

Comparison of subspace analysis methods for fault detection in industrial systems

V.-H. Nguyen, C. Rutten, J.-C. Golinval

University of Liege, Aerospace & Mechanical Engineering Department, Structural Dynamics Research Group, Chemin des chevreuils, 1 B52/3, B-4000 Liège 1, Belgium
email: VH.Nguyen@doct.ulg.ac.be

Abstract

In the field of structural health monitoring or machine condition monitoring, the activation of nonlinear dynamic behavior complicates the procedure of damage or fault detection. Principal Component Analysis (PCA) is known as an efficient method for damage diagnosis. However, two drawbacks of PCA are the assumption of the linearity of the system and the need of many sensors. This article presents industrial applications of two possible extensions of PCA: Null subspace analysis (NSA) and Kernel PCA (KPCA). The advantages of NSA rely on its rapidity of use and its reliability. The KPCA method, through the use of nonlinear kernel functions, allows to introduce nonlinear dependences between variables. The objective is to address the problem of fault detection in mechanical systems using subspace-based methods. The detection is achieved by comparing the subspace features between the reference and a current state through statistics. Industrial data are used as illustration of the methods.

Keywords: KPCA; NSA; subspace; condition monitoring; statistics.

1 Introduction

Principal Component Analysis (PCA) is a linear multivariable statistical method that can be used for fault diagnosis in mechanical systems. It is known as an efficient method to compress large sets of random variables and to extract interesting features from a dynamical system. However, this method is based on the assumption of linearity. To some extent, many systems show a certain degree of nonlinearity and/or non-stationarity, and PCA may then overlook useful information on the nonlinear behavior of the system. As reported in [1], there are many types of damage that make an initially linear structural system respond in a nonlinear manner. Therefore, detection problem may necessitate methods which are able to study nonlinear systems.

Efforts have been made to develop nonlinear damage detection methods based on PCA. For example, the nonlinear PCA method proposed in [2], [3] is able to generate nonlinear features by use of artificial neural network training procedures. In reference [4], local PCA is used to perform piecewise linearization in the cluster of nonlinear data in order to split it into several regions, and then to carry out PCA in each sub-region.

Alternatively, Kernel Principal Component Analysis (KPCA) is a nonlinear extension of PCA built to authorize features such that nonlinear dependence between variables. The method is “flexible” in the sense that different kernel functions may be used to better fit the testing data. In the beginning, KPCA has interested many scientists in the domain of image processing [5, 6]. These researchers showed that KPCA may be more advantageous than other techniques such as PCA or Wavelet Transform etc. in encoding image structure. In the last five years, KPCA has been introduced in other fields of research and has shown its ability in the monitoring of nonlinear process. J.M. Lee *et al.* [7] used KPCA to detect fault in

the biological wastewater treatment process by means of statistics charts. R. Sun *et al.* [8] achieved fault diagnosis in a large-scale rotating machine through classification techniques. A. Widodo and B.-S. Yang [9] extracted nonlinear features in support vector machines (SVM) to classify the faults of induction motor. Q. He *et al.* [10] monitored gearbox conditions by extracting the nonlinear features with low computational complexity based on subspace methods.

Another drawback of PCA is the need of several sensors. If the number of sensors is too small, modal identification and/or damage detection may not be performed in good conditions using PCA. An alternative PCA-based method named Null Subspace Analysis (NSA) was proposed to detect damages in bearings [11] and on an airplane mock-up [12]. The NSA method generates data by means of block Hankel matrices and has been proven to be efficient when the number of available sensors is small or even reduced to one sensor only [13].

This paper focus on the damage detection problem in mechanical structures and its content is as follows. First the definition of the block Hankel matrix is recalled and the PCA method is described briefly as it constitutes the background of the proposed method. Next, the kernel PCA method is introduced to deal with fault detection. The method is illustrated on two applications which consist in detecting damage in a rotating device and in controlling quality of welded joints. NSA results serve as reference for the comparison.

2 Block Hankel matrix [14]

The block Hankel matrix plays an important role in subspace system identification [14]. The covariance-driven block Hankel matrix is defined as:

$$\mathbf{X} = \begin{bmatrix} \Delta_1 & \Delta_2 & \dots & \dots & \Delta_c \\ \Delta_2 & \Delta_3 & \dots & \dots & \Delta_{c+1} \\ \dots & \dots & \dots & \dots & \dots \\ \Delta_r & \Delta_{r+1} & \dots & \dots & \Delta_{r+c+1} \end{bmatrix} \quad (c \geq r) \quad (1)$$

where r, c are user-defined parameters ($r = c$ in this paper) and Δ_i represents the output covariance matrix defined by:

$$\Delta_i = \frac{1}{N-i} \sum_{k=1}^{N-i} \mathbf{y}_{k+i} \mathbf{y}_k^T \quad (0 \leq i < N-1) \quad (2)$$

where \mathbf{y}_k is the measurement vector at time step k and N is the number of sampling points.

The Hankel matrix characterizes the dynamics of the analyzed signals and has been used for modal identification [14] and damage detection [11-13].

3 Principal Component Analysis (PCA)

Let us assume that a dynamical system is characterized by a set of vibration features collected in matrix $\mathbf{X} \in \mathbf{R}^{n \times M}$. In general, PCA involves a data normalization procedure, which leads to a set of variables with zero-mean and unitary standard deviation.

PCA, known as Karhunen-Loève transform or Proper Orthogonal Decomposition (POD), provides a linear mapping of data from the original dimension n to a lower dimension p [4] using the transformation:

$$\mathbf{Y} = \mathbf{T} \mathbf{X} \quad (3)$$

where $\mathbf{Y} \in \mathbf{R}^{p \times M}$ is called the score matrix and $\mathbf{T} \in \mathbf{R}^{p \times n}$ the loading matrix. The dimension p represents the physical order of the system or the number of principal components which affect the vibration features. The loading matrix may be found from the main p eigenvectors of the matrix \mathbf{X} . In practice, PCA is often computed by Singular Value Decomposition (SVD) of matrix \mathbf{X} , i.e.

$$\mathbf{X} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (4)$$

where \mathbf{U} is an orthonormal matrix whose columns define the principal components (PCs) and \mathbf{V} is an orthonormal matrix. The order p of the system is determined by selecting the first p singular values in $\mathbf{\Sigma}$ which have a significant magnitude ('energy') as described in [12]. A threshold in terms of cumulated energies is often fixed to select the effective number of PCs that is necessary for a good representation of matrix \mathbf{X} . In practice, a cumulated energy of 75% to 95% is generally adequate for the selection of the active PCs.

4 Kernel Principal Component Analysis (KPCA) [5, 10]

The key idea of KPCA is first to define a nonlinear map $\mathbf{x}_k \mapsto \Phi(\mathbf{x}_k)$ with $\mathbf{x}_k \in \mathbf{R}^n, (k=1, \dots, M)$ which defines a high dimensional feature space F , and then to apply PCA to the data in space F .

Suppose that the data are centered i.e. $\sum_{i=1}^M \Phi(\mathbf{x}_i) = 0$, then the covariance matrix in the space F is:

$$\mathbf{C} = \frac{1}{M} \sum_{i=1}^M \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i)^T \quad (5)$$

Eigenvalues and eigenvectors may then be extracted by solving the equation:

$$\lambda \mathbf{V} = \mathbf{C} \mathbf{V} = \frac{1}{M} \sum_{i=1}^M [(\Phi(\mathbf{x}_i)^T \mathbf{V}) \Phi(\mathbf{x}_i)] \quad (6)$$

All eigensolutions \mathbf{V} with $\lambda \neq 0$ must lie in the subspace spanned by $\Phi(\mathbf{x}_1), \dots, \Phi(\mathbf{x}_M)$ which means that coefficients $\alpha_i, (i=1, \dots, M)$ exist such that:

$$\mathbf{V} = \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i) \quad (7)$$

Pre-multiplication of equation (6) by $\Phi(\mathbf{x}_k)^T$ gives:

$$\lambda \Phi(\mathbf{x}_k)^T \mathbf{V} = \Phi(\mathbf{x}_k)^T \mathbf{C} \mathbf{V} \quad (k=1, \dots, M) \quad (8)$$

Let us define the kernel matrix \mathbf{K} of dimensions $M \times M$ such that [5]:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j) \quad (9)$$

For instance, the following kernel functions proposed by Vapnik [15] may be used:

- *polynomial kernel function*,

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + 1)^d, \quad (10)$$

where d is a positive integer

- *radial basis function (RBF)*,

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2\right) \quad (11)$$

where $2\sigma^2 = w$ is the width of the Gaussian kernel

It is worth noting that in general, the above kernel functions give similar results if appropriate parameters are chosen. The last function may present advantages because of its flexibility in the choice of the associated parameter. For example, the width of the Gaussian kernel can be very small (<1) or quite large. Contrarily the polynomial function requires a positive integer for the exponent.

The combination of equations (7) to (9) yields:

$$M \lambda \mathbf{a} = \mathbf{K} \mathbf{a} \quad (12)$$

where \mathbf{a} is the column vector collecting the coefficients $\alpha_1, \dots, \alpha_M$. KPCA requires the data points $\Phi(\mathbf{x}_i)$, ($i = 1, \dots, M$) to be centered in the feature space F , so that the centered kernel matrix is defined:

$$\tilde{K}_{ij} = (K - 1_M K - K 1_M + 1_M K 1_M)_{ij} \quad (13)$$

where 1_M is a matrix of dimensions $M \times M$ with $(1_M)_{ij} = 1/M$. Further details may be found in reference [5].

The eigenvalues associated to the eigenvectors of \mathbf{K} : $\alpha_1, \alpha_2, \dots, \alpha_M$, are ordered in descending order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \gg \lambda_{p+1} \geq \dots \geq \lambda_M \rightarrow 0$ where $\lambda_{p+1}, \dots, \lambda_M$ are negligible with respect to the first p eigenvalues (assuming that $\Phi \neq 0$). Normalization of $\alpha_1, \dots, \alpha_p$ results from the normalization of the corresponding vectors in F i.e.

$$\mathbf{V}_k^T \mathbf{V}_k = 1 \quad (k = 1, \dots, p) \quad (14)$$

According to equations (7) and (12), it follows that:

$$\lambda_k (\alpha_k^T \alpha_k) = 1 \quad (15)$$

The eigenvectors identified in the feature space F can be considered as kernel principal components (KPCs) which characterize the dynamical system in each working state.

Note that, since the number of eigenvectors (i.e. nonlinear PCs) is the same as the number of samples, it is higher than the number of (linear) PCs given by PCA. The KPCA method is termed “nonlinear” since the feature mapping in the space F is achieved by a nonlinear function (Eqs. 10, 11). Thanks to that property, extracted KPCs should be able to reflect nonlinear or high order features which allow the representation and classification of varied states. It is also interesting to note that the method is able to characterize a dynamic behavior even if the number of measurements is small (for example, if only one sensor is available).

The kernel principal component representation \mathbf{u}_k may be obtained by projecting all the observations on the direction of the k th eigenvector:

$$\mathbf{u}_k = \mathbf{V}_k \tilde{\Phi}(\mathbf{x}_i) = \sum_{v=1}^M \alpha_{k,v} \tilde{K}(\mathbf{x}_v, \mathbf{x}_i) \quad (i = 1, \dots, M) \quad (16)$$

where \mathbf{V}_k is the k th eigenvector of equation (6), \mathbf{x}_i indicates a sample feature vector in the data matrix.

Let us consider now a set of current test data $\mathbf{t}_1, \dots, \mathbf{t}_L$. The kernel matrix $L \times M$ and its centering are defined by $K^{test}(\mathbf{t}_i, \mathbf{x}_j) = (\Phi(\mathbf{t}_i) \cdot \Phi(\mathbf{x}_j))$ and:

$$\tilde{\mathbf{K}}^{test} = \mathbf{K}^{test} - 1'_M \mathbf{K} - \mathbf{K}^{test} 1_M + 1'_M \mathbf{K} 1_M,$$

where $\mathbf{1}_M^1$ is a $L \times M$ matrix of which all entries are $1/M$. With a set of data \mathbf{t}_r , one can extract a nonlinear component by:

$$\mathbf{u}_k = \mathbf{V}_k \tilde{\Phi}(\mathbf{t}_r) = \sum_{v=1}^M \alpha_{k,v} \tilde{K}^{test}(\mathbf{x}_v, \mathbf{t}_r) \quad (r=1, \dots, L) \quad (17)$$

5 KPCA monitoring using statistics

A popular method for dynamical monitoring is the chart of Hotelling's T^2 statistics and the Q -statistics, also known as the squared prediction error (SPE) that is depicted in Figure 1. The definitions and formulation presented here follow closely the ones used by Lee *et al.* in reference [7].

T^2 measures the distance in the principal subspace (spanned by the first p PCs), whereas SPE characterizes the distance in the residual subspace (spanned by the remaining PCs, which is orthogonal to the principal subspace). In general, the use of T^2 and SPE distances in detection methods based on linear PCs (PCA method) is expressed in terms of *Mahalanobis* and *Euclidian distances* [4], respectively. Those indexes are built in the feature subspace in order to exploit the results of KPCA.

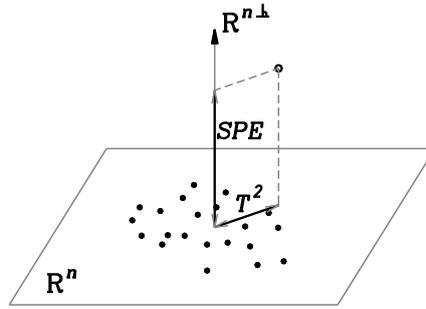


Figure 1: Graphical description of two monitoring statistics T2 and SPE

T^2 measures the variation within the model subspace, which is the sum of the normalized squared scores defined as:

$$T^2 = [\mathbf{u}_1, \dots, \mathbf{u}_p] \mathbf{\Lambda}^{-1} [\mathbf{u}_1, \dots, \mathbf{u}_p]^T \quad (18)$$

where \mathbf{u}_k is obtained by (16) or (17) and $\mathbf{\Lambda}^{-1}$ is the diagonal matrix of the inverse of the eigenvalues associated with the retained PCs. Assuming that the scores have a multivariate normal distribution, the confidence limit for T^2 is calculated using the Snedecor's F distribution (F -distribution) as:

$$T_{p,M,\alpha}^2 \approx \frac{p(M-1)}{M-p} F_{p,M-p,\alpha} \quad (19)$$

where p is the number of PCs. $F_{p,M-p,\alpha}$ is a F -distribution with degrees of freedom p and $M-p$ with level of significance α .

In the KPCA method, the observations are not analyzed in the input space, but in the feature space F . The measure of the quality of fit of a sample to the KPCA model was suggested by Lee *et al.* [7] through a simple calculation of SPE in the feature space F . The conceptual framework of the KPCA method is shown schematically in Figure 2. First, an input vector \mathbf{x} is projected using a nonlinear mapping Φ in a high-dimensional feature space F . Then linear PCA is performed in this feature space and gives score values \mathbf{u}_k in a lower p -dimensional KPCA space (eq. (16) or (17)). A feature vector $\Phi(\mathbf{x})$ may be

reconstructed from \mathbf{u}_k by projecting \mathbf{u}_k into the feature space via \mathbf{V}_k and it results in a reconstruction with p PCs in the feature space: $\hat{\Phi}_p(\mathbf{x}) = \sum_{k=1}^p \mathbf{u}_k \mathbf{V}_k$.

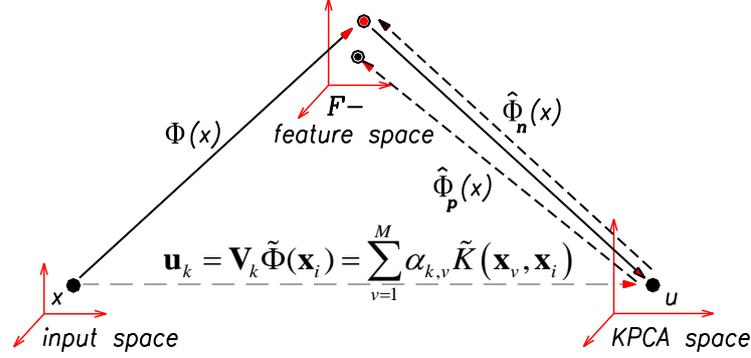


Figure 2: Conceptual diagram of KPCA

The SPE statistic in the feature space is defined by: $\text{SPE} = \|\Phi(\mathbf{x}) - \hat{\Phi}_p(\mathbf{x})\|^2$. $\Phi(\mathbf{x})$ is just identical to $\hat{\Phi}_n(\mathbf{x}) = \sum_{k=1}^n \mathbf{u}_k \mathbf{V}_k$ where n is the number of nonzero eigenvalues generated from (12) among all the M eigenvalues. The SPE can be deduced from the expression:

$$\begin{aligned}
 \text{SPE} &= \|\Phi(\mathbf{x}) - \hat{\Phi}_p(\mathbf{x})\|^2 = \|\hat{\Phi}_n(\mathbf{x}) - \hat{\Phi}_p(\mathbf{x})\|^2 = \hat{\Phi}_n(\mathbf{x})^T \hat{\Phi}_n(\mathbf{x}) - 2\hat{\Phi}_n(\mathbf{x})^T \hat{\Phi}_p(\mathbf{x}) + \hat{\Phi}_p(\mathbf{x})^T \hat{\Phi}_p(\mathbf{x}) \\
 &= \sum_{k=1}^n \mathbf{u}_k \mathbf{V}_k^T \sum_{i=1}^n \mathbf{u}_i \mathbf{V}_i - 2 \sum_{k=1}^n \mathbf{u}_k \mathbf{V}_k^T \sum_{i=1}^p \mathbf{u}_i \mathbf{V}_i + \sum_{k=1}^p \mathbf{u}_k \mathbf{V}_k^T \sum_{i=1}^p \mathbf{u}_i \mathbf{V}_i \\
 &= \sum_{k=1}^n \mathbf{u}_k^2 - 2 \sum_{k=1}^p \mathbf{u}_k^2 + \sum_{k=1}^p \mathbf{u}_k^2 = \sum_{k=1}^n \mathbf{u}_k^2 - \sum_{k=1}^p \mathbf{u}_k^2
 \end{aligned} \tag{20}$$

as $\mathbf{V}_k^T \mathbf{V}_i = 1$ when $k=i$, otherwise $\mathbf{V}_k^T \mathbf{V}_i = 0$.

Assuming that the prediction errors are normally distributed, the $100(1-\alpha)\%$ confidence limit for the SPE is calculated by fitting a weighted χ^2 -distribution and is given by:

$$\text{SPE}_\alpha \approx \frac{v}{2m} \chi_{2m^2/v, \alpha}^2 \tag{21}$$

where m and v are the estimated mean and variance of the SPE, respectively [16]. This approximating distribution works well even in cases for which the errors do not follow a Gaussian distribution.

Given the T^2 and SPE statistics as well as their control limits, we can compute outlier statistics to count how many prediction errors (represented in percentage) overpass the control limits for each state. In normal condition, i.e. without damage, the outlier statistics of the current data should be similar to the statistics of the reference state. Conversely, the feature subspace associated with a damage state should present significant changes from the reference state. Thus, the outlier statistics in this case should increase clearly. In order to compare the statistics related to two different sets of data, the ratio S_d/S_r is also computed, where d and r denote the damage and the reference states, respectively, S describes the average SPE or T^2 in according statistic measurements. While $S_d/S_r \rightarrow 1$ indicates a normal state, a higher value of this ratio can indicate the occurrence of a possible damage.

6 Applications

In the following, the covariance-driven Hankel matrix in Eq. (1) is considered as the input data matrix for the KPCA procedure. Here, the number of block rows r and the number of block columns c are equal, hence \mathbf{X} is a square matrix.

6.1 Damage detection in electro-mechanical devices

This industrial application concerns the case of electro-mechanical devices for which the overall quality at the end of the assembly line has to be assessed.

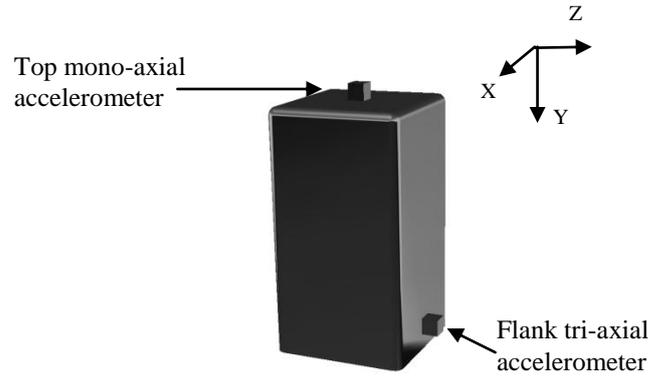


Figure 3: Location of the accelerometers on the electro-mechanical device: one mono-axial accelerometer on the top and one tri-axial accelerometer on the flank of the device

A set of five good (healthy) devices and four damaged devices was considered. Dynamic responses were collected by one mono-axial accelerometer on the top and one tri-axial accelerometer on the flank of the device as illustrated in Figure 3. Only data measured in one direction on the flank (X, Y, Z in Figure 3) or on the top of the device is used for the detection.

It is worth recalling that with only one sensor response, detection cannot be performed by a subspace method like PCA. For this reason, fault detection was realized by Rutten *et al.* [13] through Null subspace analysis (NSA) [12], using the Novelty Index [4] based on the Mahalanobis norm. The detection in [13] also consists in computing outlier statistics where a confidence limit for NI was admitted: $NI_{lim} = \overline{NI} + 6\sigma$, \overline{NI} and σ are the mean value and standard deviation of NI for the reference test. As it was found that detection is better when using the data in the Y direction, the data in this direction is exploited here to test the KPCA-based detection method. In order to increase the sensitivity to damage, 35 block rows are used to construct the Hankel matrix. The radial basis function (RBF) is used with the width of Gaussian kernel $w = 3$. The eigenvalue diagram for the kernel matrix \mathbf{K} in Figure 4 helps to choose an appropriate number of PCs which defines the optimal dimension of the KPCA space. This dimension was finally chosen equal to 3 as the first three eigenvalues present about 90% of variance percentage and the next value shows a clear decrease.

First, a healthy device is examined and the results are shown in Figure 5. The monitoring diagram is split into two parts: the left part trains the reference data and the right part analyzes the current data. The 99% confidence limits are shown by dashed lines. As expected, Figure 5 does not reveal any fault because the outlier statistics of the current data are not relevant and the ratios S_d/S_r remain close to unity for both SPE and T^2 . The confidence limit for T^2 is much bigger than all T^2 values and not shown in the figure.

In Figure 6, one damaged device is considered. In this case, the SPE statistic shows 100% outlier; the ratios S_d/S_r reach 13.75 for the SPE and 2.49 for the T^2 index.

Figure 7 reports the monitoring of the set of eight rotating devices which are compared to the healthy device considered as reference. Among the eight devices, four of them may be considered as good (Ok 1-Ok 4) and four of them as damaged (NOk 1-NOk 4). Figures 7a and b show that both NSA and KPCA are able to detect accurately the four damaged devices (NOk). Among the four NOk-cases, three of them are diagnosed with 100% of SPE monitoring. The SPE ratios S_d/S_r in Figure 7c gives also a clear distinction between healthy and damaged devices. However, the T^2 norm does not provide useful information in this problem (Figure 7d).

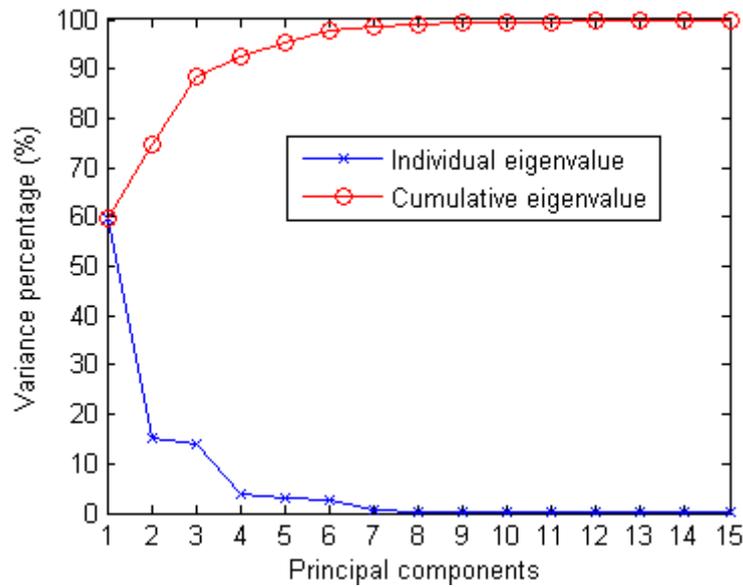


Figure 4: Eigenvalues of the kernel matrix vs. the number of PCs

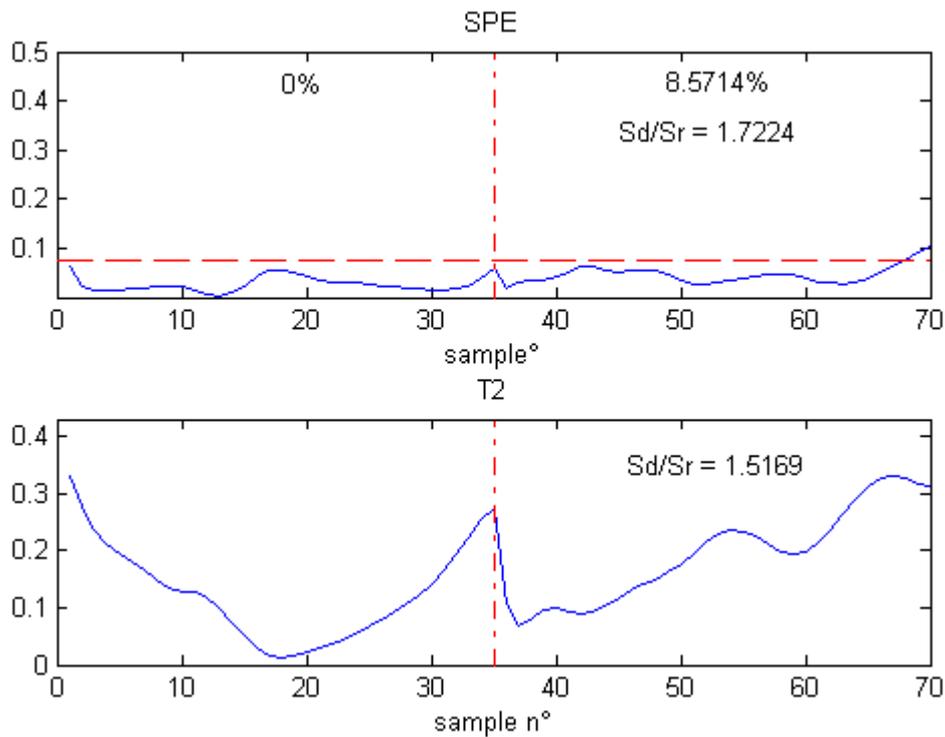


Figure 5: SPE and T^2 statistics - no damage

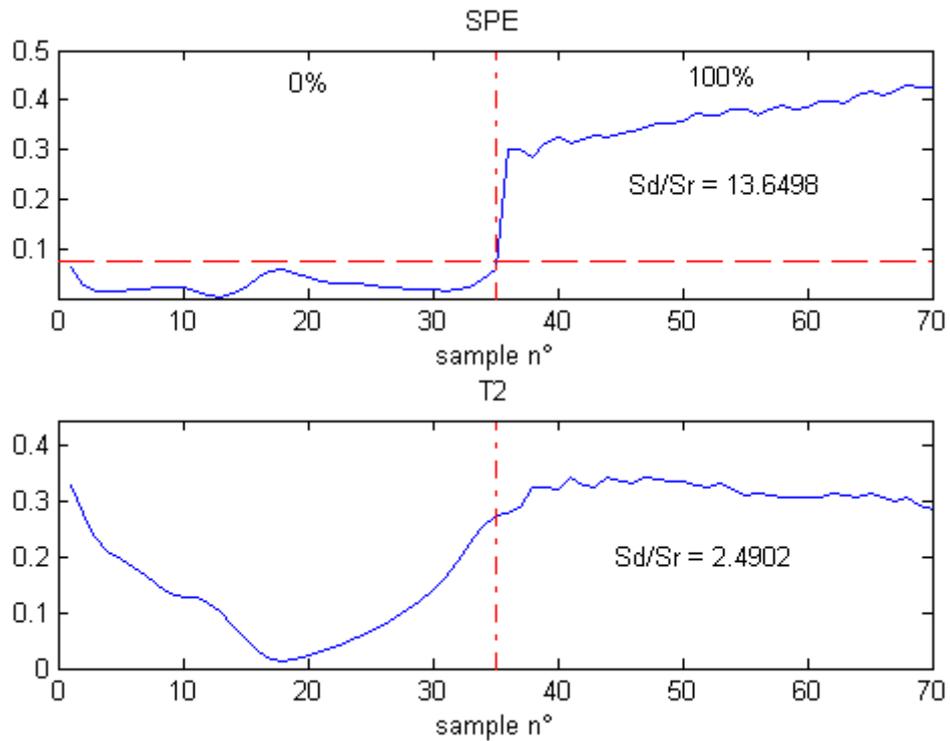


Figure 6: SPE and T^2 statistics - with damage

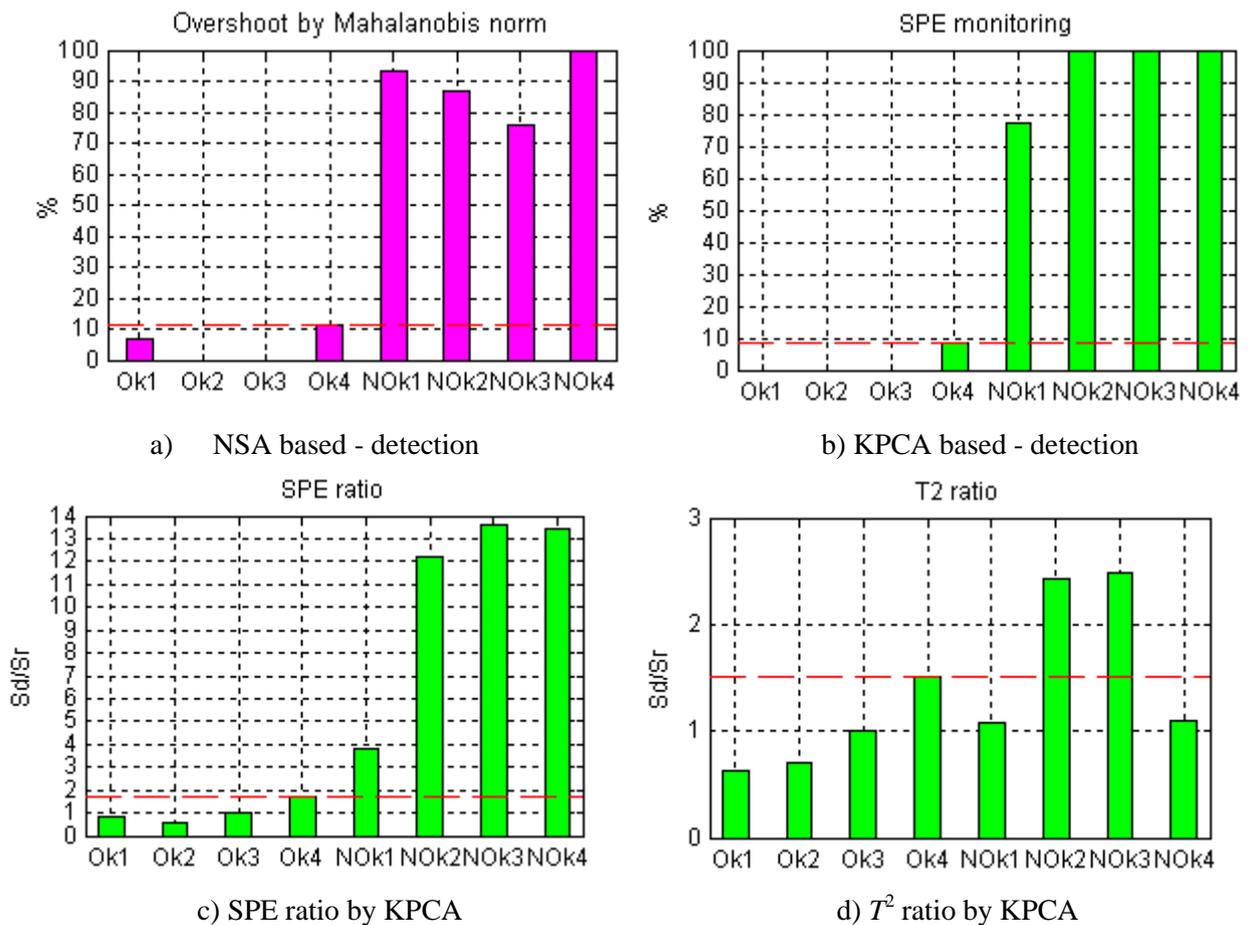


Figure 7: Damage detection by NSA and KPCA – based methods (dashed lines present the maximal values for good devices)

6.2 Quality control of welded joints

The second example involves an industrial welding machine from a steel processing plant. The machine was instrumented with a mono-axial accelerometer on the forging wheel, as illustrated in Figure 8. The purpose of this wheel is to flatten the welded joint.

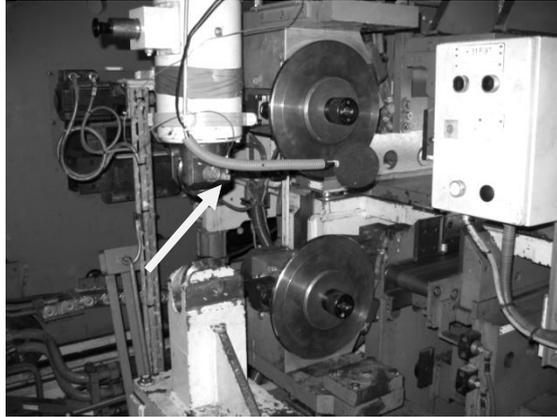


Figure 8: Location of the accelerometer on the forging wheel of the welding machine

The quality of the weld depends on several parameters. In [13], six welded joints with nominal welding parameters (OK 1- OK 6) and 27 joints with out-of-range parameters were studied. Some welding parameters were altered, namely covering, compensation, current and forging pressure, as reported in Table 1. From a microscope quality control, welded joints C and G were diagnosed good, welded joints A, D, E, H were diagnosed acceptable and welded joints named B, F, I were diagnosed bad.

Name	Modified parameter	Weld quality
Welding A	-33% covering	Acceptable
Welding B	-66% covering	Bad
Welding C	-33% compensation	Good
Welding D	-66% compensation	Acceptable
Welding E	-10% current	Acceptable
Welding F	-20% current	Bad
Welding G	-10% forging pressure	Good
Welding H	+5% forging pressure	Acceptable
Welding I	-66% covering and compensation	Bad

Table 1: Welds realized with altered parameters (with respect to the nominal parameters).

The parameter modifications were clearly identified using the NSA-based detection method as presented in Figure 9. It is observed that all OK welded joints do not present non-zero outlier statistics, and so do welded joints C and G.

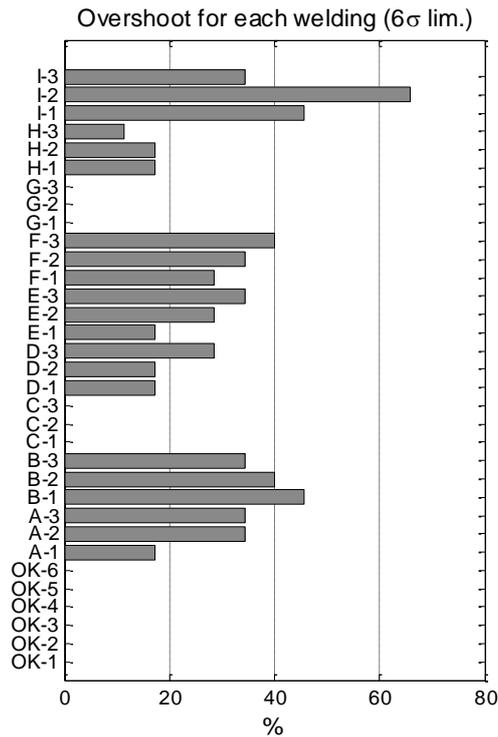


Figure 9: NSA-based detection (from [13] , $NI_{lim} = \overline{NI} + 6\sigma$)

In this example, the welded joints are examined by means of the KPCA method. From only one response vector, the Hankel matrix is built using 30 block rows. The width of Gaussian kernel $w = 150$ was chosen for the RBF kernel. The eigenvalue diagram is presented in Figure 10. It reveals that the first three PCs concentrate 70% of the accumulated variance percentage. It is interesting to note that, as long as the number of principal components is higher than 3, detection results lead to the same conclusion. The best results are obtained within the range of 5-7 PCs which accumulate 80-90% of the variance percentage.

The SPE-based indexes given in Figure 11 allow to identify the alterations reported in Table 1. The indexes of welded joints C1-C3, G1, G3 are in the same range as for healthy welded joints OK1-OK6 (i.e. with nominal parameters). This is in good agreement with microscopic quality control inspections. Weld G2 gives an index close to with the one of welds A, D, E, H which are diagnosed as acceptable. Finally, welds B, F and I are characterized by significant overshoots and high ratios S_d/S_r . Their diagnosis as bad joints is corroborated by the quality control inspection. It should be noted also that index T2 does not lead to significant detection results in this problem. Furthermore, the detection indicators used in the KPCA method as well as in the NSA method are not proportional to the severity of the defects.

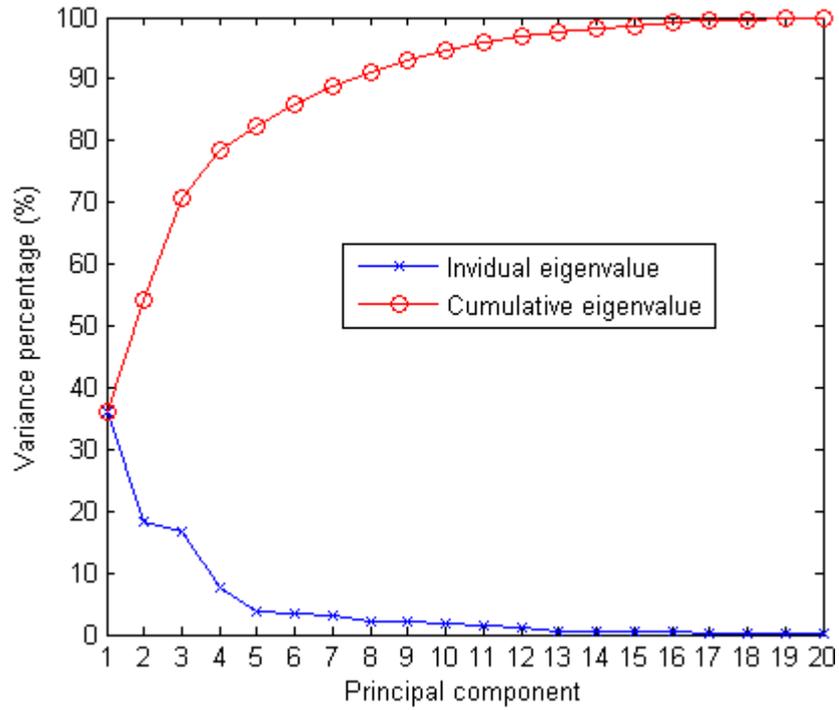


Figure 10: Eigenvalues of the kernel matrix vs. the number of PCs

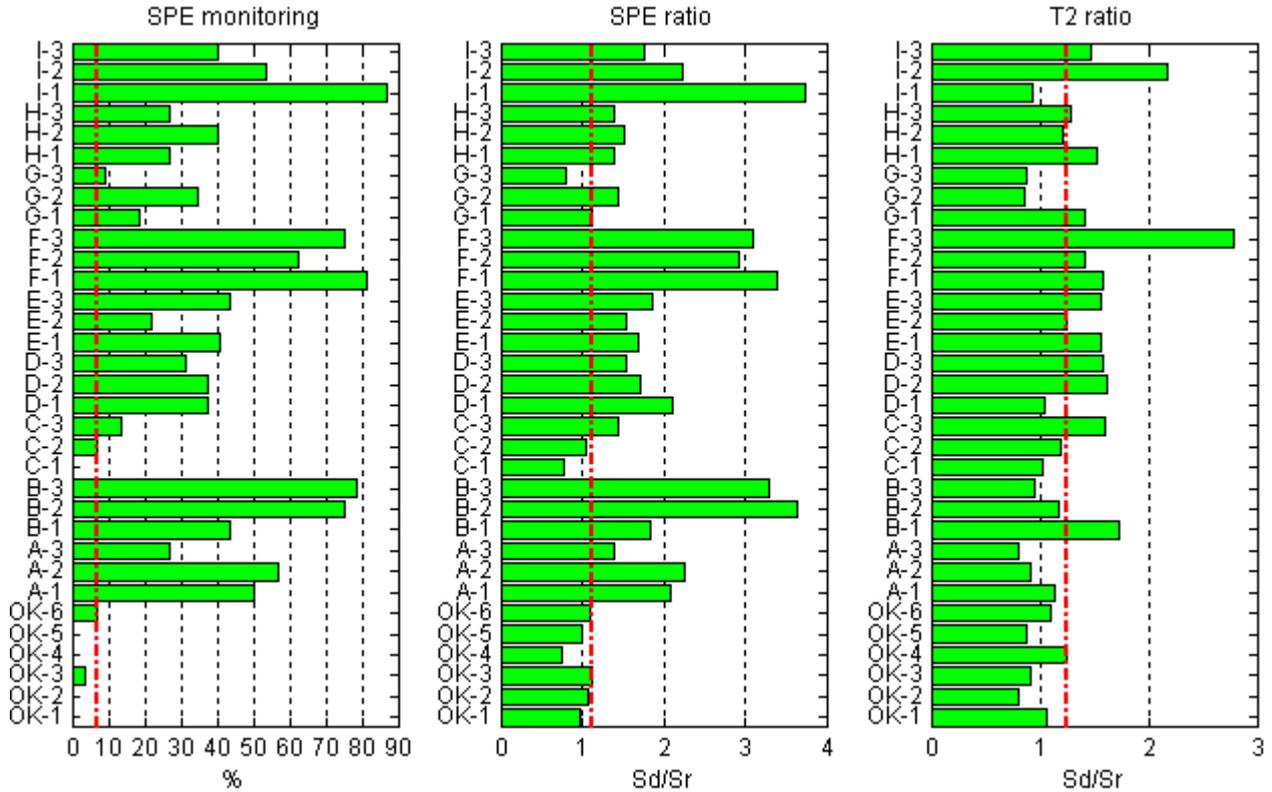


Figure 11: KPCA-based detection (dash-dot lines present the maximal values for healthy welded joints)

7 Conclusion

In this paper, two subspace analysis methods for fault detection in industrial systems were compared. Both the NSA-based and the KPCA-based methods are extension of the PCA-based method but they differ from it by making use of the Hankel matrix. In the case of KPCA, the data are processed in a nonlinear way. The efficiency of NSA and KPCA in damage detection problems has been illustrated using experimental data from two industrial applications. The advantage of those methods is that they are able to work when the set of sensors is limited and even reduced to one single sensor.

For both methods, the choice of the number of blocks in the Hankel matrix and the selection of the order (number of active principal components - PCs) are important. In this paper, the number of active PCs relies on their variance percentage which is representative of the system dynamics. An alternative method for the NSA method [13] is to select the order on the basis of a mapping of the space [number of blocks in the Hankel matrix, number of PCs].

References

- [1] Charles R. Farrar, Keith Worden, Micheal D. Todd, Gyuhae Park, Jonathon Nochols, Douglas E. Adams, Matthew T. Bement, Kevin Farinholt, 2007, *Nonlinear System Identification for Damage Detection*, Los Alamos National Laboratory 11(2007).
- [2] M.A. Kramer, *Nonlinear principal component analysis using autoassociative neural networks*, A.I.Ch.E. Journal 37 (2), (1999), pp. 233-243.
- [3] H. Sohn, K. Worden, C.F Farrar, *Novelty detection under changing environmental conditions*, SPIE's Eight Annual International Symposium on Smart Structures and Materials (2001), Newport Beach, CA (LA-UR-01-1894).
- [4] Aimin Yan, Gaëtan Kerschen, Pascal De Boe, Jean-Claude Golinval, *Structural damage diagnosis under changing environmental conditions – Part II: local PCA for nonlinear analysis*, Mechanical System and Signal Processing 19(4) (2005), pp. 865-880.
- [5] Bernhard Schölkopf, Alexander Smola, Klaus-Robert Müller, *Nonlinear component analysis as a kernel eigenvalue problem*, Neural Computation 10 (1998), pp.1299-1319.
- [6] In Kim Kwang, O. Franz Matthias and Bernhard Schölkopf, *Iterative Kernel Principal Component Analysis for Image Modeling*, IEEE Transactions on Pattern Analysis and Machine Intelligence 27(9) (09-2005), pp. 1351-1366.
- [7] J.M. Lee et al., *Nonlinear process monitoring using kernel principal component analysis*, Chemical Engineering Sciences 59, 2004, pp. 223-234.
- [8] Ruixiang Sun, Fugee Tsung, Liangsheng Qu, *Evolving kernel principal component analysis for fault diagnosis*, Computers & Industrial Engineering 53 (2007), pp. 361-371.
- [9] Achmad Widodo, Bo-Suk Yang, *Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motor*, Expert Systems with Applications 33 (2007), pp. 241-250.
- [10] Qingbo He, Fanrang Kong, Ruqiang Yan, *Subspace-based gearbox condition monitoring by kernel principal component analysis*, Mechanical Systems and Signal Processing 21 (2007), pp. 1755-1772.
- [11] C. Thiry, A.-M. Yan, J.-C. Golinval, *Damage detection in rotating machinery using statistical methods: PCA analysis and autocorrelation matrix*, Surveillance 5 CETIM Senlis 11-13 October, (2004).
- [12] A-M. Yan, J.C. Golinval, *Null subspace based damage detection of structures using vibration measurements*. Mechanical Systems and Signal Processing 20 (2006), pp. 611-626.

- [13] C. Rutten, C. Loffet, J.C. Golinval, *Damage detection of mechanical components using null subspace analysis*, 2th International Symposium ETE'2009 – Brussels, Belgium.
- [14] Peter van Overschee, Bart De Moor, *Subspace identification for linear systems – Theory – Implementation – Applications*, Kluwer Academic Publishers (1997).
- [15] Vapnik V., *The nature of statistical learning theory*, New York: Springer Verlag (1995).
- [16] Nomikos P., MacGregor J.F., *Multivariate SPC charts for monitoring batch processes*. *Technometrics* 37 (1995), pp. 41-59.