



Principal component analysis for fast and automated thermographic inspection of internal structures in sandwich parts

D. Griefahn, J. Wollnack, and W. Hintze

Institute of Production Management and Technology (IPMT), TU Hamburg-Harburg, Denickestrasse 17,
21073 Hamburg, Germany

Correspondence to: D. Griefahn (dominik.griefahn@tuhh.de)

Received: 15 August 2013 – Revised: 3 March 2014 – Accepted: 7 April 2014 – Published: 14 May 2014

Abstract. Rising demand and increasing cost pressure for lightweight materials – such as sandwich structures – drives the manufacturing industry to improve automation in production and quality inspection. Quality inspection of honeycomb sandwich components with infrared (IR) thermography can be automated using image classification algorithms. This paper shows how principal component analysis (PCA) via singular value decomposition (SVD) is applied to compress data in an IR-video sequence in order to save processing time in the subsequent step of image classification. According to PCA theory, an orthogonal transformation can project data into a lower dimensional subspace with linearly uncorrelated principal components preserving all original information. The effect of data reduction is confirmed with experimental data from IR-video sequences of simple square-pulsed thermal loadings on aramid honeycomb-sandwich components with CFRP/GFRP (carbon-/glass-fiber-reinforced plastic) facings and GFRP inserts. Hence, processing time for image classification can be saved by reducing the dimension of information used by the classification algorithm without losing accuracy.

1 Introduction

Lightweight materials – such as sandwich structures – experienced and are forecasted to see a rising demand due to overall increasing transportation volumes especially in aviation. Driven by fuel efficiency requirements, the higher share of lightweight materials also in the traditional transportation industry will further augment this demand. This overall increase continuously drives the manufacturing industry to improve automation in production and quality inspection. Today, sandwich is – thanks to its excellent combination of mechanical strength but also damping properties and the low average material density – a commonly used macro- and micro-composite construction. Sandwich components with carbon- or glass-fiber-reinforced plastic (CFRP/GFRP) facings are typically deployed in rough environments with locally high loadings. In order to cope with heavy concentrated loads or to connect with other structures, components are designed with molded-in inserts, e.g., made from short glass fiber-

reinforced plastic. These inserts replace the honeycomb core to absorb stresses in a defined way (Bitzer, 1997). Quality inspection requires controlling these inserts for presence, correct type, and deviation of geometrical location inside the component after the fabrication step. Due to the mostly in-transparent sandwich facings, normal visual inspection methods fail to perform the described tasks, whereas infrared (IR) thermography combined with image classification algorithms delivers promising results for facing thicknesses below half a millimeter.

Active thermography methods are typically classified by excitation method – namely optical, electromagnetic and mechanical excitation. Most commonly applied methods for composite materials use optical excitation, since these do not require electrical conductivity and are contactless. Lock-in thermography with modulated optical excitation is typically deployed for defects at high depth relative to their size, being comparatively time consuming due to the load modulation.

Pulsed thermography methods are characterized by a shorter cycle time but lower depth resolution. Both are commonly used methods and established for the testing of small lots at laboratory level (Maldague, 2001; Ibarra-Castanedo et al., 2009).

This study evaluates the potential of the square-pulsed thermography for detection of macroscopic subsurface structures in large sandwich components and shows an approach for automated inspection.

2 Background on principal component analysis and automated detection in IR sequences

As described, thermography is a very well investigated NDT (nondestructive testing) method with many different technical variants for the active testing approach (Maldague, 2001; Ibarra-Castanedo et al., 2009). All techniques aim at maximizing contrast directly in the thermal image or to apply algorithms to create or improve contrast in a second step. Principal component thermography (PCT) is a computational approach for analyzing thermal material behavior over time (Rajic, 2002), further improvement can be obtained with contrast enhancement methods and thermal behavior modeling (Omar et al., 2010; Feuillet et al., 2012). An automation of the qualitative PCT approach can be achieved by adding a supervised learning step for image classification (Marinetti et al., 2004).

2.1 Principal component analysis (PCA) using singular value decomposition (SVD)

Principal component analysis is a technique widely used in the context of machine vision (e.g., face recognition or remote sensing), but also for image and video compression. PCA applies a linear transformation to a group of correlated variables in such a way that the obtained set of transformed variables is uncorrelated (Jackson, 1991). The principal components are typically computed via a SVD.

In order to perform a PCA using SVD on infrared video sequences (spatial temperature information over time) the 3-D thermographic data need to be rearranged into a 2-D matrix. Image information (n_x -by- n_y), where n_x and n_y represent the number of photosensitive elements on the sensor in x and y direction, is reshaped into an $n_x \cdot n_y$ -by-1 matrix for every time step. This operation preserves the original spatial information of temperature on the specimen surface, since the reverse transformation is unique. The subsequent transformation of all n_t time steps in the video sequence creates an $n_x \cdot n_y$ -by- n_t matrix \mathbf{A} in which time variations are stored column-wise and spatial variation row-wise.

According to the theory of SVD, any matrix \mathbf{X} (P -by- Q , $P \leq Q$) can be factorized as follows:

$$\mathbf{X} = \mathbf{\Omega}\mathbf{\Gamma}\mathbf{V}^T, \quad (1)$$

where $\mathbf{\Omega}$ is a Q -by- Q matrix, $\mathbf{\Gamma}$ is a P -by- Q matrix with positive or zero diagonal elements representing the singular values and \mathbf{V}^T is the transposed of a P -by- P matrix. The decomposition of IR-data in the matrix \mathbf{A} (M -by- N , $M = n_x \cdot n_y$ and $N = n_t$ and therefore $M > N$) can be determined by computing and decomposing $\mathbf{A}\mathbf{A}^T$ or using the “reduced” or “economy” SVD form to obtain

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (2)$$

where \mathbf{U} is an M -by- N matrix containing spatial information in the orthogonal space. Since spatial information in \mathbf{A} are arranged vertically, the columns of \mathbf{U} represent a set of orthogonal statistical modes called empirical orthogonal functions (EOF) (Emery and Thomson, 2004). The rows of \mathbf{V}^T describe the characteristic time behavior of the corresponding orthogonal function – called principal component vectors building the principal component space. The vectors can provide a measure for time behavior and characterize the defect depths in the material. The matrix \mathbf{S} is an N -by- N diagonal matrix with the singular values s_j of \mathbf{A} . The principal components are obtained scaling the EOFs by multiplying \mathbf{U} with \mathbf{S} or by projecting \mathbf{A} via a multiplication with \mathbf{V} into the principal component space.

It can be shown that

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{S}^2\mathbf{U}^T \quad (3)$$

to derive that the singular values s_j are the square roots of the positive eigenvalues of $\mathbf{A}\mathbf{A}^T$, which is the co-variance matrix of \mathbf{A} multiplied with the factor $(M - 1)$. This relationship allows creating a relative measure for the share of cumulated variance included in the first i EOFs.

$$v_{\text{EOF}}(i) = \frac{\sum_{j=1}^i s_j}{\sum_{j=1}^{n_t} s_j} \quad i \in [1, n_t] \quad (4)$$

Earlier investigations state that more than 95 % of variance can be contained in the first three to five statistical modes and respective components (Marinetti et al., 2004).

2.2 Instance-based learning with k -nearest neighbor

Instance-based learning is used for classification when an explicit description of the target function is not available. The instance-based algorithms store training data for classification of future instances. The k -nearest neighbor algorithm is the most basic and very common kind of instance-based learning classifiers.

The k -nearest neighbor algorithm classifies points in an n -dimensional space based on the Euclidean distance to the k nearest points in the training sample. Depending on the selection of k the classification result can differ (Mitchell, 1997). The algorithm is a powerful tool for classification of multi-attribute instances with high robustness to noise for sufficiently large sets of training data.

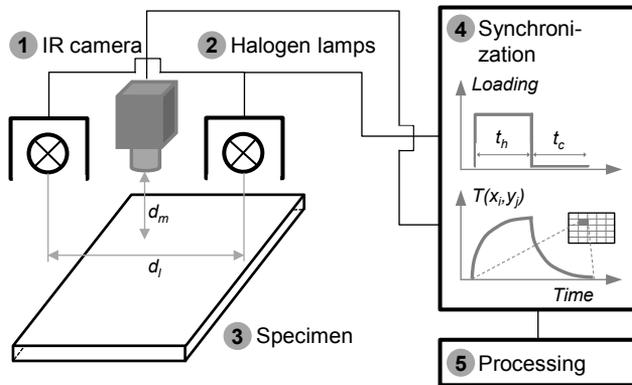


Figure 1. Test field for infrared inspection.

The approach for supervised learning on IR sequences is to generate sets of training data from EOFs with known geometries for different materials and corresponding test settings.

3 Experimental setup for square-pulsed thermographic inspection

The following section describes the experimental setup for the square-pulsed thermographic inspection including test field configuration, the deployed sandwich specimen, and the test settings

3.1 Test field configuration

Figure 1 shows the test field setup used for the described experiments. An IR camera (1) is installed equilaterally with two 400 W halogen lamps at a lateral distance of $d_l = 250$ mm (2), which are used as heat sources. The halogen lamps as well as the IR camera have a distance of $d_m = 200$ mm to the tested sandwich specimen. The deployed IR camera is an Optris PI400 with sensor resolution of 382 pixels by 288 pixels, a thermal sensitivity of 80 mK, and maximum measurement frequency of 80 Hz. The halogen lamps are equipped with a reflector plate in order to homogenize the radiation on the specimen (3) surface. The camera captures the radiation emitted by the specimen's surface.

The synchronization unit (4), which is also linked to the IR-camera recording software, triggers the halogen lamp via a relay and applies the thermal loading during the heating phase. The camera records the heating and the cooling phase. The algorithms for SVD and k -nearest neighbor described in Sect. 2 subsequently perform the processing (5) using MATLAB (version 2012b). LabVIEW (version 2011) coupled with a digital I/O (input/output) device synchronizes the measurements.

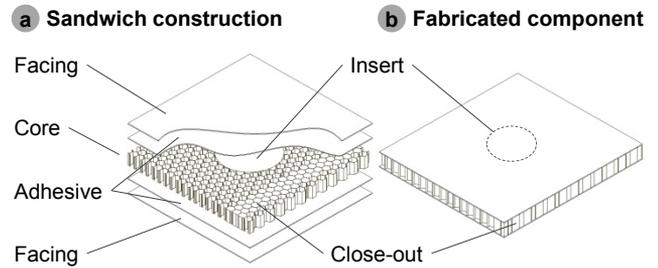


Figure 2. Schematic setup of sandwich structures with close-out and inserts (a) and fabricated component (b) according to Bitzer (1997).

3.2 Tested specimens

Sandwich panels (Fig. 2) are generally built from a dense and strong facing, an adhesive layer and a core. The role of the adhesive layer is to bond the facing to the top and bottom sides of the lightweight core.

Facing material can be metallic such as steel, titanium or aluminum as well as nonmetallic material such as glass fiber, Kevlar-reinforced plastic or carbon-fiber-reinforced plastic. For composite materials such as prepregs, the matrix material may substitute the effect of the adhesive layer (Bitzer, 1997). The earliest core material used for aviation purposes was balsa wood after World War I, and is still in use for some applications. Mostly for nonaerospace applications, expanded polymer foams and aluminum foams can be found as core material today. Honeycomb cores clearly dominate all other cores in aerospace. Honeycomb core structures can be produced from almost all typical lightweight materials such as aluminum, regular, and reinforced polymers or paper. Aramid-fiber paper impregnated with phenolic resin is today's most used honeycomb material (Karlsson and Åström, 1997).

Figure 2a additionally shows an example for a close-out and a high-strength insert element. Close-outs fulfill the function of mechanical protection of the component's edges and a barrier for humidity penetration. Close-outs are added cofabricated during master shaping as polymer filling of the honeycombs as shown in Fig. 2 (Bitzer, 1997).

3.2.1 Sandwich panel fabrication

Autoclave and pressing are the two main methods for bonding sandwich components at industrial scale. Autoclaves are used for curve-shaped components. Hydraulic presses are deployed for flat components and can produce large panels with sizes of up to 3000 mm \times 15 000 mm (Euro-Composites[®], 2013). Several smaller components are typically fitted onto larger master plates hence separated and trimmed in a secondary machining step. The tested specimens are produced on a multilayer press and are machined into manageable sizes

of less than $800 \text{ mm} \times 800 \text{ mm}$, but only smaller sections are inspected.

3.2.2 Specimens material

All specimens used for the experiments are fabricated from typical aircraft-grade materials and produced under condition of mass production for the aviation industry.

For test purposes, two types of pressed honeycomb sandwich modifications with intransparent facings have been selected – one with a CFRP-based facing and the other with a GFRP-based facing. The aramid honeycomb core used has a cell-size of 3.2 mm. The CFRP specimen, with a total thickness of 9.7 mm, is covered with a 0.5 mm facing based on woven carbon-fiber phenolic resin and bonded to the aramid honeycomb core. The GFRP specimen, with a total thickness of 15.5 mm, is composed of a 0.3 mm woven fiberglass facing with phenolic resin and also bonded to an aramid honeycomb. For in-service reasons, the GFRP sample is covered with a thin but intransparent polymer protection foil. Both are equipped with GFRP inserts and in the potting step locally filled with thermoplastic polymer for edge close-out. The diameter of the inserts, locally replacing the honeycomb core of the specimens, is 18 mm for the CFRP specimen and 45 mm for the GFRP specimen. The dimensions of edge close-outs range from approximately 10 mm for the CFRP specimen to 20 mm for the GFRP specimen; given the accuracy of the potting and the filling behavior of the honeycomb cells the width varies a few millimeters.

3.3 Test settings

Heating time t_h and cooling time t_c have to be selected depending on the material of the sandwich facing (see Sect. 4.1 for the exact settings) and the facing layer thickness. CFRP facings require increased heating time or higher power of halogen lamps. This is due to the high heat flow transversal to the test direction given the higher conductivity of carbon compared to glass fibers. The required spatial resolution for the purpose of quality assurance defines the distance between camera and specimen resulting from the field of view.

4 Experimental results

4.1 Contrast improvement via PCA

Video sequences are acquired with heating time $t_{h,\text{GFRP}}$ and cooling time $t_{c,\text{GFRP}}$, each of 10 s, on a GFRP sample resulting in a sequence length $t_{m,\text{GFRP}}$ of 20 s for the total measurement. The CFRP sample was tested at a heating time $t_{h,\text{CFRP}}$ and cooling time $t_{c,\text{CFRP}}$, each of 20 s, resulting in a sequence length $t_{m,\text{CFRP}}$ of 40 s for the total measurement. In order to investigate the influence of the amount of provided input data on the PCA, measurement frequency f_m is varied with the steps 0.5, 1, 2, and 3 Hz. The number of resulting images or

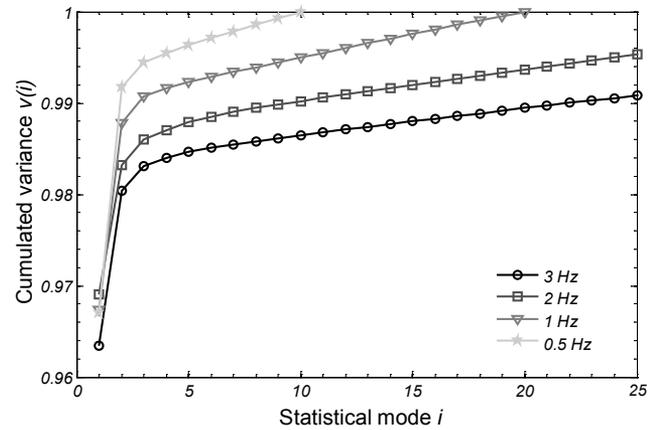


Figure 3. Cumulated and normalized variance v_{EOF} for the first 25 statistical modes.

dimension n_t of the matrix \mathbf{A} is given by the following equation:

$$n_t = (t_h + t_c) \cdot f_m = t_m \cdot f_m. \quad (5)$$

Figure 3 shows the normalized variance v_{EOF} cumulated in the first i so-called statistical modes and corresponding EOFs for the GFRP sample. According to Eq. (3), this measure cumulates the first i singular values in \mathbf{S} corresponding to the first i spatial components or EOFs in columns of the matrix \mathbf{U} . The value is normalized with the total variance.

The first dimension of the video matrix \mathbf{A} contains the total number of $n_x \cdot n_y = 382 \cdot 288 = 110\,016$ elements. At the maximum frequency of 3 Hz and the given recording time, the number of time steps and second dimension n_t of the video matrix \mathbf{A} equals to 60 elements.

The analysis shows that the relatively slow process of heat conduction through materials with partially very low thermal conductivity does not require measurement frequencies above 1 Hz to cover more than 99 % of the time behavior of sandwich material in the first three EOFs. At minimum, the Shannon theorem in the time domain must be fulfilled.

For a measurement frequency of 1 Hz raw, thermal data without emissivity correction are shown on a grayscale in Fig. 4 (GFRP specimen (a) and CFRP specimen (b)) for three selected and representative instances. The first image of the heating phase as well as the first and the last image of the cooling phase are displayed. The SVD is performed in a subsequent step to obtain the EOFs from the thermal data.

Figures 5 and 6 show the effect of decreasing variance on the specimen with GFRP and CFRP facings respectively. It visualizes that the information from an IR-video sequence of square-pulsed thermal images are compressed into three to four EOFs. The retransformation of the columns of matrix \mathbf{U} as described in Sect. 2.1 delivers the spatial information that are scaled to a grayscale image. In contrast to the images from the GFRP sample, the second statistical mode from the

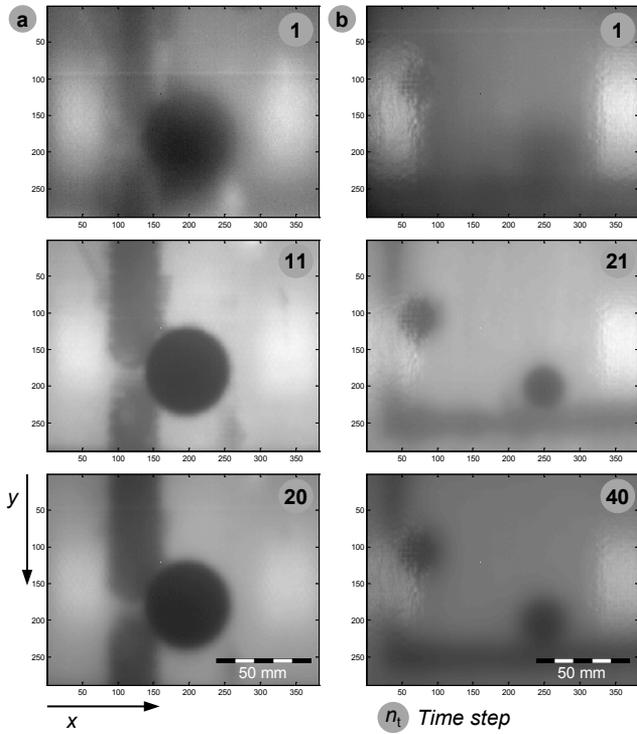


Figure 4. Thermal images at the time steps n_t for the GFRP specimen (a) and the CFRP specimen (b) (first image in heating, first and last image in cooling phase).

CFRP sample contains the direct reflections from the heat sources occurring at the shiny specimen surface.

This section demonstrated that PCA can be applied to square-pulsed thermography for inspection of sandwich components by delivering compressed spatial information with improved contrast between subsurface elements in sandwich structures by preserving time-variation data.

4.2 Automated detection

The aim of the automated detection is to obtain segmented images for the purpose of further inspection. The k -nearest neighbor algorithm requires an amount of preclassified data. These so-called training data are generated from the sandwich samples from known geometric locations (e.g., center of the insert or close-out, plain honeycomb) that are manually classified.

Figure 7 visualized such set of training data for the first three EOFs of the GFRP specimen for approximately 5000 preclassified pixels, which are a subset in the dimension M ($n_x \cdot n_y$). Each data point corresponds to the intensity values from the first three images in Fig. 5 for the same selected pixel. The k -nearest neighbor algorithm uses the training data to classify a “new” instance (in this case the nonclassified data from the sequence) of data – pixel by pixel – based on the Euclidian distance to the k -nearest neighbors. For the ex-

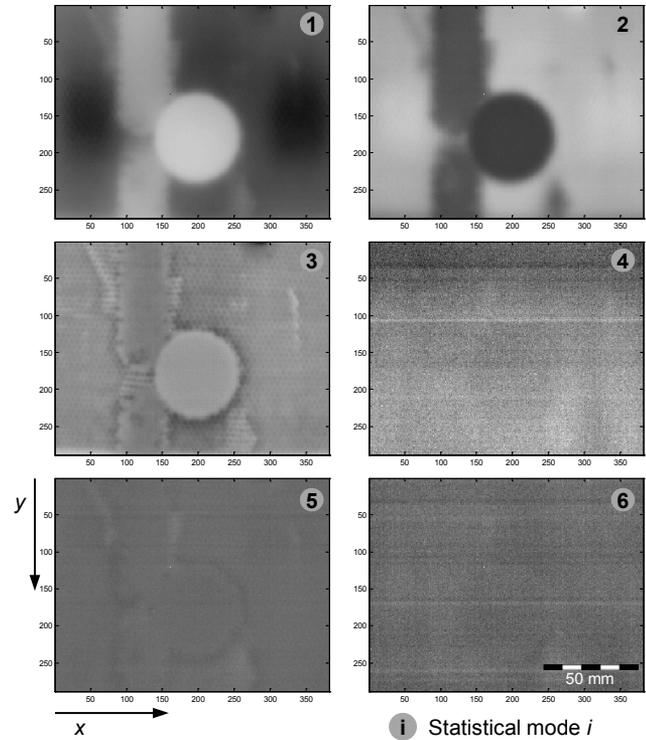


Figure 5. First 6 EOFs of the GFRP specimen.

periments, k varies from three to five depending on the other test settings. In theory, the algorithm can perform the classification tasks in real numbers \mathbb{R}^n , whereas only $i = n_t$ features for classification are available from the IR sequence. Based on the result that 99 % or more of the variance is retained in the first few EOFs, even a reduction of the dimension of the feature space has to be considered.

Two types of definitions for classification errors are used to assess the performance of the algorithm depending on the dimension of the feature space in terms of classification accuracy and computation speed. An algorithm implements the definitions to obtain repeatable and automated results. All inspected parts contain subsurface elements that can be assumed as closed contours on the level of pixel size. If all (or all but one) neighboring pixels in a classified image differ from the class of the selected pixel, it is obviously falsely classified. Figure 8 illustrates the definition of classification errors type I and type II in a segmented image.

Figure 9 shows the results from the classification performance analysis. All data are normalized to 100 % for $i = 1$ to evaluate the relative performance to the smallest possible dimension of feature space. The experiment varies the dimension of the feature space used for classification from one to the maximum possible n_t , and evaluates the number of errors in the classified image as well as meters the elapsed computation time. Using the ten first instead of only the first statistical mode as feature space for classification increases

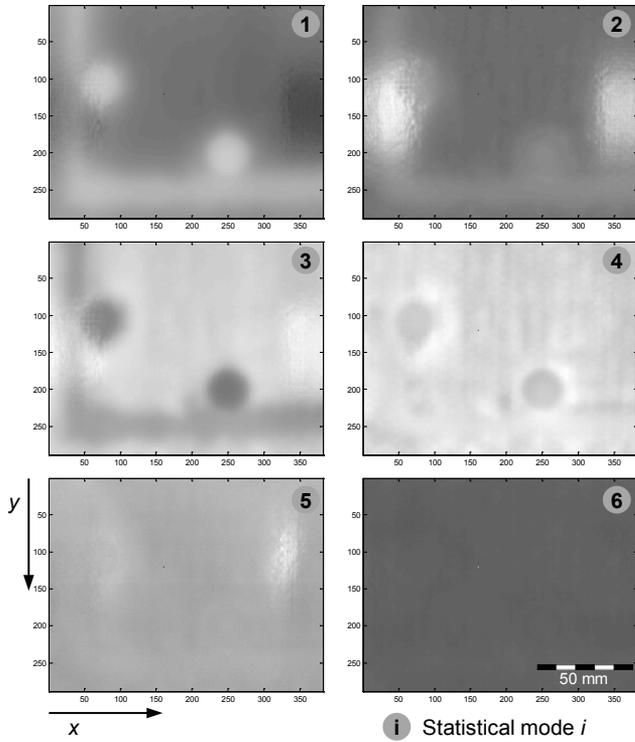


Figure 6. First 6 EOFs of the CFRP specimen.

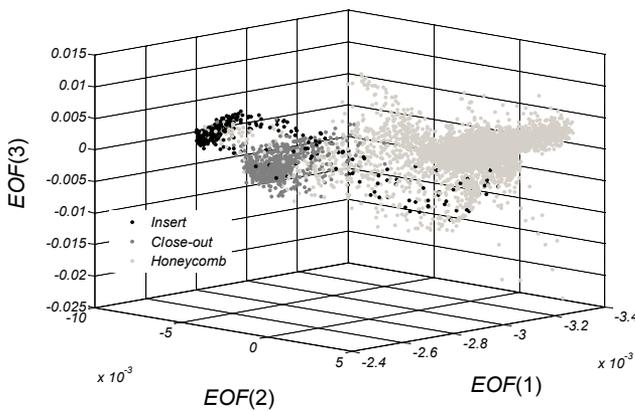


Figure 7. Visualization of first three EOFs of the GFRP-training sample.

the computation effort by a factor of 100, while significantly increasing the computation effort, the classification accuracy does not improve but worsens by 15 % for error type I. Using only the first EOF as feature space for the k -nearest neighbor classification is comparable to applying a histogram-based approach with multiple thresholds. Feature space dimensions between two and five deliver up to 30 % improved results for classification accuracy regardless of the type of error definition and show the advantage of reduced ambiguity. Results from Fig. 3 explain the described effect of falling accuracy when adding statistical modes beyond the fifth one.

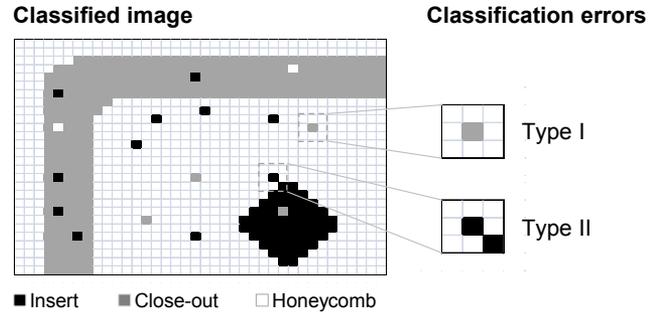


Figure 8. Definition of detectable classification errors by type.

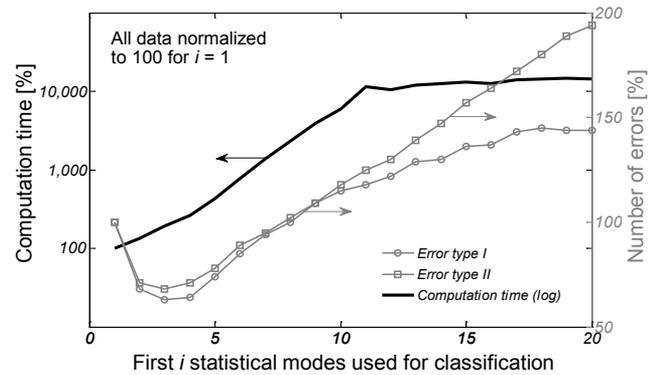


Figure 9. Normalized plot of computation time and number of occurring errors for classification for the GFRP sample based on the first i statistical modes.

Every additional mode contains only a very small amount of additional information useful for classification. It mostly increases noise in the image due to the high rate of data compression.

Figure 10 shows the final results for the classification of the GFRP specimen (a) and the CFRP specimen (b). The GFRP specimen – tested at the setting specified above – is processed using the three first EOFs and based on the three nearest neighbors relationship. Heating and cooling time for the CFRP sample are both increased by 5 s to obtain sufficient contrast. Processing requires including the first five EOFs and using the five nearest neighbors to improve classification results. The color code in the images for the different subsurface structures reflects the classification result. Both tests show that subsurface structures in sandwich components with fully intransparent facing materials are detectable. The images show a specimen of surface of approximately 100 cm², which is tested in less than 30 s including processing time for classification. Decreasing the spatial resolution by pixel in the test setup further improves this ratio.

5 Conclusions

The described experiments demonstrate that PCA on simple square-pulsed thermography in combination with an

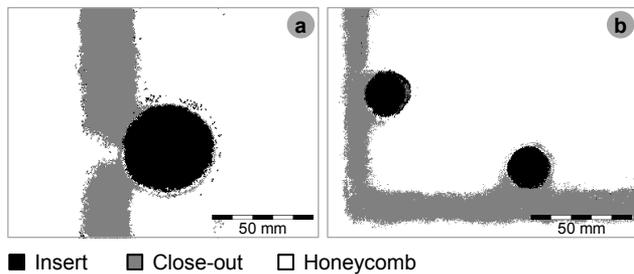


Figure 10. Classification result for the GFRP specimen (a) and the CFRP specimen (b).

instance-based classification algorithm detects and separates subsurface structures in sandwich components with intransparent facing material.

All investigations aimed at identifying suitable test settings to ensure fast and reliable results for the detection. The experiments confirm the effect of data reduction via PCA into the first three to five statistical modes from previous investigations and suggest limiting the measurement frequency to avoid noise from oversampling of a slow thermal process. The numerical evaluation of the cumulated variance in the transformed sequences fortifies the result. Principal component thermography on sandwich components is a very robust technique to improve contrast on square-pulsed-tested IR sequences and has a lower sensitivity to inhomogeneous lighting than e.g., simple threshold methods in image processing.

The combination with the k -nearest neighbor algorithm enhances the setup to a method for automated detection and classification of subsurface structures. Three to five statistical modes covering more than 99 % of variance deliver clearly an optimum result with respect to classification accuracy and a relatively low computation effort.

Future investigations focus on a prediction of thermal behavior of sandwich material based on numerical simulations. This will help to improve the current set of training data in the transition between different subsurface structures and will show approaches for automated population of training data.

Acknowledgements. The project on which this paper is based was funded by the German Federal Ministry of Economics Affairs and Energy under funding code 20W1115C. The authors assume all responsibility for the content of this publication.

Edited by: R. Tutsch

Reviewed by: two anonymous referees

References

- Bitzer, T.: Honeycomb technology: Materials, design, manufacturing, applications and testing, 1st Edn., Chapman & Hall, London, 1997.
- Emery, W. J. and Thomson, R. E.: Data analysis methods in physical oceanography, 2. and rev. ed., 3. impr., Elsevier, Amsterdam, 2004.
- Euro-Composites[®], Infrastructure and production technologies – Panel Production: <http://www.euro-composites.com/en/technology/Seiten/panel.html> (last access: 13 March 2013), 2009.
- Feuillet, V., Ibos, L., Fois, M., Dumoulin, J., and Candau, Y.: Defect detection and characterization in composite materials using square pulse thermography coupled with singular value decomposition analysis and thermal quadrupole modeling, *NDT&E Int.*, 51, 58–67, 2012.
- Ibarra-Castanedo, C., Piau, J.-M., Guilbert, S., Avdelidis, N. P., Genest, M., Bendada, A., and Maldague, X. P. V.: Comparative Study of Active Thermography Techniques for the Nondestructive Evaluation of Honeycomb Structures, *Res. Nondestruct. Eval.*, 20, 1–31, 2009.
- Jackson, J. E.: A user's guide to principal components, Wiley series in probability and mathematical statistics, Wiley, 1991.
- Karlsson, K. F. and Åström, T. B.: Manufacturing and applications of structural sandwich components, *Compos. Part A-Appl. S.*, 28, 97–111, 1997.
- Maldague, X. P. V.: Theory and practice of infrared technology for nondestructive testing, Wiley series in microwave and optical engineering, Wiley, New York, NY, 2001.
- Marinetti, S., Grinzato, E., Bison, P. G., Bozzi, E., Chimenti, M., Pieri, G., and Salvetti, O.: Statistical analysis of IR thermographic sequences by PCA, *Infrared Phys. Techn.*, 46, 85–91, 2004.
- Mitchell, T. M.: Machine learning, International ed., McGraw-Hill series in computer science, McGraw-Hill, New York, NY, 1997.
- Omar, M. A., Parvataneni, R., Zhou, Y.: A combined approach of self-referencing and Principle Component Thermography for transient, steady, and selective heating scenarios, *Infrared Phys. Techn.*, 53, 358–362, 2010.
- Rajic, N.: Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures, *Composite Structures*, 58, 521–528, 2002.