Statistical analysis of IR thermographic sequences by PCA

S. Marinetti[a], E. Grinzato[a], P.G. Bison[a], E. Bozzi[b], M. Chimenti[b], G. Pieri[b], O. Salvetti[b]

[a]*CNR- Institute of Building Technologies Section of Padova, C.so Stati Uniti 4, 35127 Padova, Italy*
[b] *CNR-Institute of Information Science and Technology, via G. Moruzzi 1, 56124 Pisa, Italy*

Abstract

Automatic processing of IR sequences is a desirable target in Thermal Non Destructive Evaluation (TNDE) of materials. Unfortunately this task is made difficult by the presence of many undesired signals that corrupt the useful information detected by the IR camera. In this paper the Principal Component Analysis (PCA) is used to process IR image sequences to extract features and reduce redundancy by projecting the original data onto a system of orthogonal components. As a thermographic sequence contains information both in space and time, the way of applying PCA to these data cannot be straightforwardly borrowed from typical applications of PCA where the information is mainly spatial (e.g. Remote Sensing, Face Recognition). This peculiarity has been analysed and the results are reported. Finally, in addition to the use of PCA as an unsupervised method, its use in a "learning and measuring" configuration is considered.

Key words: IR image sequence, Principal Component Analysis, learning and measuring, data compression, feature extraction.

1. Introduction

PCA is a quite old method, originated in 1901 by Pearson [1] and later developed by Hotelling [2]. It has found application in fields such as face recognition, remote sensing, image compression and is a common technique for summarising data of high dimension. It is a classical multivariate analysis that is useful for data compression and detection of linear relationships. It is essentially equivalent to Karhunen-Loeve transformation and closely related to factor analysis. All these methods are based on 2nd order statistics of the data. Recently PCA has been introduced in the TNDE field for discriminating optical from thermal effects in open crack detection [3]. In such an application PCA was used to qualitatively enhance the thermal signal due to the open-crack and to reduce the optical effects regarded as false alarms. A more recent application called Principal Component Thermography (PCT) [4] has been proposed as a method for defect depth characterisation. In this quantitative approach, a link between some principal components and the thermal contrast was found. This made it possible to formulate a calibration function for the defect depth estimation. These first applications of PCA to thermal data showed some interesting potentialities that it is worth investigating. The question is whether the PCA can be straightforwardly borrowed from its typical fields of application, or there is a better way to take advantage of the specific nature of the thermal signal. Indeed a set of IR images coming from a dynamic thermal test contains useful information both in space and time. Furthermore, most of the processing algorithms in TNDE are based on the analysis of the thermal contrast evolution in time, while spatial analysis is rarely used. This peculiarity makes thermal sequences different, for instance, from multispectral images used in remote sensing. This paper reports on the possible ways of applying PCA to IR sequences. Example of application to experimental data will be reported. It will be shown that the signal decomposition operated by the PCA provides good, but generally not predictable, results. This is acceptable for qualitative analysis carried out by an operator, but prevents the use of PCA in automated tasks. To overcome this problem and take a step toward quantitative applications, a learning phase is considered.

2. Basic principles of PCA

PCA is a linear projection technique for converting a matrix **A** of dimension $m \times q$ to a matrix $\mathbf{A_p}$ of lower dimension $s \times q$ ($s<m$) by projecting **A** onto a new set of principal axis. This can be done by a matrix multiplication $\mathbf{A_p}=\mathbf{U}^T\mathbf{A}$ where the columns of **U** are the projection vectors that maximise the variance retained in the projected data $\mathbf{A_p}$. This operation can be also seen as a linear transformation that minimizes the reconstruction error or a procedure to obtain uncorrelated projected distributions. Each principal axis corresponds to the normalized orthogonal eigenvector of the scatter matrix $\mathbf{S}=(\mathbf{A}-\mathbf{A}_{mean})(\mathbf{A}-\mathbf{A}_{mean})^T$ of $m \times m$ elements. One simple approach to PCA is to use singular value decomposition (SVD) of **S**:

$$(1)\quad \mathbf{S}=\mathbf{U}\mathbf{D}\mathbf{U}^T=\begin{bmatrix}\mathbf{U}_s & \mathbf{U}_n\end{bmatrix}\begin{bmatrix}\mathbf{D}_s & \\ & \mathbf{D}_n\end{bmatrix}\begin{bmatrix}\mathbf{U}_s & \mathbf{U}_n\end{bmatrix}^T$$

where **U** is the eigenvector matrix (i.e. modal matrix) and **D** is the diagonal matrix whose diagonal elements correspond to the eigenvalues of **S** (in descending order). Then the PCA transformation from m-dimensional data to s-dimensional subspace (with $s+n=m$) is given by choosing the first s column vectors. The matrix $\mathbf{A_p}$ taking into account the first s principal components is given by:

$$(2)\qquad \mathbf{A_p}=\mathbf{U}_s^{\mathrm{T}}\mathbf{A}$$

The choice of s is based on the desired amount of the variance proportion retained in the first s eigenvalues:

$$(3)\qquad r=\frac{\sum_{i=1}^{s} d_i}{\sum_{i=1}^{m} d_i}\cdot 100$$

where d_i is the i[th] element (eigenvalue) of the diagonal matrix **D**. In many cases more than 95% of variance is contained in the first 3-5 components.

3. PCA applied to IR image sequences

Let us consider a typical thermal non-destructive test: a sample is heated on one surface and the transient thermal process is observed by an IR camera. The acquired sequence of n_t IR images ($n_x \times n_y$ pixel) represents the source data volume **V** to be processed with the PCA algorithm. A pre-processing phase is needed to structure this three-dimensional data set in a suitable way for the SVD. Before doing that, it is worth explaining the meaning of the columns and rows of the above mentioned matrix **A** with $m \times q$ elements. **A** represents a set of q measurements of m-dimensional data. For the PCA to work properly the average across each of the data dimensions must be subtracted to compute the matrix **S**. Hence $\mathbf{A}_{mean}$ is a $m \times 1$ vector that is subtracted column-wise from **A**.

As the information contained in the original volume **V** is both in space (e.g. defects geometry and location) and time (thermal contrast evolution), there are two possible ways to convert **V** into the matrix **A**.

Case 1: regarding **V** as a sequence of thermograms, $\mathbf{A}_1$ has $n_x \cdot n_y$ rows (each column is an unrolled image) and n_t columns. The data dimension is $n_x \cdot n_y$ and the number of cases (or measurements) is n_t. $\mathbf{A}_{1mean}$ is the mean image. From a dimensional point of view, the principal axes are images and the projected data are temporal profiles.

Case 2: **V** is considered as a sequence of thermal contrast profiles, $\mathbf{A}_2$ has n_t rows (each column is a time profile) and $n_x \cdot n_y$ columns. The data dimension is n_t and the number of cases is $n_x \cdot n_y$. $\mathbf{A}_{2mean}$ is the mean temporal profile. Dimensionally speaking, the principal axes are temporal profiles and the projected data are images.

3.1 Computational aspects

From the computational point of view the main difference between case *1* and case *2* is the dimension of the matrix **S**. As an example, let us consider a sequence of *150* images *320 × 240* pixels each. In the first case $\mathbf{S}_1$ is a huge matrix of *76800 × 76800* elements which requires an amount of RAM hardly available in normal computers, in the second case $\mathbf{S}_2$ is a much smaller *150 × 150* matrix. In practice it is always possible to use the second approach because there is a simple relationship to recover the eigenvector matrix of the case *1* from that obtained in the case *2*:

$$(4)\qquad \mathbf{U}_1=\left(\mathbf{A}_1-\mathbf{A}_{1mean}\right)\mathbf{U}_2\mathbf{D}_2^{-\frac{1}{2}}$$

where $\mathbf{U}_2$, $\mathbf{D}_2$ are respectively the eigenvectors and eigenvalues relative to the second case. It is worth noting that eq. (4) is based on the assumption that

$$(5) \qquad (\mathbf{A}_1 - \mathbf{A}_{1mean}) = (\mathbf{A}_2 - \mathbf{A}_{2mean})^T .$$

Hence, in such a condition, apart from the dimensional exchange between projection vectors and projected data, computing the PCA on $\mathbf{A}_1$ or $\mathbf{A}_2$ does not make any essential difference. For the sake of simplicity, from here on, the components having the same dimensions as the temporal profiles will be referred to as *Temporal Components* (*TC*) independently of their nature of principal vectors or projected data. Moreover the first *TC (TC1)* will be that relative to the largest eigenvalue and the following *TC* will follow the descending order of the respective eigenvalues. Analogously, the results of PCA having the same dimensions as an image will be called *Spatial Components* (*SC*). All the following considerations will refer to the case *2* without using any subscript to denote the matrix symbols.

3.2 Mean subtraction

Differently from the transposition of the matrix **A**, the way $\mathbf{A}_{mean}$ is computed influences the PCA results. As it was mentioned before, it is possible to subtract from each image the mean image or from each temporal profile the mean profile. To better analyse these two alternatives, experimental data are processed and results discussed. A test was carried out on a 3 mm thick steel plate with six circular bottom holes (10 mm in diameter) located at different depth, to simulate material loss due to corrosion from 50% to 2% of the total thickness. The sample was heated by two flash lamps delivering an energy pulse of 4800 J in 5 ms. An IR camera FLIR® SC3000 was used to image the specimen response. A sequence of 150 images was acquired at a frequency of 50 Hz amounting to an observation interval of 3 s.

3.2.1 Mean image subtraction

The first step was the normalization of the images by the second image in order to reduce the effects of a possible uneven heating pattern or absorption distribution. The choice of the second image instead of the first one available after the flash, was made to reduce the reflection signals coming directly from the heat source. Then the volume of raw data was reduced to a two-dimensional matrix **A** by a raster-like operation, the mean image (i.e. the averaged row) was finally subtracted from each row.

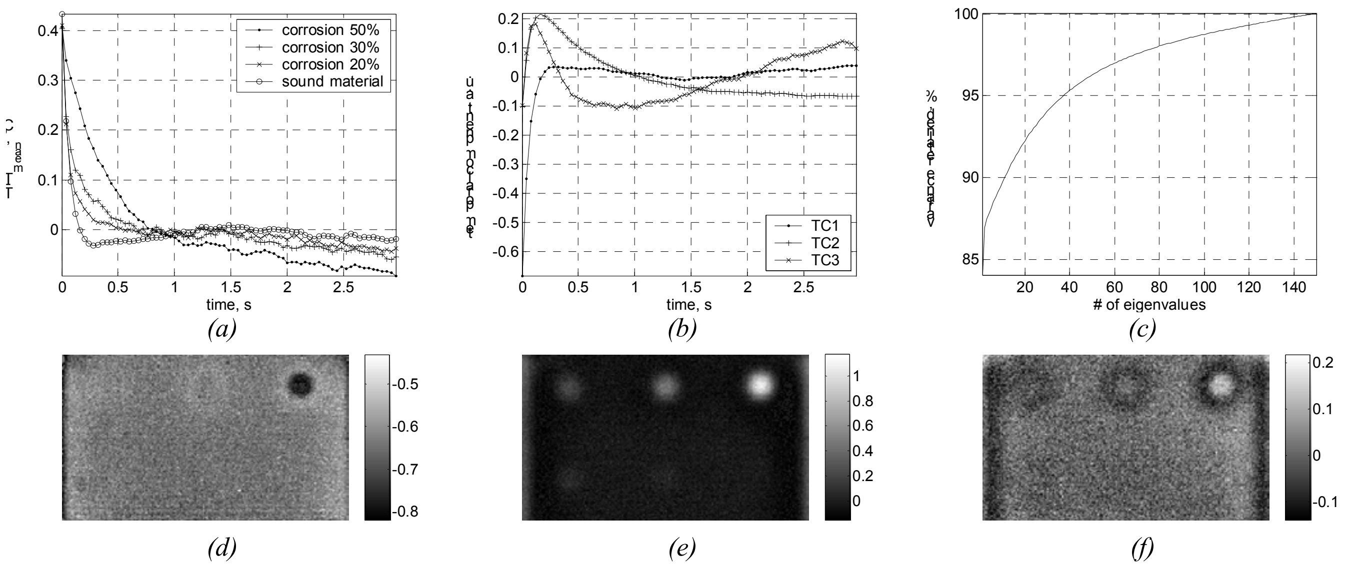


Fig. 1: temperature profiles after subtracting the mean image (a); first 3 temporal components provided by PCA (b); dimensionality curve (c); first 3 spatial components (d, e, f).

In such a way, the temporal profiles (columns of **A)** are the original temperature evolutions centred on zero (F*ig. 1a*). PCA applied to **A** yields as a first results the *TC* (F*ig. 1b*) that represent the uncorrelated decomposed profiles. In *Fig. 1c* the percentage of retained variance (eq. (3)) is plotted against the number of eigenvalues considered. It can be seen that considering 3 components the 87.5% of variance is maintained (37 eigenvalues are needed to keep 95% of variance). Moreover, *TC1* appears to be very similar to the mean temperature decay of the sample, while the following components look like thermal contrasts. *Fig. 1d, 1e* and *1f* show the *SC* relative to the first three *TC*. The *SC1* (*Fig. 1d*) is a quite uniform image, except for the shallowest defect already visible in the normalizing image. This means that the contribution of *TC1* is about the same for all the profiles. The *SC2* (*Fig. 1e*), on the contrary, exhibits low values for the background and higher values for defects (decreasing according to the severity of the material loss). Finally *SC3* still shows the marks of the three largest defects (boundary effects on the edges of the sample are visible as well).

3.2.2 Mean profile subtraction

The same procedure has been applied to the same data set but subtracting the mean temporal profile instead of the mean image. Results are reported in *Fig. 2*. *Fig. 2a* shows how the input profiles are now similar to thermal contrasts. Indeed, after normalization, as the most part of the sample is defect free, the mean profile is very close to the normalized temperature evolution over a sound area. Hence, the curves in *fig. 2a* could be regarded as normalized contrasts. With respect to the previous case, now the *SC1* is noisier (*fig. 2d*) and the three main defects are barely visible. No marked evidence of the upper-right defect is seen. While before the *TC1* was similar to the mean thermal decay, in this case it is a quite constant profile (*fig. 2b*). As for *SC2* and *SC3*, the considerations of the previous case hold. In particular the defect visibility in the *SC2* was evaluated through the SNR in both cases and the results were the same. The amount of variance retained considering three eigenvalues is now 68%. To keep the 95% of information 80 eigenvalues are needed.

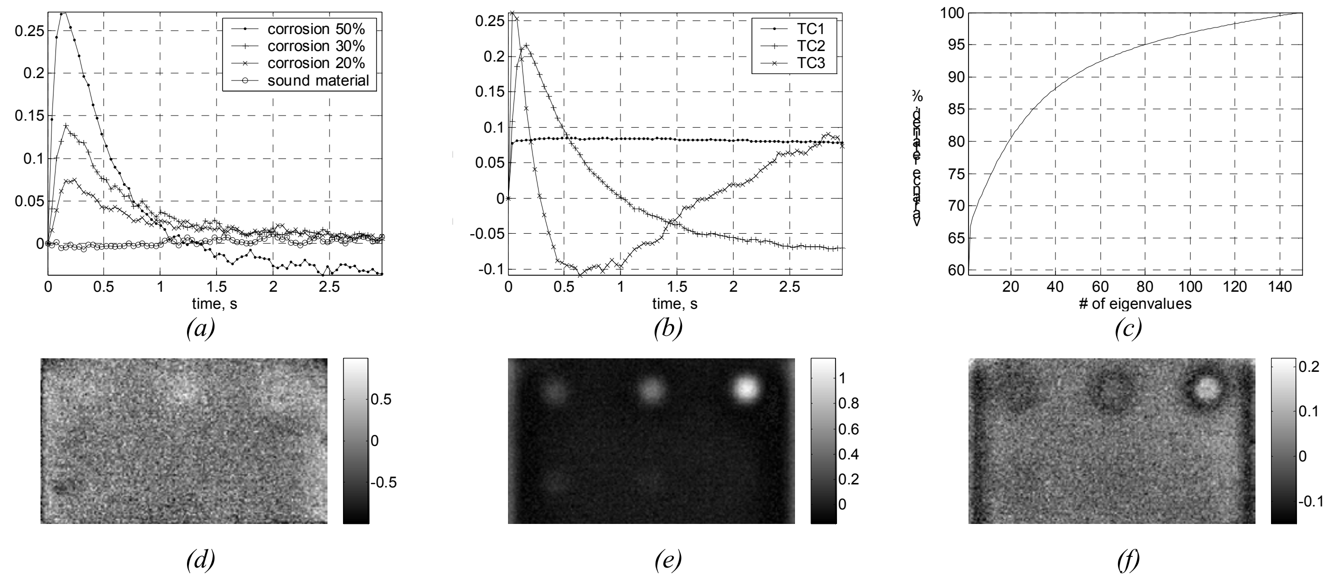


Fig. 2: temperature profiles after subtracting the mean profile (a); first 3 temporal components provided by PCA (b); dimensionality curve (c); first 3 spatial components (d, e, f).

3.3 Other examples of application

Let us consider now another example of application of PCA to a non-destructive testing of a plastic sample with 9 bottom holes at different depth. The test was carried out using another IR camera. The results are depicted in *Fig. 3* where 4 *SC* are considered. In this case the difference between the two cases is emphasized. It is worth noting that the *SC4* shows a very regular pattern maybe caused by the not properly working camera synchronization device.

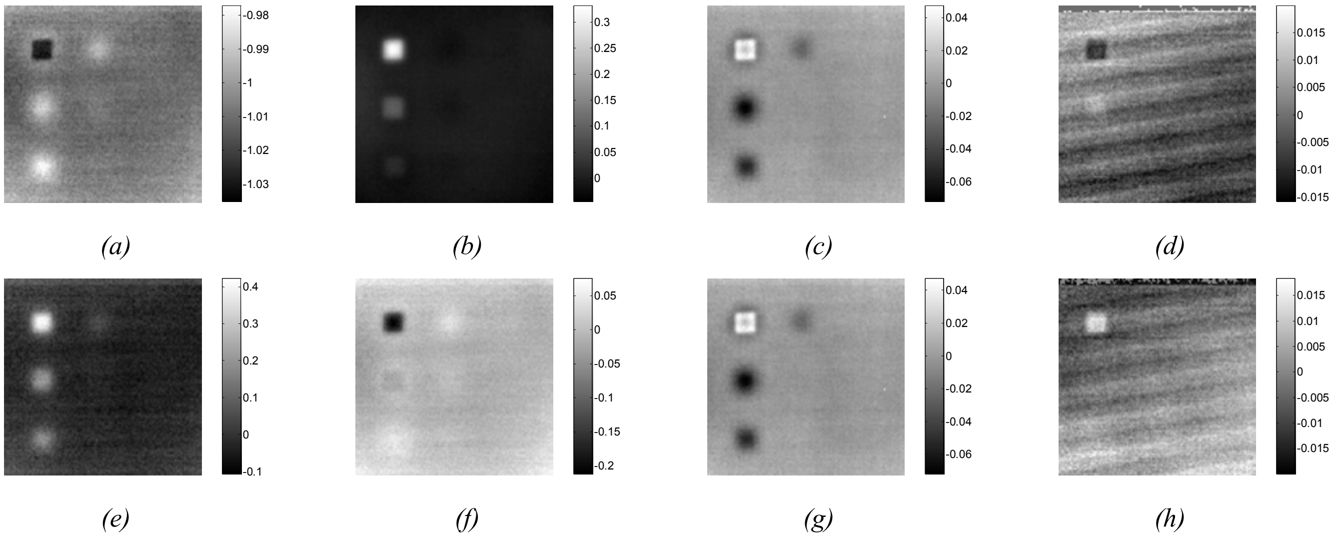


Fig. 3. NDT of a plastic sample: spatial components provided by PCA after subtracting the mean image (a, b, c, d); spatial components obtained subtracting the mean temporal profile (e, f, g, h).

So far the examples reported referred to a transient regime. In the following example, a CFRP plate with 9 Teflon® inserts was heated with a harmonic heat flux until the periodic regime was reached. Results are shown in *Fig. 4*. It can be noticed that in this case, compared to the tests in transient regime, for both the qualitative and quantitative aspects the two processing methods provide quite different results. For instance the *SC1* in the second row seems to show only the distribution of the uneven heating, while this information is totally absent in the first row.

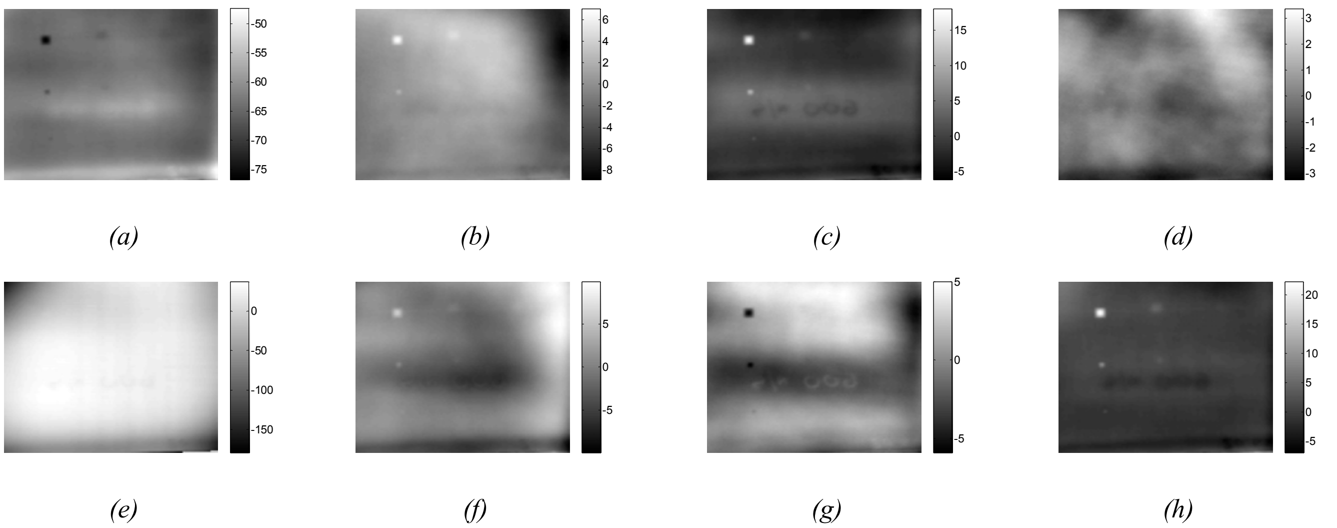
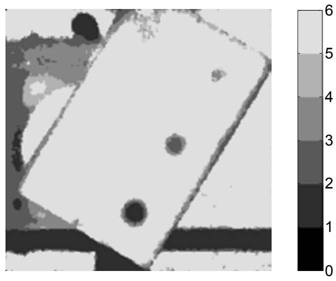


Fig. 4. NDT of a CFRP in periodic regime: spatial components provided by PCA after subtracting the mean image (a, b, c, d); spatial components obtained subtracting the mean temporal profile (e, f, g, h).

4. Learning and testing

PCA can be also used in a "*learning and testing scheme*". A training sequence is used to compute the new system of principal axes that, dimensionally, are *temporal components*. Afterwards, the training sequence is projected onto a specific subset of them so that each original profile is represented by its *n* coordinates, where *n* is the number of the projection vectors considered. These *n*-dimensional points (reference points) can be subdivided in subsets each one having a specific meaning (for instance denoting a certain defect depth). A testing sequence is then projected onto the same principal vectors used before and the distances between each projected profile and the reference points are computed. The closest reference points will determinate the class assigned to the profile under test. This procedure was applied to the steel sample described before. The second and third TC were considered (*n=2*). Classes from 1 to 5 represented the defect depths and the class n. 6 was related to the sound material. Two tests were carried out: one with the sample in a horizontal position (used for the learning phase) and one with the same sample rotated (testing phase). *Fig. 5* shows the classification results. All the four defects were assigned to the correct class. This procedure, that can be easily made automatic, is useful when several tests have to be performed on the same kind of samples. Moreover, being the testing phase based on a matrix multiplication, requires a very short computation time.

(a)



(b)

Fig. 5. Learning phase: after assigning the classes to defects depending on their severity, the training data were used for verification (a). Testing phase: result of the classification procedure applied to a testing sequence (b).

5. Conclusion

The application of PCA to TNDT has been studied. The peculiarity of this application stems from the fact that the information is both in space and time. This led to the dual interpretation of the input data volume as a set of images or a set of temporal profiles. To better understand how the results are influenced by these two ways of thinking, PCA was applied to both cases. Computational problems stemmed from the large dimensions of the scatter matrix. It has been shown that in practice this problem can be overcome thanks to a property of the eigenvectors matrix. PCA was then applied to experimental data in transient and periodic regime. It was verified that considering the initial sequence as a set of images or a set of temporal profiles influence the final results. In any case, PCA showed its ability to extract features and to condense information in a few images. Anyway, so far, no apparent connections have been found between the principal components and the physical processes involved in the test. Finally, PCA was used for learning and testing. A preliminary training stage provided the principal components used as a reference system for the classification algorithm. This procedure was applied with promising results to experimental data for material loss estimation due to corrosion.

References

[1] K. Pearson, "*On lines and planes of closest fit to systems of points in space*", The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, 2. pp. 559-572, 1901.

[2] H. Hotelling, *"Analysis of a complex of statistical variable into principal components"*, J. Educ. Psych., vol. 24, pp. 417-441, 1933.

[3] S. Hermosilla-Lara, P. Y. Joubert, D. Placko, F. Lepoutre, M. Piriou, "*Enhancement of open-cracks detection using a principal component analysis/wavelet technique in photothermal non-destructive testing*", Proceedings of QIRT Conference, Dubrovnik (Croatia), 2002 (to be published).

[4] N. Rajic, "*Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures*", Elsevier, Composite Structures 58, pp. 521-528, 2002.