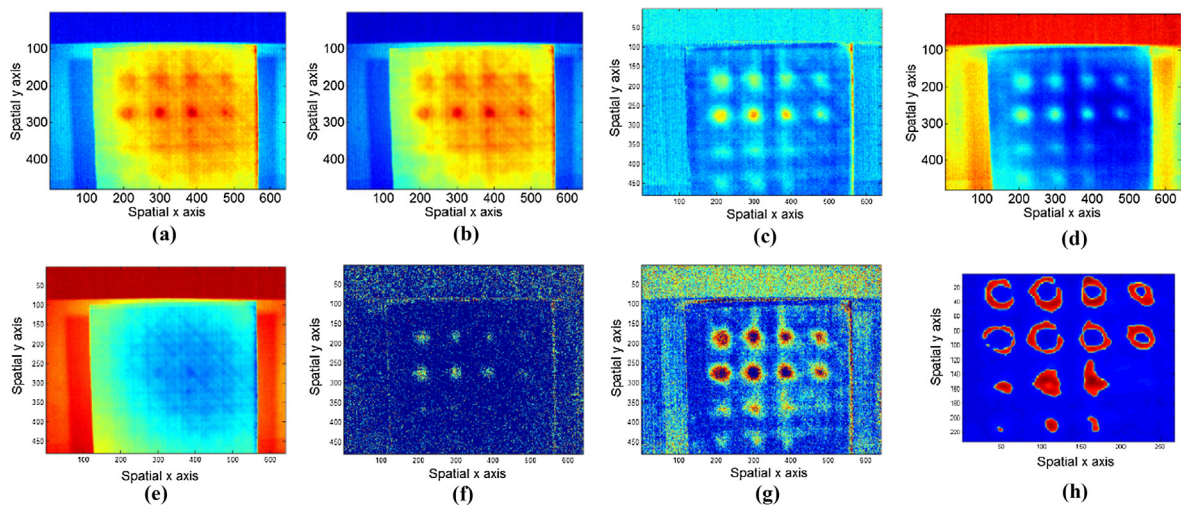


Fig. 7. Layering results for specimen 1 (a) First Layer; (b) Second Layer; (c) Third Layer; (d) Fourth Layer.

Table 3
Performance comparison of F-score (Left) [34] and consumption time (Right in seconds).

Specimen Number	PPT		TSR		PCA		ICA		[34]		[35]		Proposed layering		Proposed filtering	
1	0.66	564	0.66	642	0.85	153	0.85	212	0.30	1019	0.54	277	0.93	466	0.93	328
2	0.46	151	0.75	579	0.46	52	0.75	26	0.75	905	0.75	85	0.75	169	0.75	128
3	0.94	135	0.94	271	0.94	43	0.94	24	0.94	1342	0.75	89	0.94	173	1	137
4	0.66	130	0.66	319	0.80	21	0.66	9	0.66	838	0.66	73	1	120	1	102
5	0.66	129	0.85	241	1	15	1	7	0.85	766	0.85	53	0.66	86	1	74
6	0.66	105	0.66	322	1	15	0.85	7	0.85	686	0.85	57	1	91	1	76
Average	67%	206	75%	396	84%	50	84%	48	72%	926	73%	106	88%	184	94%	140

defects based on the F-score over the PPT algorithm. The percent increase of 19%, 10%, 10%, 22%, and 21% over the TSR, PCA, ICA, [34] and [35] algorithms. Table 3 shows the consumption time for each algorithm. From the Table, it can be seen that the ICA algorithm gives the least consumption time with an average defect detection rate of 84% based on F-score. The average consumption time taken by the PPT, TSR and ICA algorithms are 206 s, 396 s and 50 s. Their average defect detection rates are 67%, 75% and 84%, respectively. The algorithm of [34] gives the highest consumption time on average of 926 s with 72% average defect detection rate. The average time taken by the proposed algorithm is 140 s with 94% (highest) defect detection rate. The algorithm of [35] takes the average time of 106 s with 73% defect detection rate based on F-score. It can be concluded that, with some increase in the computational cost, the proposed algorithm gives good defect detection performance.

This justifies the fact that the proposed algorithm performs better in removing the noise from the thermal images and is able to detect defects clearly.

Figs. 7 and 8, give the layering comparison for specimen 1 and 2. We have shown up to four layer comparison. Layering the algorithm further provides over fitting of the data and hence results are not good. It can

be seen from the Fig. 7 that the defects are more sharp and clear at the layer 3 and also the background noise is significantly removed.

Figs. 9 and 10 show the visual results for a new and more challenging CFRP sheets, respectively. Rather than the standard square or rectangular sheet, this sample is an elbow-shaped as it can be seen in Table 2. The locations of defects in this specimen are at the angular joint of the elbow and defects are difficult to be detected. The advantage of the proposed algorithm can be clearly illustrated. It provides clear and enhanced image detecting for all the defects where the results from the other algorithms suffer from poor resolution and noise.

To validate the capability of the MoG-LRMF algorithm for estimating the complex noise, we show a simple experimental analysis. Here we generate the synthetic data with rank 5 and add different types of noises into it. The task is to recover the original data. For performance evaluation we use the relative reconstruction error (RRE) parameter along with the estimated rank (ER). For comparison other low rank matrix factorization algorithms are chosen such as PCA [11], RPCA [39], and BRPCA [40]. The results are shown in Table 4. The best results are highlighted in bold. It can be seen in Table 1, when there is no noise, the PCA gives the best RRE and the proposed algorithm gives the second best RRE. In case of the sparse noise and Gaussian noise, the

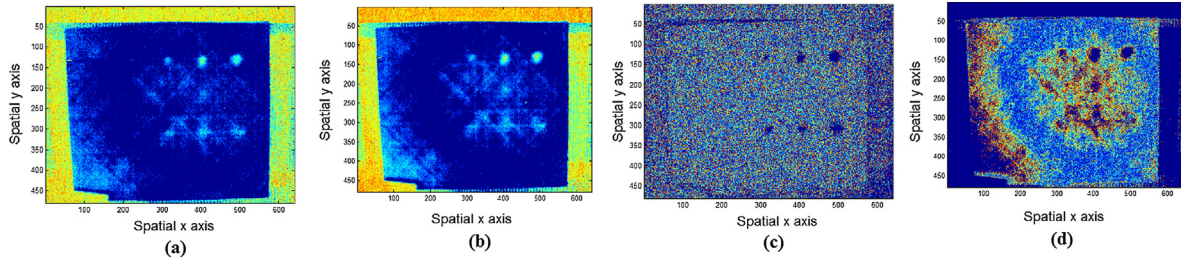


Fig. 8. Layering results for specimen 2 (a) First Layer; (b) Second Layer; (c) Third Layer; (d) Fourth Layer.

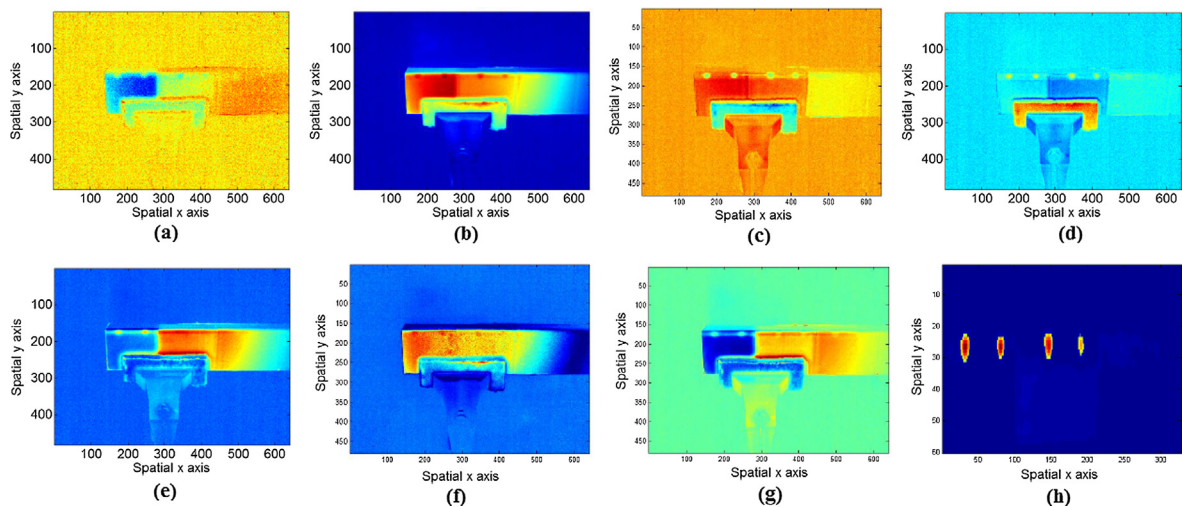


Fig. 9. Visual results for specimen 5 (a) PPT; (b) TSR; (c) PCA; (d) ICA; (e) [34]; (f) [35]; (g) Proposed Layering; (h) Proposed Filtering.

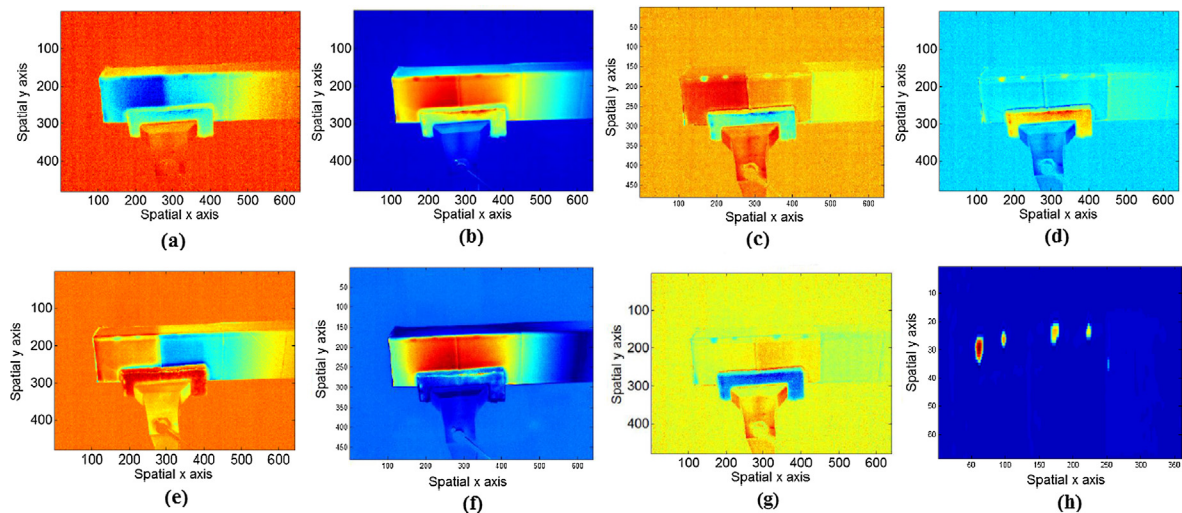


Fig. 10. Visual results for specimen 6 (a) PPT; (b) TSR; (c) PCA; (d) ICA; (e) [34]; (f) [35]; (g) Proposed Layering; (h) Proposed Filtering.

Table 4
Comparative analysis for different types of noise.

Rank (5)		PCA	RPCA	BRPCA	MoG	Proposed (Sparse-MoG)
No noise	RRE	1.11e-13	2.38e-7	0.324	4.99e-5	5.69e-6
	ER	–	5	5	5	5
Sparse noise	RRE	0.678	1.21e-5	6.99e-2	8.19e-5	9.58e-6
	ER	–	5	5	5	5
Gaussian noise	RRE	0.0041	5.11e-2	3.11e-2	3.11e-2	2.96e-3
	ER	–	57	5	5	5
Mixture of noise	RRE	1.10	8.66e-2	6.66e-2	2.40e-3	2.11e-4
	–	–	58	5	5	5

proposed algorithm gives the best RRE of (9.58e-6 and 2.96e-3, respectively). Thus, it recovers the data more accurately. In case of the complex noise, the proposed algorithm gives the RRE of (2.11e-4) which is better than other algorithms. It can be concluded here that Sparse-MoG algorithm is able to model the complex noise more accurately than other algorithms.

5. Conclusion and future work

A sparse ensemble MoG-LRMF approach for detecting the sub-surface debond defects in the CFRP composite is presented. In addition, a

Gaussian filtering and enhancement approach is proposed. The sparse and low-rank projection of the thermal data onto a layered decomposed low dimensional subspace enables to remove the noise and improve its contrast. Further, filtering and non-linear enhancement improved the visual resolution of the defects. In the comparative analysis with other algorithms F-score is used for evaluation. The results indicate that the proposed algorithm is able to quantify the sub-surface debond defects in the CFRP specimen that are both smaller and deeper. Regarding the future work, the algorithm can be tested on more complicate structure specimen apart from CFRP i.e. GFRP, Plexiglas also for other kinds of defects such as delaminations and cracks. In addition, it can be utilized with other active thermography techniques apart from OPT i.e. electromagnetic thermography.

6. Conflict of interest statement

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgements

The work was supported by the Science & Technology Department of Sichuan Province, China (Grant No. 2016GZ0185), National Natural

Science Foundation of China (Grant No. 51377015, No. 61401071 and No. 61527803), Supported by NSAF (Grant No. U1430115) and China Postdoctoral Science Foundation (No. 136413).

References

- [1] R. Nikolas, et al., Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures, *Compos. Struct.* 58 (4) (2002) 521–528.
- [2] F. Henrique, et al., Carbon fiber composites inspection and defect characterization using active infrared thermography: numerical simulations and experimental results, *Appl. Opt.* 55 (34) (2016) D46–D53.
- [3] P. Anish, et al., Non-destructive evaluation of composite repairs by using infrared thermography, *J. Compos. Mater.* 50 (3) (2016) 351–363.
- [4] David K. Hsu, *Nondestructive Inspection of Composite Structures: Methods and Practice*, 2008.
- [5] Rachael C. Waugh, *Development of Infrared Techniques for Practical Defect Identification in Bonded Joints*, Springer, 2015.
- [6] K. Elliott Cramer, et al., *Quantitative NDE of Composite Structures at NASA*, 2015.
- [7] R. Yang, Y. He, Optically and non-optically excited thermography for composites: a review, *Infrared Phys. Technol.* 75 (2016) 26–50.
- [8] Chung-Hyeon Ryu, et al., Nondestructive evaluation of hidden multi-delamination in a glass-fiber-reinforced plastic composite using terahertz spectroscopy, *Compos. Struct.* 156 (2015) 338–347.
- [9] Chulkov, Arseny Olegovich, Vladimir Platonovich Vavilov, *Hardware and software for thermal nondestructive testing of metallic and composite materials*. *Journal of Physics: Conference Series*, vol. 671. No. 1. IOP Publishing, 2016.
- [10] G. Mayr et al., Characterization of defects in curved CFRP samples using pulsed thermography and 3d finite element simulation, in: 9th International Conference on Quantitative Infrared Thermography, Krakow-Poland, 2008.
- [11] B. Chiu, et al., Inspection of cfrp sheet with subsurface defects using pulsed thermographic technique, *Infrared Phys. Technol.* 65 (2014) 117–121.
- [12] Christiane Maierhofer, et al., Characterizing damage in CFRP structures using flash thermography in reflection and transmission configurations, *Compos. B Eng.* 57 (2014) 35–46.
- [13] H. Zhang, et al., Optical excitation thermography for twill/plain weaves and stitched fabric dry carbon fiber perform inspection, *Compos. A Appl. Sci. Manuf.* 107 (2018) 282–293.
- [14] H. Zhang, et al., Optical and mechanical excitation thermography for impact response in Basalt-Carbon Hybrid Fiber-Reinforced composite laminates, *IEEE Trans. Ind. Inf.* 14 (2) (2018) 514–522.
- [15] L. Zhang, et al., Simultaneous spectral-spatial feature selection and extraction for hyper-spectral images, *IEEE Trans. Cybern.* 48 (1) (2018) 16–28.
- [16] L. Zhang, et al., Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding, *Pattern Recogn.* 48 (10) (2015) 3102–3112.
- [17] Liu Yuanlin, et al., Pulsed infrared thermography processing and defects edge detection using FCA and ACA, *Infrared Phys. Technol.* 72 (2015) 90–94.
- [18] B. Yousefi, et al., Comparative analysis on thermal non-destructive testing imagery applying candid covariance free incremental principal component thermography (ccipct), *Infrared Phys. Technol.* 85 (2017) 163–169.
- [19] Xavier Maldague, Sergio Marinetti, Pulse phase infrared thermography, *J. Appl. Phys.* 79 (5) (1996) 2694–2698.
- [20] Sergio Marinetti, et al., Statistical analysis of IR thermographic sequences by PCA, *Infrared Phys. Technol.* 46 (1) (2004) 85–91.
- [21] K. Pearson, On lines and planes of closest fit to systems of point in space, *Phil. Mag.* 2 (1901) 559–572.
- [22] Hyun-Chul Kim, Daijin Kim, Sung Yang Bang, An efficient model order selection for PCA mixture model, *Pattern Recogn. Lett.* 24 (9) (2003) 1385–1393.
- [23] Tao Liang, et al., Low energy impact damage detection in CFRP using eddy current pulsed thermography, *Compos. Struct.* 143 (2016) 352–361.
- [24] S.M. Shepard, *Advances in pulsed thermography*, *Proc. SPIE Thermosense XXIII*, 2001, pp. 511–516.
- [25] Steven M. Shepard, Maria Frendberg Beemer, *Advances in thermographic signal reconstruction*, *SPIE Sensing Technology+ Applications*, International Society for Optics and Photonics, 2015.
- [26] Letchuman Sripragash, Mannur J. Sundaresan, A normalization procedure for pulse thermographic nondestructive evaluation, *NDT & E Int.* 83 (2016) 14–23.
- [27] F. López, et al., Comparative study of Thermographic signal reconstruction and partial least-squares thermography for detection and evaluation of subsurface defects, in: 12th International Conference on Quantitative Infrared Thermography, Bordeaux, France, 2014.
- [28] Bo Peng, Lei Zhang, David Zhang, Automatic image segmentation by dynamic region merging, *IEEE Trans. Image Process.* 20 (12) (2011) 3592–3605.
- [29] Gerald Zauner, M.A.Y.R. Guenther, Guenther Hendorfer, Subsurface defect characterization in pulsed phase thermography by means of wavelet analysis, in: 2nd International Symposium on NDT in Aerospace, Tu, vol. 1.
- [30] Wenwei Ren, et al., Quantitative non-destructive evaluation method for impact damage using eddy current pulsed thermography, *Compos. B Eng.* 54 (2013) 169–179.
- [31] Liu Junyan, et al., Inverse methodology for identification the thermal diffusivity and subsurface defect of CFRP composite by lock-in thermographic phase (LITP) profile reconstruction, *Compos. Struct.* 138 (2016) 214–226.
- [32] Aapo Hyvärinen, Erkki Oja, Independent component analysis: algorithms and applications, *Neural Netw.* 13 (4) (2000) 411–430.
- [33] Qibin Zhao, et al., Bayesian robust tensor factorization for incomplete multiway data, *IEEE Trans. Neural Netw. Learn. Syst.* 27 (4) (2016) 736–748.
- [34] P. Lu, et al., Ensemble variational bayes tensor factorization for super-resolution of cfrp debond detection, *Infrared Phys. Technol.* 85 (2017) 335–346.
- [35] H. Yong, et al., Robust online matrix factorization for dynamic background subtraction, *IEEE Trans. Pattern Anal. Mach. Intell.* PP (99) (2018) 1–1.
- [36] Y.C. Pati, et al., Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, *Signals, Systems and Computers*, Conference Record of the Twenty-Seventh Asilomar Conference on, IEEE, 1993, pp. 40–44.
- [37] D. Meng, et al., Robust matrix factorization with unknown noise, *Computer Vision (ICCV)*, IEEE International Conference, 2013, pp. 1337–1344.
- [38] Y. He, et al., Volume or inside heating thermography using electromagnetic excitation for advanced composite materials, *Int. J. Therm. Sci.* 111 (2017) 41–49.
- [39] Liu, et al., Robust subspace segmentation by low-rank representation, *Machine Learning (ICML)*, IEEE International Conference, (2010).
- [40] Ding, et al., Bayesian robust principal component analysis, *IEEE Trans. Image Process.* 20 (12) (2011) 3419–3430.