10.21611/qirt.2016.128

Sensitive infrared detection of thermal and optical non-uniformities using correlation and PCA processing

by Z. Suszyński*, R. Świta*, M. Bednarek** and Joanna Łoś***

* Koszalin Univ. of Technology, 70-453, Śniadeckich 2, Koszalin, Poland, *zbigniew.suszynski@tu.koszalin.pl*
** The Main School of Fire Service, 01-629, Słowackiego 52/54, Warszawa, Poland
*** State Higher Vocational School in Koszalin, Leśna 1, 75-582 Koszalin, Poland

Abstract

One of the major tasks of active thermography is the detection and analysis of thermal and optical heterogeneity features in solids. Detection can be an independent or preceding task in the process of identifying the characteristics of heterogeneities (recognition). This article demonstrates that the correlation analysis and independently PCA analysis can increase the sensitivity of heterogeneity detection and classification due to their characteristics. Correlation method is used for calculating correlation coefficient between temperature contrast and reference signals. PCA analysis of the thermographic sequence allows processing to a form of principal images, which classify different types of heterogeneities and improve signal-to-noise ratio. The paper presents examples of optical and thermal heterogeneity analysis.

1. Introduction

Identifying characteristics of thermal and optical heterogeneities in active thermography is based on analysis of changes in temperature or temperature contrast (alternatively amplitude, phase, or composite contrast) of each pixel in a thermographic sequence in time or frequency domain. In order to identify heterogeneity with specific characteristics one can use correlation analysis [1]. It involves measuring the correlation coefficient between contrast signal X for consecutive sequence pixels and the reference contrast X_{ref}. If vectors $\hat{X}$ and $\hat{X}_{ref}$ contain contrast signal values and reference contrast values respectively, decreased by their mean value:

$$\hat{X} = X - \frac{1}{N}\sum_{i=1}^{N} X[i]$$
$$\hat{X}_{ref} = X_{ref} - \frac{1}{N}\sum_{i=1}^{N} X_{ref}[i] \quad (1)$$

then Pearson correlation coefficient is equal to the cosine function of the angle between these two vectors, hence it defines the likelihood of both:

$$Corr\left(X, X_{ref}\right) = \frac{Cov(X, X_{ref})}{\sqrt{Cov(X,X)Cov(X_{ref}, X_{ref})}} = \frac{\hat{X}^T \hat{X}_{ref}}{\|\hat{X}\|\|\hat{X}_{ref}\|} = \cos\sphericalangle(\hat{X}, \hat{X}_{ref}) \quad (2)$$

The correlation image is created based on correlation coefficients, determined for each pixel of the sequence to the reference signal. Such images can be obtained for any contrast sequence and have a very high signal to noise ratio. This allows detection of areas with specific physical properties.

Principal Component Analysis (PCA) is a linear transformation of signal comprised of observations described by correlated parameters into the new coordinate system with parameters uncorrelated and maximizing variance of data. No correlation between parameters means orthogonality between their versors. Using covariance analysis and parameters orthogonality, the method tries to uncover natural data structure and finds important, from describing object properties point of view, new parameters, which are linear combination of original parameters. The number of new parameters is the same as the number of original parameters, because it defines data dimensionality. Contribution of the new parameters to the data variation is a monotonic decreasing function. One shows this function on so called scree plot, which can be drawn after sorting principal components in order of their decreasing influence on data variance. PCA allows data dimensionality reduction by rejection of parameters with minor influence on signal magnitude (low variances) – the data is orthogonally projected on the dimensions that's left. The method is also used in effective data compression, as well as in signal filtering by means of separation of uncorrelated with the signal noises and distortions. It is very often used for visualization of multidimensional data.

For the matrix $\boldsymbol{X}$ of size *N*×*M*, whose rows are the vectors of observation and in which columns contain parameter values describing these observations, PCA is centering parameter values in each column in order to become independent of the shift associated with the mean value of parameters:

$$\widehat{\boldsymbol{X}}[:,k] = \boldsymbol{X}[:,k] - \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{X}[i,k]\ ,\quad k = 1..K \tag{3}$$

The method transforms n-dimensional data $\widehat{\boldsymbol{X}}$ into the orthogonal coordinate system $\boldsymbol{W}$. The coordinates $\boldsymbol{T}$ in the new system are obtained simply by matrix multiplication:

$$\boldsymbol{T} = \widehat{\boldsymbol{X}}\boldsymbol{W} \tag{4}$$

Because matrix $\boldsymbol{W}$ is orthogonal, it's easy to invert. The inverse is determined by the transposition of this matrix ($\boldsymbol{W}^{\boldsymbol{T}}$):

$$\widehat{\boldsymbol{X}} = \boldsymbol{T}\boldsymbol{W}^{T} \tag{5}$$

The linear dependence of parameters describing observations is characterized by the covariance matrix $\widehat{\boldsymbol{X}}^{T}\widehat{\boldsymbol{X}}$, which can be now described as:

$$\widehat{\boldsymbol{X}}^{T}\widehat{\boldsymbol{X}} = (\boldsymbol{T}\boldsymbol{W}^{T})^{T}\boldsymbol{T}\boldsymbol{W}^{T} = \boldsymbol{W}(\boldsymbol{T}^{T}\boldsymbol{T})\boldsymbol{W}^{T} \tag{6}$$

If the observation parameters are scaled differently or are incomparable, one should use normalization of the matrix $\widehat{\boldsymbol{X}}$ columns with their variance. Covariance matrix $\widehat{\boldsymbol{X}}^{T}\widehat{\boldsymbol{X}}$ becomes then a correlation matrix.

No correlation between the parameters in the new coordinate system means that the covariance matrix $\boldsymbol{T}^{T}\boldsymbol{T}$ must be a diagonal matrix. The matrix $\boldsymbol{W}$ is therefore an orthogonal matrix of eigenvectors of the covariance matrix $\widehat{\boldsymbol{X}}^{T}\widehat{\boldsymbol{X}}$, and the diagonal of matrix $\boldsymbol{T}^{T}\boldsymbol{T}$ - vector of its eigenvalues. The PCA transformation therefore is associated with the problem of determining eigenvectors of the data covariance matrix. Geometric illustration of such a problem can be presented as a multi-dimensional ellipsoid describing cloud of data with axis directions consistent with the eigenvectors and axis lengths equal to the eigenvalues of the data covariance matrix. PCA transformation can be then interpreted as a combination of rotations of this ellipsoid, after which all of its orthogonal axes coincide with coordinate system versors, whereby each versor from the next dimension corresponds to the shorter axis of ellipsoid.

2. Principal images

Thermographic sequences are characterized by the constancy of the scene with dynamically changing contrasts of areas which have different physical properties. Measuring the temperature signal is made independently for each pixel. These measurements can therefore be considered as independent observations in the principal component analysis. It can be expected, that due to the continuity of the temperature signal, information included in subsequent samples of the signal, as well as temperature variations in the neighboring pixels, are strongly correlated. This should allow PCA achieving high processing efficiency, by means of strong concentration of information in the first principal components (images) [2, 3], improving the signal to noise ratio and image segmentation, which takes into account the differences in physical properties.

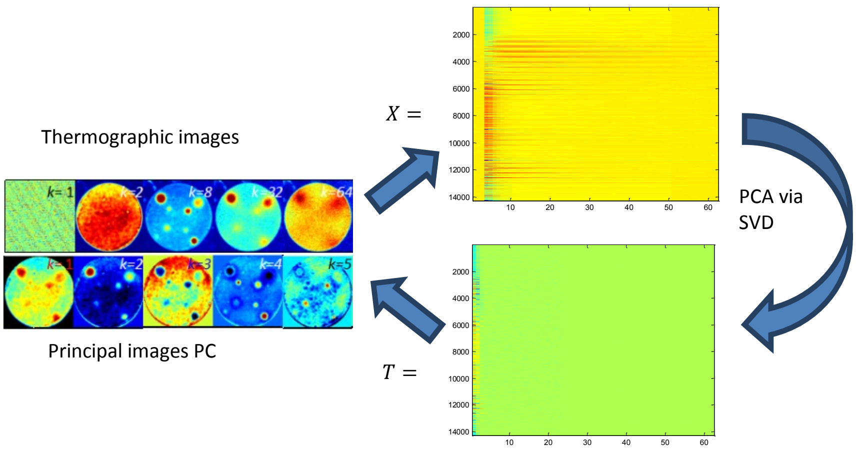

Fig. 1. Construction of principal images

For the thermal sequence with K frames and the frame size of MxN pixels, processed observation matrix $\boldsymbol{X}$ contains MxN rows of pixels and K columns with pixel values in successive frames of the sequence (Fig.1). PCA performs shift of the pixel values by their mean value in each frame, making transformation independent from the average intensity in each frame $\overline{\boldsymbol{X}}[:,k]$:

$$\widehat{\boldsymbol{X}}[:,k] = \boldsymbol{X}[:,k] - \frac{1}{MN}\sum_{i=1}^{MN}\boldsymbol{X}[i,k] = \boldsymbol{X}[:,k] - \overline{\boldsymbol{X}}[:,k]\,, \quad k = 1..K \tag{7}$$

In case of filtration, the average intensities of the frames $\overline{\boldsymbol{X}}[:,k]$ should be remembered, they will be needed for the proper reconstruction of the signal. Sequence encoded in such matter is orthogonalized using eigenvectors of the covariance matrix $\widehat{\boldsymbol{X}}^T\widehat{\boldsymbol{X}}$. If it is important to compare tested heterogeneities qualitatively, not quantitatively, parameters normalization with the variance values may be used, which means use of the correlation rather than covariance matrix. Due to the high computational complexity, the covariance matrix is not constructed explicitly. Its eigenvectors are determined using SVD matrix factorization $\widehat{\boldsymbol{X}}$. If the regions of interest are labeled, the data in the new coordinate system can be visualized by examining the similarity of pixels in selected areas (Fig.2)

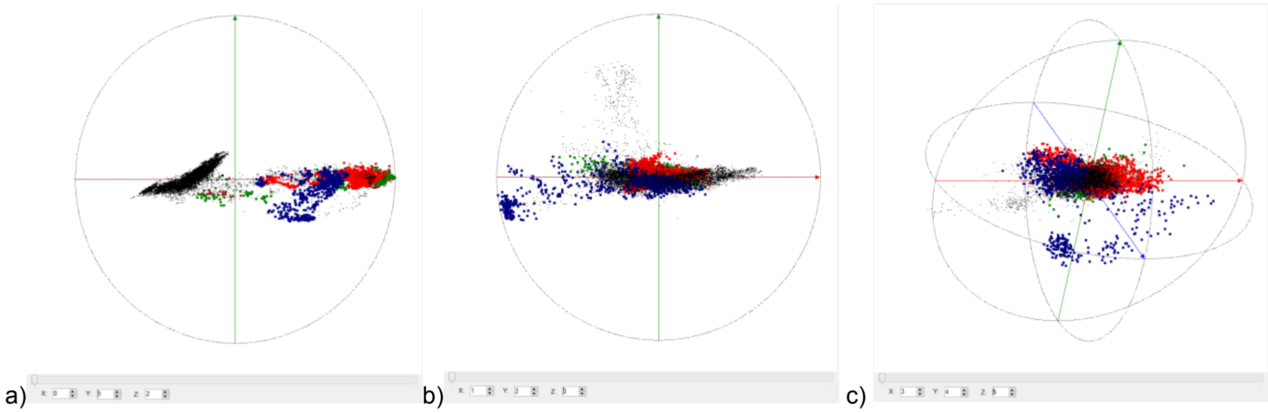

Fig. 2. Projections of the principal components of the thermal sequence with marked regions of interest Components 0-1 a), Components 1-2 b), Components 3-4-5 c)

Data coordinates in a new coordinate system are encoded back into a sequence of frames while keeping the size and frame rate. Signal in the form of principal images (PC) depict the essential contents of the sequence (in the sense of feature extraction - signal variance) in successive, independent frames. The first principal image contains information characteristic to all frames, therefore thermal heterogeneities are extracted often only in subsequent frames. Most of the frames, however, are unnecessary and are associated with high data redundancy. The scree plot presents Importance of principal images. It's a basis for determining the number of important principal images, which are used for saving information in compressed form or to filter data by means of reconstruction from reduced number of PC. In case of filtration one should be aware of eventual undoing of normalization (if the correlation matrix was used instead of covariance) and shifting data by the average values of the original parameters

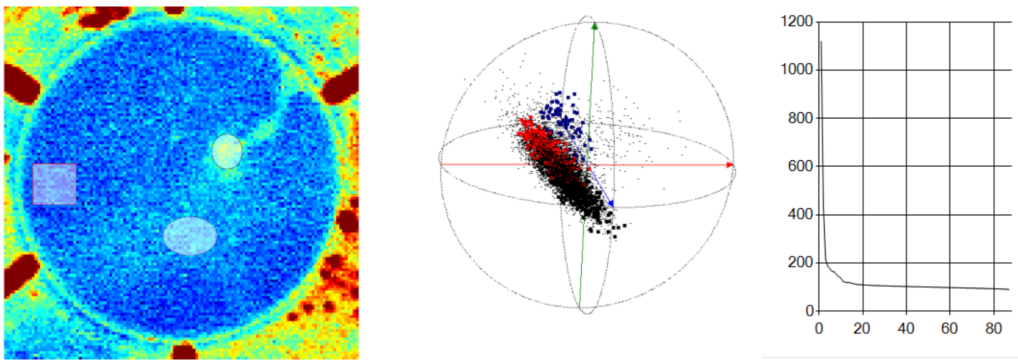

Fig. 3. Frame of the thermal sequence with marked regions of interest a) The PCA space with marked pixel regions b) Scree plot c)

10.21611/qirt.2016.128

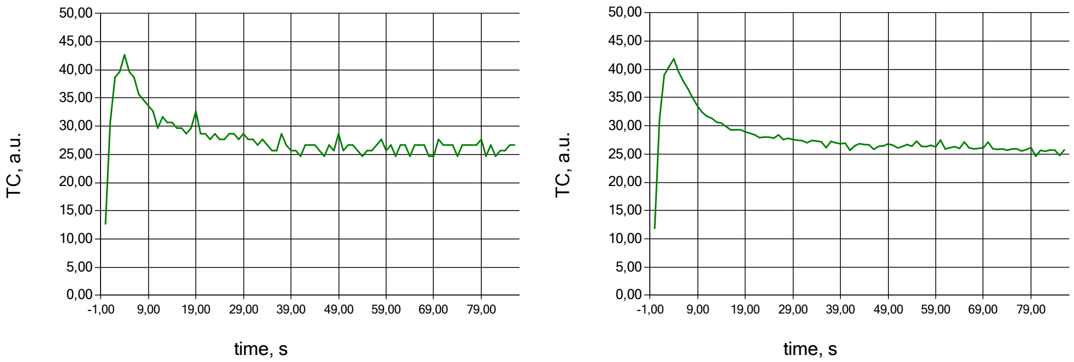

Fig. 4. Filtration using PCA transformation of an exemplary pixel of the sequence from Fig. 3. in the frames direction. The waveform before a) and after filtering, reducing the number of the PC from 87 to 10 b)

The PCA transformation of the sequence can be performed not only in the frames direction, treating it's pixel as an independent observations. Also, the pixels adjacent in the plane of the frame contain highly correlated data. Filtration of the signal can be improved by using all three directions. Due to the smaller number of vertical and horizontal pixels compared to the number of recorded frames, compression and filtration efficiency in these directions is however usually smaller.

3. Analysis

It can be assumed, that the PCA, based on analysis of correlation or covariance data matrixes and correlation analysis have similar thermographic data processing capabilities. This is confirmed by cases described below. Correlation analysis and PCA were applied to two sequences obtained in infrared optical excitation pulse mode. The first sequence was registered for PCB with electronic components and second one for TLC plates with separated analytes (fullerenes C60 and C70) [4].

Images of PCA were determined by the scheme described above and to determine correlation of images reference signal was taken from the specific areas of chosen principal images of PCA sequence. For the PCB it was: a temperature signal on a surface of one of the integrated circuits, the left corner of the PCB (paint), and temperature changes of the resistors. For TLC plate the reference signal was signal for fullerenes temperature (first PCA image) and thickening of the active layer TLC (right and bottom edges of the image).

Fig. 2. shows PCA and correlation mages, obtained from the IR sequence. Both types of analysis enable to separate from images, i.a. integrated circuits, heterogeneity of paint covering the plate and resistors. Contrasts of these groups of elements in correlation and PCA images are clearly better than in the exemplary IR image.

.

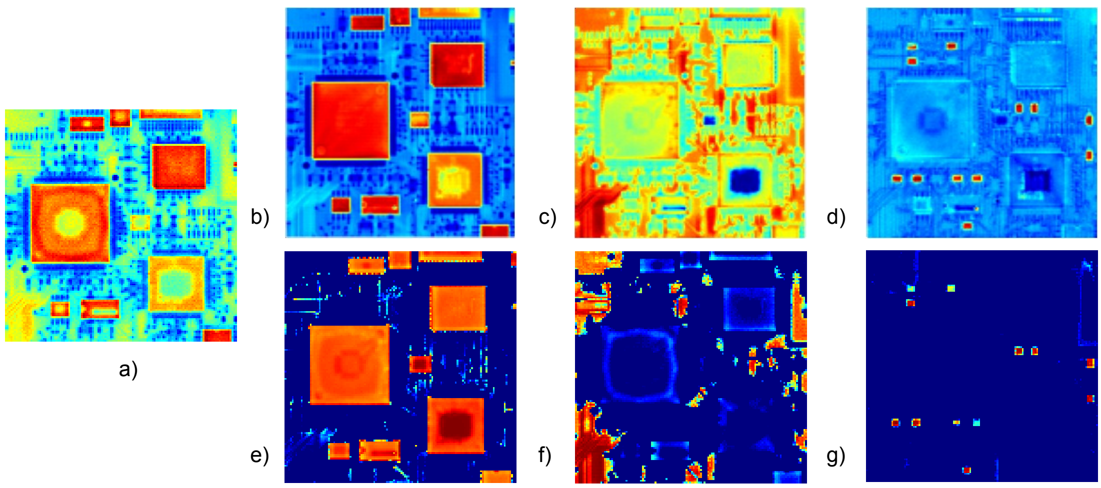

Fig. 3. Image of IR sequence (a), PCA three main images (b, c, d) and three correlation images (e, f, g) of IR sequence obtained for printed circuit board. Individual images reveals: integrated circuits (b, e), non-uniformities of paint (c, f), resistors (d, g)

Fig. 3. shows the corresponding images obtained for thin layer chromatography plate with fullerenes. Images show two types of heterogeneities: optical and thermal [4]. The correlation and PCA analysis allowed separation of both types of heterogeneity and increase in signal to noise ratio. The advantage of the PCA method is lack of necessity to know a priori about the thermal response of heterogeneity - thermal sequence is automatically divided into principal component frames presenting information about objects with different thermal properties. Scree plot can be used to determine the actual number of independent parameters describing the pixels in the thermal sequence and performing optimized segmentation or signal filtering.

PCA image sequences and correlation sequence show similar divisions of images contents (content analysis), and will unveil, among others, details not visible on the image of the original sequence. This is understandable, because each correlation image and each PCA image is a result of transformation condensing information from the entire thermal sequence, while a single sequence image shows only a temporary temperature distribution and dynamic of changes in the temperature of each pixel vary. For this reason the temperature information of each pixel is dispersed in time.

Both methods of thermographic sequence analysis can be used for detection of heterogeneities and their segmentation.

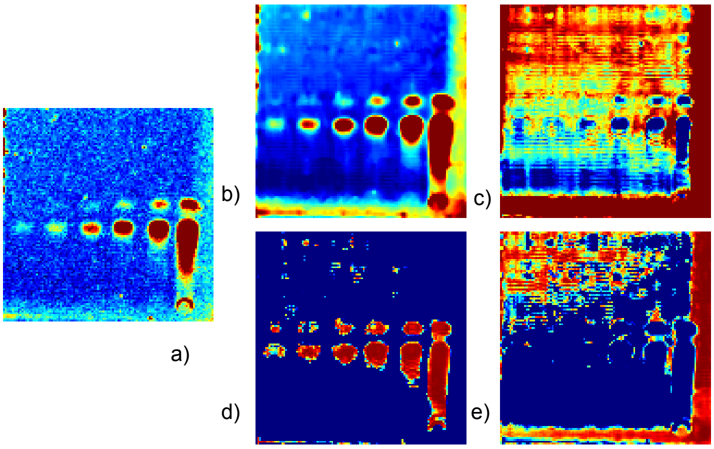

Fig. 4. Image of IR sequence acquired for thin layer chromatography plate with fullerenes (a), two main images of PCA sequence (b, d) and correlation images of IR sequence. Individual images reveal: c60 and c70 fullerene clusters (b,d) and inhomogeneities in thickness and structural changes of stationary phases.

.

Conclusions

Correlation images and principal images (after PCA) enable highly sensitive detection of changes in the thermal properties based on thermographic sequence. Correlation images detect areas, which properties (temperature in the time or frequency domain) correlate with the reference signal. Therefore, their interpretation is relatively simple - the following images show areas with specific thermal characteristics, represented by the reference signal. The reference signal may be calculated on the basis of the model or can be obtained from the analyzed thermographic sequence. PCA images are the result of statistical processing operations, so the final result depends on the context of the image. The first image shows areas with dominant in statistical sense characteristics. Other images show details of the mutually independent (orthogonal), and therefore different, statistical characteristics. In the case of PCA analysis, we deal with the analysis of the thermal sequence based on the covariance or correlation matrix, and in case of correlation images, correlation of observation with reference observation is studied.

Images analysis using correlation or PCA simplify detection of thermal heterogeneities but their interpretation in the case of the PCA analysis is not always obvious.

REFERENCES

[1] Z. Suszyński, M. Bednarek, "Application of the Correlation Function in the Detection of Thermal Heterogeneity Using Pulsed Thermography", International Journal of Thermophysics, September 2012
[2] Suszyński Z., Świta R.: "Processing of thermographic sequence using Principal Component Analysis", Measurement Automation Monitoring, tom 61, nr 06, pp. 215-218, 2015
[3] Rajic N.: "Principal component thermography for flaw contrast enhancement and flaw depth characterization in composite structures". Composite Structures, vol. 58, pp. 521-8, 2002
[4] Z.Suszyński, P.K. Zarzycki, New approach for sensitive photothermal detection of C60 and C70 fullerens on micro-thin-layer chromatographic plates, Analitica Chimica Acta, vol. 863 pp. 70–77, 2015