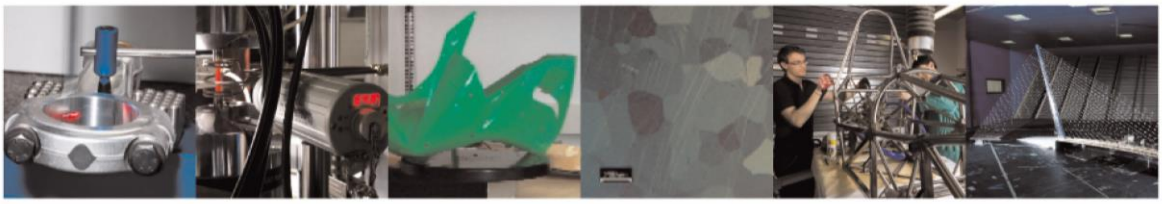

Multilinear principal component analysis for statistical modeling of cylindrical surfaces: a case study

Pacella, Massimo; Colosimo, BIANCA MARIA

This is an Accepted Manuscript of an article published by Taylor & Francis in QUALITY TECHNOLOGY & QUANTITATIVE MANAGEMENT on 06 Sep 2016, available online: https://doi.org/10.1080/16843703.2016.1226710

Revised Article

Multilinear Principal Component Analysis for Statistical Modeling of Cylindrical Surfaces: a Case Study

Massimo Pacella[1], Bianca M. Colosimo[2]

[1]Universita' del Salento, Piazza Tancredi 7, 73100 Lecce, Italy

[2]Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

March 2016

Abstract

This paper focuses the problem of modeling manufactured surfaces for statistical process control. The application of Multilinear Principal Component Analysis (MPCA) is introduced. MPCA is the generalization of the regular Principal Component Analysis (PCA) where the input can be not only vectors, but also tensors. The objective of this work is basically to explore the MPCA, as well as some basic concepts of multilinear algebra, for modeling manufactured surfaces. A real case study concerning cylindrical surfaces obtained by a lathe-turning process is taken as reference. The measurements related to a specific surface are stored in a matrix addressed by 2 index variables, while the observed data set related to several surfaces is stored in a 3rd-order tensor addressed by 3 indexes. Since the targeted application involves only the use of 3rd-order tensors of real entries, in this study the implementation of MPCA is limited to this specific case. Although a specific geometry is used herein as reference case study, any 2.5-dimensional surface (i.e., where scalar measurements are sampled for each item by using a fixed grid of two spatial index variables) can be modeled with the proposed MPCA-based approach.

Key Words: Orthogonal tensor decomposition, Multilinear algebra, Dimensionality reduction, Feature extraction, Turning process' parameters

Dr. Pacella is Assistant Professor in the Department of "Ingegneria dell'Innovazione" at the Universita' del Salento, Italy. His email address is `massimo.pacella@unisalento.it`

Dr. Colosimo is Full Professor in the Department of Mechanical Engineering at the Politecnico di Milano, Italy. She is a Senior Member of ASQ. Her email address is `biancamaria.colosimo@polimi.it`

Introduction

In modern statistical process control, an increasing attention is being devoted to *profile monitoring*, where the quality characteristic of interest can be represented as a function of one or more explanatory variables.

Starting from the simplest case of linear profiles, the literature on profile monitoring has now covered different parametric and nonparametric approaches, as well as several application domains (an extensive overview of profile monitoring can be found in the book of (Noorossana et al., 2011)). Among the different application domains, one of growing interest in the manufacturing field concerns geometric tolerances (e.g., roundness or circularity, straightness, cylindricity, flatness or planarity), for which the quality characteristic of interest consists of a profile or surface describing the shape of the manufactured item (Colosimo and Pacella, 2007; Colosimo et al., 2008; Colosimo and Pacella, 2010; Colosimo and Pacella, 2011; Colosimo and Senin, 2011; Colosimo et al., 2014).

Although most of the profile monitoring researches dealt with curves modeled as a scalar function of one single location variable (i.e., $y = f(x)$), when the quality is related to the shape of manufactured item constrained by a geometric tolerance, the techniques and methods presented in the existing literature cannot be simply applied. Recently, two cases were highlighted in the literature, with reference to geometric specification often used in industry (Colosimo and Pacella, 2011; Colosimo et al., 2010; Colosimo et al., 2014). In (Colosimo and Pacella, 2011), the straightness of cylinder axes was considered. According to the ISO standard (ISO, 2012), a spatial straightness error must be considered and axis must be modeled as a curve not lying in a plane. To this aim, performances of different modeling approaches, all of them based on the Principal Component Analysis (PCA), were compared by the authors. In (Colosimo et al., 2010; Colosimo et al., 2014), cylindricity of lathe-turned items was considered. According to the standards (ISO, 2012), the cylindricity tolerance refers to the deviation of the actual shape from an ideal cylindrical surface. In (Colosimo et al., 2010), a method was presented, which combined linear regression model and spatial autoregressive errors with the objective to represent the spatial correlation characterizing adjacent points on each cylindrical surface. Control charts were implemented in order to monitor model coefficients to detect changes of the surface pattern. In (Colosimo et al., 2014), a Gaussian Process (GP) was introduced for cylindrical surface modeling. The use of GPs instead of regression allows one to simplify the modeling step because selection of the appropriate regressor variables is not required. In order to detect change of the surface pattern, control charts were implemented to monitor the GP-predicted deviations of the surface from the in-control pattern.

The present research goes in the same direction of the cited ones, as it faces the problem of modeling a 2.5-dimensional (2.5D) surface, where measurements z are sampled for each item using a fixed grid of two spatial index variables, i.e., $z = f(x, y)$, where (x, y) represents the location point where z is measured. To this aim a PCA-based technique is explored.

It is known that regular PCA can be used as an effective tool for dimension reduction and variation characterization of one-dimensional arrays of data, e.g., curves lying in a plane and sampled on a fixed grid of spatial locations (Colosimo and Pacella, 2007). However, regular PCA cannot be directly applied to 2.5D surfaces since measurements in each sample are arranged as two-dimensional arrays of data (matrices). One approach to overcome this limitation is to reshape measurements in each sample to a vector and then apply regular PCA. However, this approach breaks the correlation structure in the original data, and loses potentially more compact or useful representations that can be obtained in the original form.

The Multilinear PCA (MPCA) (Lu et al., 2008) is a generalization of the regular PCA where the input can be not only vectors, but also tensors. A tensor is defined as a multidimensional array of data. For example, a data set of 2.5D surfaces, which share the same grid of spatial location variables, be viewed as an 3rd-order tensor. Tensor-based PCA methods, such as MPCA, directly analyse a tensor without reshaping it into a vector, thus preserving the tensor structure of the original data set. Recently, tensor-based PCA methods were used in manufacturing both for statistical process control and fault diagnosis (Paynabar et al., 2013; Yan et al., 2015). This paper aims to explore the MPCA in order to define an approach for modeling 2.5D surfaces for statistical process control in manufacturing. The case study of lathe-turned items presented in (Colosimo and Pacella, 2011) is considered herein. For each lathe-turned item, a 2.5D geometric feature is taken into account, namely the cylindrical surface of each item (subjected to cylindricity tolerance).

The remainder of the paper is organized as follows. In the next section, the real case study of cylindrical surfaces is presented. Then, the approach for multilinear truncation and best approximation for dimensionality reduction, which follows the classical PCA paradigm, are briefly introduced. Application of the MPCA to the reference case study is discussed in the subsequent section. Conclusions and final remarks will be provided in the last section.

Data collection

The case study consists of a set of 90 samples, where each sample is a bar of titanium alloy Ti-6Al-4V supplied in 20 mm diameter and machined using a lathe-turning process to a final diameter of 16.8 mm by implementing two cutting steps. With reference to the second cutting step (finishing operation), three levels of two process parameters (the cutting speed and the depth of cut) were considered, according to a 3^2 full design. Table (1) shows the 9 treatments of the experiment, where each treatment was randomly replicated 10 times. Each

Treatment #	**Sample #**	**depth** (mm)	**speed** (m/min)
1	1-10	0.4	80
2	11-20	0.4	70
3	21-30	0.4	65
4	31-40	0.8	80
5	41-50	0.8	70
6	51-60	0.8	65
7	61-70	1.2	80
8	71-80	1.2	70
9	81-90	1.2	65

Table 1: Cutting parameters for 9 experimental conditions.

surface was measured by a Coordinate Measuring Machine - CMM (Dowling et al., 1997) on a given cylindrical grid of $P \times Q$ equally spaced positions: P measurements along the vertical direction and Q along the angular direction. The CMM touch trigger approached the nominal position of each grid point, and then stored the coordinates of the probe when it touched the surface. The measurements were taken in 42 mm along the bar length and a set of 210×64 radii was available for each item. Surface data were computed as deviations of the measured radii from the substitute geometry: i.e., the ideal cylinder. For each surface, the substitute geometry was computed by minimizing the sum-of-squared distances between the observed radii and the ideal feature. Hence, the final set of data consists of radial deviations (either positive or negative) from a perfect cylinder, measured at each position of the fixed grid. A registration procedure was also implemented on the surface data set, as required in industrial engineering when dealing with form tolerances (Colosimo and Pacella, 2011).

An introduction to multilinear PCA

PCA is a low-rank decomposition approach for feature extraction and analysis of one-dimensional arrays of data. In practice, the objective of PCA is to reduce the dimensionality of a large number of interrelated variables of the array by linearly transforming them to a new set of variables, called principal components (PCs), which are uncorrelated and ordered so that the first few retain most of the original data variation (Jolliffe, 2002).

Application of PCA to tensor data could be implemented (Henke et al., 1999; Summerhays et al., 2002) by reshaping tensors into one-dimensional arrays with a large number of variables. One issue in implementing this approach is that it breaks the structure and correlation in the original data, removing higher order dependencies present in the data set and losing potentially more compact or useful representations that can be obtained in the original form. Recently, the *Multilinear Principal Component Analysis* (MPCA) formulation for tensor feature extraction was proposed in the literature (Lu et al., 2008). This method operates directly on tensor data. The objective of MPCA is to perform linear transformations in all tensor modes seeking those bases in each mode that allow projected tensors to capture most of the variation present in the original data set.

In the remainder of this section first some basic concepts of multilinear algebra are introduced and then the MPCA method is presented with reference to the case study considered in this paper. For a more general discussion, the reader is referred to the original references (Lathauwer et al., 2000a; Lathauwer et al., 2000b; Lathauwer and Vandewalle, 2004).

Basic notation

The following convention is adopted herein: 1-order tensors (vectors) are denoted by lower-case boldface letters ($\mathbf{a}$), 2-order tensors (matrices) by upper-case boldface ($\mathbf{A}$) and 3rd-order tensors by calligraphic letters ($\mathcal{A}$). Their elements are denoted with indexes in brackets, where the indexes are denoted by lower-case letters and span the range from 1 to the upper-case letter index ($\mathbf{a}(p)$, $\mathbf{A}(p,q)$, $\mathcal{A}(p,q,n)$ where $p = 1,\ldots,P$, $q = 1,\ldots,Q$ and $n = 1,\ldots,N$).

Let (z,θ) denote the location at which the actual radius is measured, where z and θ are the vertical and angular coordinates associated with the sampled point. The same sampling strategy is assumed for all the surfaces, defining an ideal grid of $P \times Q$ equally spaced locations. Angular and vertical locations are selected among fixed levels, i.e., $z \in [z_1, \dots, z_p, \dots, z_P]$ and $\theta \in [\theta_1, \dots, \theta_q, \dots, \theta_Q]$. Therefore, the coordinates (z, θ) can be represented by the couple of indexes (p, q).

Let $v_{n,(p,q)}$ represent the difference between the actual radius and the nominal radius observed at location of indexes (p, q) on the n-th $(n = 1, \dots, N)$ surface. Note that $v_{n,(p,q)}$ represents the response variable of interest as a function of other location variables, similar to what is usually done for representing 2.5D surfaces, where the deviation of the real surface from its ideal pattern is commonly modeled as a function of the other two coordinates. The matrix $\mathbf{V}_n \in \mathbb{R}^{P \times Q}$, whose element of row p and column q is $\mathbf{V}_n(p, q) = v_{n,(p,q)}$, represents the surface data the n-th item.

A data set of N cylindrical surfaces is stored in a 3rd-order tensor denoted as $\mathcal{V} \in \mathbb{R}^{P \times Q \times N}$ and is addressed by 3 indexes, p, q and n related to the so-called 1-mode, 2-mode and 3-mode of $\mathcal{V}$. By definition, the j-mode vectors $(j = 1, 2, 3)$ of $\mathcal{V}$ are defined as the vectors obtained from by *varying* the j-th index while keeping all the others fixed. *Unfolding* the tensor $\mathcal{V}$ along the 1-mode (2-mode or 3-mode) is denoted by $\mathbf{V}_{(1)} \in \mathbb{R}^{P \times (Q \cdot N)}$ ($\mathbf{V}_{(2)} \in \mathbb{R}^{Q \times (N \cdot P)}$ or $\mathbf{V}_{(3)} \in \mathbb{R}^{N \times (P \cdot Q)}$) and represents the matrix whose column vectors are the 1-mode (2-mode or 3-mode) vectors of $\mathcal{V}$.

SVD with tensor notation

Consider a data set of N one-dimensional arrays of P variables, where the data set is summarized in a matrix $\mathbf{V} \in \mathbb{R}^{P \times N}$ (the rows represent the variables and the columns are the samples). Let $\mathbf{V}^c$ represent the centered data matrix obtained from $\mathbf{V}$ by subtracting the mean data samples from each sample. The covariance matrix of $\mathbf{V}$ is $\hat{\mathbf{\Sigma}} = \frac{1}{N-1}\mathbf{V}^c\mathbf{V}^{c^T}$, and is obtained re-scaling by a constant $1/(N-1)$ the *scatter* matrix of $\mathbf{V}^c$ defined as $\mathbf{V}^c\mathbf{V}^{c^T}$. The *Singular Value Decomposition* (SVD) of $\mathbf{V}^c$ is as follows.

$$\mathbf{V}^c = \mathbf{U}\mathbf{S}\mathbf{Z}^T = \mathbf{U}^{(1)}\mathbf{S}\mathbf{U}^{(2)^T} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)}. \tag{1}$$

$\mathbf{U} \in \mathbb{R}^{P \times P}$ is a unitary matrix (i.e., it has columns which form an orthonormal basis of $\mathbb{R}^P$), whose columns are the eigenvectors of the scatter matrix $\mathbf{V}^c\mathbf{V}^{c^T}$ (and of $\hat{\mathbf{\Sigma}}$). These eigenvectors are called the *left singular vectors* of $\mathbf{V}^c$. $\mathbf{S} \in \mathbb{R}^{P \times N}$ is pseudodiagonal and contains the *singular values* of $\mathbf{V}^c$. Without loss of generality, the elements in $\mathbf{S}$ are assumed arranged in a decreasing order (columns of $\mathbf{U}$ and $\mathbf{Z}$ are correspondingly arranged). $\mathbf{Z} \in$

$\mathbb{R}^{N\times N}$ is a unitary matrix, whose columns are the eigenvectors of the scatter matrix $\mathbf{V}^{c^T}\mathbf{V}^c$, also called the *right singular vectors* of $\mathbf{V}^c$.

Tensor notation on the right-hand in Equation (1) is as follows. The j-mode product ($j = 1, 2$) of matrix $\mathbf{S} \in \mathbb{R}^{P\times N}$ by matrices $\mathbf{U}^{(1)} \in \mathbb{R}^{P\times P}$ and $\mathbf{U}^{(2)} \in \mathbb{R}^{N\times N}$, is denoted by $\mathbf{S} \times_j \mathbf{U}^{(j)}$ and results in a matrix with entries $(\mathbf{S} \times_1 \mathbf{U}^{(1)})(j_1, i_2) = \sum_{i_1=1}^{P} \mathbf{S}(i_1, i_2)\mathbf{U}^{(1)}(j_1, i_1)$ and $(\mathbf{S} \times_2 \mathbf{U}^{(2)})(i_1, j_2) = \sum_{i_2=1}^{N} \mathbf{S}(i_1, i_2)\mathbf{U}^{(2)}(j_2, i_2)$. As the columns of matrix $\mathbf{U} = \mathbf{U}^{(1)}$ form a unitary basis for the space $\mathbb{R}^P$, the coordinates in this new basis for the data set stored in $\mathbf{V}^c$ are $\mathbf{U}^T\mathbf{V}^c = \mathbf{S}\mathbf{Z}^T$, i.e., $\mathbf{V}^c \times_1 \mathbf{U}^{(1)^T} = \mathbf{S} \times_2 \mathbf{U}^{(2)}$.

The aim of PCA is to solve the problem of approximating the data matrix $\mathbf{V}^c$ with another matrix $\widetilde{\mathbf{V}}^c$ which has a lower rank ($rank(\widetilde{\mathbf{V}}^c) < rank\,(\mathbf{V}^c)$ - a low-rank representation of data), where the approximation objective is to minimize the distance between $\mathbf{V}^c$ and $\widetilde{\mathbf{V}}^c$:

$$\underset{\widetilde{\mathbf{V}}^c}{min}(\|\mathbf{V}^c - \widetilde{\mathbf{V}}^c\|_F) \quad subject\ to \quad rank\left(\widetilde{\mathbf{V}}^c\right) \leq P'. \tag{2}$$

$\|\cdot\|_F$ denotes the *Frobenius norm*, which is by definition the squared root of the sum of each squared entry of its argument. P' is a given upper bound for the rank of matrix $\widetilde{\mathbf{V}}^c$ ($P' < P$). By the SVD in Equation (1), matrix $\widetilde{\mathbf{V}}^c$ can be obtained as $\widetilde{\mathbf{V}}^c = \widetilde{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)}$ where $\widetilde{\mathbf{S}}$ is the same matrix as $\mathbf{S}$ except that it contains only the P' largest singular values (the other singular values are replaced by zero). Hence, denoting by $\widetilde{\mathbf{U}}^{(1)}$ the matrix formed by the first P' columns of $\mathbf{U}^{(1)}$, which correspond to the first P' larger singular values of $\mathbf{V}^c$, a data sample vector of P points $\mathbf{v}_n$ ($n = 1, \ldots, N$) is projected to a feature space as $\widetilde{\mathbf{U}}^{(1)^T}(\mathbf{v}_n - \overline{\mathbf{v}}) = (\mathbf{v}_n - \overline{\mathbf{v}}) \times_1 \widetilde{\mathbf{U}}^{(1)^T}$. This is a set of P' coordinates which represent the so-called *scores* (PC-features) of vector $\mathbf{v}_n$ in the subspace $\widetilde{\mathbf{U}}^{(1)}$.

HOSVD as a generalization of SVD

In the targeted application the data set is summarized in a 3rd-order *tensor* with real entries. Let $\mathcal{V} \in \mathbb{R}^{P\times Q\times N}$ represent the tensor of data, and $\mathcal{V}^c$ the centered data tensor obtained from $\mathcal{V}$ by subtracting the mean $\overline{\mathbf{V}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{V}_n$ from each sample $\mathbf{V}_n$ ($n = 1, \ldots, N$). Tensor $\mathcal{V}^c$ is decomposed using the generalization of the SVD in Equation (1) called *Higher Order Single Value Decomposition* (HOSVD) in multilinear algebra (Lathauwer et al., 2000a). In particular, the centered 3rd-order tensor $\mathcal{V}^c$ can be decomposed as follows.

$$\mathcal{V}^c = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}, \tag{3}$$

where $\mathcal{S} \in \mathbb{R}^{P\times Q\times N}$, $\mathbf{U}^{(1)} \in \mathbb{R}^{P\times P}$, $\mathbf{U}^{(2)} \in \mathbb{R}^{Q\times Q}$ and $\mathbf{U}^{(3)} \in \mathbb{R}^{N\times N}$. The j-mode product ($j = 1, 2, 3$) of the tensor $\mathcal{S}$ by matrix $\mathbf{U}^{(j)}$ is denoted as $\mathcal{S} \times_j \mathbf{U}^{(j)}$ and results in a tensor with entries $(\mathcal{S} \times_1 \mathbf{U}^{(1)})(j_1, i_2, i_3) = \sum_{i_1=1}^{P} \mathcal{S}(i_1, i_2, i_3)\mathbf{U}^{(1)}(j_1, i_1)$, $(\mathcal{S} \times_2 \mathbf{U}^{(2)})(i_1, j_2, i_3) = \sum_{i_2=1}^{Q} \mathcal{S}(i_1, i_2, i_3)\mathbf{U}^{(2)}(j_2, i_2)$, and $(\mathcal{S} \times_3 \mathbf{U}^{(3)})(i_1, i_2, j_3) = \sum_{i_3=1}^{N} \mathcal{S}(i_1, i_2, i_3)\mathbf{U}^{(3)}(j_3, i_3)$ respectively.

$\mathcal{S}$ is called the *core tensor* for the HOSVD of $\mathcal{V}^c$. It is a 3rd-order tensor that contains the so-called 1-mode, 2-mode and 3-mode *singular values* of $\mathcal{V}^c$. $\mathbf{U}^{(1)} = \left(\mathbf{u}_1^{(1)} \cdots \mathbf{u}_p^{(1)} \cdots \mathbf{u}_P^{(1)}\right)$, $\mathbf{U}^{(2)} = \left(\mathbf{u}_1^{(2)} \cdots \mathbf{u}_q^{(2)} \cdots \mathbf{u}_Q^{(2)}\right)$ and $\mathbf{U}^{(3)} = \left(\mathbf{u}_1^{(3)} \cdots \mathbf{u}_n^{(3)} \cdots \mathbf{u}_N^{(3)}\right)$ in Equation (3) are unitary matrices in $\mathbb{R}^{P\times P}$, $\mathbb{R}^{Q\times Q}$ and $\mathbb{R}^{N\times N}$ respectively.

Considering the columns of matrices $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ and $\mathbf{U}^{(3)}$, the HOSVD of Equation (3) can be also rewritten as follows.

$$\mathcal{V}^c = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{n=1}^{N} \mathcal{S}(p, q, n) \cdot \left(\mathbf{u}_p^{(1)} \circ \mathbf{u}_q^{(2)} \circ \mathbf{u}_n^{(3)}\right), \tag{4}$$

where $\mathbf{u}_p^{(1)} \circ \mathbf{u}_q^{(2)} \circ \mathbf{u}_n^{(3)}$ denotes the *outer product* of vectors $\mathbf{u}_p^{(1)}$, $\mathbf{u}_q^{(2)}$ and $\mathbf{u}_n^{(3)}$ and results in a 3rd-order tensor in which the generic element of indexes i_1, i_2, i_3 is obtained as $\mathbf{u}_p^{(1)}(i_1) \cdot \mathbf{u}_q^{(2)}(i_2) \cdot \mathbf{u}_n^{(3)}(i_3)$ for all values of indexes i_1,i_2 and i_3.

A way to compute the HOSVD of the tensor $\mathcal{V}^c$ is through the common SVD procedure on the results of matrix unfolding (i.e., the SVD of matrices $\mathbf{V}^c_{(1)}$, $\mathbf{V}^c_{(2)}$ and $\mathbf{V}^c_{(3)}$ in order to obtain matrices $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ and $\mathbf{U}^{(3)}$ respectively). Then, the core tensor can be obtained from Equation (3) as $\mathcal{S} = \mathcal{V}^c \times_1 \mathbf{U}^{(1)^T} \times_2 \mathbf{U}^{(2)^T} \times_3 \mathbf{U}^{(3)^T}$.

This HOSVD solution, which extends SVD to higher order tensors through the calculation of different matrix SVDs of unfolded matrices, was first formulated in (Lathauwer et al., 2000a). In our study, the *Higher-order orthogonal iteration algorithm* proposed in (Lathauwer et al., 2000b) and based on the MATLAB implementation described in (Kolda, 2001; Bader and Kolda, 2006) is exploited for computing HOSVD.

By using the HOSVD theoretical results and algorithms presented in (Lathauwer et al., 2000b), approaches for multilinear truncation and best approximation for dimensionality reduction, which follows the classical PCA paradigm, were recently proposed in the literature. MPCA, introduced by (Lu et al., 2008), is one of the multilinear dimension reduction techniques applicable to tensor data. The MPCA algorithm is summarized in the following subsection.

Multilinear Principal Component Analysis (MPCA)

The simplest method to implement dimensionality reduction in the multilinear case is based on truncation of the unitary matrices in Equation (3), analogously to the truncation of unitary matrices in PCA.

The HOSVD of $\mathcal{V}^c$ can be truncated by keeping $P' \leq P$ orthonormal columns for the unitary matrix $\mathbf{U}^{(1)}$ in the 1-mode and, simultaneously, $Q' \leq Q$ orthonormal columns for the unitary matrix $\mathbf{U}^{(2)}$ in the 2-mode, such that the projected tensor captures the most of variation observed in the original tensor.

Through the two matrices $\widetilde{\mathbf{U}}^{(1)}$ and $\widetilde{\mathbf{U}}^{(2)}$ ($\widetilde{\mathbf{U}}^{(1)} \in \mathbb{R}^{P \times P'}$ and $\widetilde{\mathbf{U}}^{(2)} \in \mathbb{R}^{Q \times Q'}$) a basis of $P' \cdot Q'$ tensors for the space $\mathbb{R}^{P \times Q}$ can be derived. Each tensor of the basis is computed as the outer product $(\mathbf{u}_p^{(1)} \circ \mathbf{u}_q^{(2)})$, where $\mathbf{u}_p^{(1)}$ is the column of index p ($p = 1, \ldots, P'$) of $\widetilde{\mathbf{U}}^{(1)}$, $\mathbf{u}_q^{(2)}$ is the column of index q ($q = 1, \ldots, Q'$) of $\widetilde{\mathbf{U}}^{(2)}$ (note that, in our case, the tensors of the basis are 2nd-order tensors, i.e., matrices). These basis tensors are called *eigentensors* in the literature (Lu et al., 2008). Thus, a centered data samples $(\mathbf{V}_n - \overline{\mathbf{V}})$ ($n = 1, \ldots, N$), is projected to a lower-dimension feature matrix as follows.

$$\widetilde{\mathbf{Y}}_n = (\mathbf{V}_n - \overline{\mathbf{V}}) \times_1 \widetilde{\mathbf{U}}^{(1)^T} \times_2 \widetilde{\mathbf{U}}^{(2)^T}, \tag{5}$$

where $\widetilde{\mathbf{Y}}_n \in \mathbb{R}^{P' \times Q'}$ ($n = 1, \ldots, N$) is the the projected matrix that, in analogy with common PCA, can be referred to as the matrix of scores.

Similarly to the optimization problem in Equation (2) for the case of regular PCA, the problem can be formulated in terms of an approximation of tensor $\mathcal{V}^c$ with another tensor $\widetilde{\mathcal{V}}^c$, where the objective is to minimize the norm of the distance between them. In particular, assuming that the targeted dimensionality (P' and Q') is specified in advance, the optimization problem can be formulated as follows (where $\widetilde{\mathbf{V}}^c_{(1)}$ and $\widetilde{\mathbf{V}}^c_{(2)}$ are the result of matrix unfolding of $\widetilde{\mathcal{V}}^c$).

$$\underset{\widetilde{\mathcal{V}}^c}{min}(\|\mathcal{V}^c - \widetilde{\mathcal{V}}^c\|_F) \quad subject\ to \quad rank\left(\widetilde{\mathbf{V}}^c_{(1)}\right) \leq P', rank\left(\widetilde{\mathbf{V}}^c_{(2)}\right) \leq Q', \tag{6}$$

Differently from the optimization problem in Equation (2), the optimal solution for Equation (6) is not univocally given by a truncation of matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$. In fact, $\mathcal{S}$ is a full tensor, instead of being pseudodiagonal as $\mathbf{S}$ in Equation (1). In the MPCA method, introduced by (Lu et al., 2008), for finding the projection matrices $\widetilde{\mathbf{U}}^{(1)}$ and $\widetilde{\mathbf{U}}^{(2)}$, an iterative algorithm was proposed in order to solve the optimization problem of Equation (6). The objective of the iterative algorithm is to define a multilinear transformation that maps the original data set into a new data set, which captures most of the variations observed in the original data, assuming that these variations are measured by the total tensor scatter.

In particular, denoting by $\psi = \sum_{n=1}^{N} \|(\mathbf{V}_n - \overline{\mathbf{V}}) \times_1 \widetilde{\mathbf{U}}^{(1)^T} \times_2 \widetilde{\mathbf{U}}^{(2)^T}\|_F^2$ the total scatter of all projected samples, and assuming that the dimensionality for each mode (P' and Q') is known or predetermined, the objective of the MPCA iterative algorithm is the determination of the two projection matrices $\widetilde{\mathbf{U}}^{(1)}$ and $\widetilde{\mathbf{U}}^{(2)}$ that maximizes:

$$\begin{aligned} &\underset{\widetilde{\mathbf{U}}^{(1)},\widetilde{\mathbf{U}}^{(2)}}{max} \left(\sum_{n=1}^{N} \|(\mathbf{V}_n - \overline{\mathbf{V}}) \times_1 \widetilde{\mathbf{U}}^{(1)^T} \times_2 \widetilde{\mathbf{U}}^{(2)^T}\|_F^2 \right), \\ &subject \;\; to \quad rank\left(\widetilde{\mathbf{U}}^{(1)}\right) \leq P', rank\left(\widetilde{\mathbf{U}}^{(2)}\right) \leq Q'. \end{aligned} \tag{7}$$

By Equation (5), the total scatter ψ can be rewritten as $\psi = \sum_{n=1}^{N} \|\widetilde{\mathbf{Y}}_n\|_F^2$. By defining $\widetilde{y}_{p',q'} = \sum_{n=1}^{N} \left(\widetilde{\mathbf{Y}}_n(p',q')\right)^2$ the portion of total scatter explained by the projection on the eigentensor of index p' and q' in the first and second mode respectively ($p' \leq P'$, $q' \leq Q'$), we have $\psi = \sum_{p'} \sum_{q'} \widetilde{y}_{p',q'}$. The ratio $\psi/\|\mathcal{V}^c\|_F^2$ is the percentage measure of the total scatter in the original data set which is projected to the feature space of multilinear PCs.

When the targeted (reduced) dimensionality in each mode (P', Q') is not specified in advance, its value has to be determined. An approximate approach for dimensionality reduction consists of truncating the 1-mode and 2-mode eigenvectors such that the retained portion of the total scatter in the 1-mode and 2-mode is about equal to a given value. This approach, which is an extension of the dimensionality reduction strategy of the regular PCA to the multilinear case, allows retaining a reduced number of features such that a given percentage of the overall variability is captured in each mode. However, in the multilinear case the dimensionality reduction in one mode cannot be determined independently from the other, since eigenvalues in other modes are affected by the eigenvector truncation in a given mode (Theorem 2 in (Lu et al., 2008)). A different approach for dimensionality reduction was proposed in (Lu et al., 2008): it is based on an iterative procedure called *Sequential Mode Truncation* (SMT), which takes into account this effect. The SMT procedure allows solving the dimensionality reduction problem by an iterative truncation of both the 1-mode and 2-mode eigenvectors. Hereafter, the STM procedure will be considered for dimensionality reduction in MPCA.

Differently from regular PCA, the extracted features by MPCA may be correlated. For this reason, other multilinear low-rank reduction techniques applicable to tensor data were also considered in the literature (Paynabar et al., 2013; Yan et al., 2015; Paynabar et al., 2015). In particular, (Lu et al., 2009) proposed an extension of the MPCA named uncorrelated multilinear PCA (UMPCA), which introduces the zero-correlation constraint among features (eigentensors) derived from an iterative procedure aimed at finding directions capturing maximum variance. However, the zero-correlation constraints introduces a limitation on the maximum number of eigentensors that may be extracted, whereas the remaining portion of data variability may be captured by removing the zero-correlation constraint. In the frame of statistical process monitoring, the existence of correlation among the extracted features does not affect the performances, since the information content is summarized by the Hotelling's T^2 statistics. Because of this, the MPCA approach is adopted in our study instead of UMPCA.

Application to the real case study

The experimental study concerns a turning process which produced cylindrical surfaces with diameter 16.8 mm and length 42 mm. A set of 90 cylinders were considered according to the 3^2 full design reported in table (1). Machined surfaces were sampled using the same regular grid of points equally spaced (210×64 points).

For each surface, the deviations of the actual shape from the reference cylinder were computed. The resulting data set was summarized in a 3rd-order tensor denoted as $\mathcal{V} \in \mathbb{R}^{210\times64\times90}$, where each entry $\mathcal{V}(p, q, n)$ represents the deviation of the actual surface from the reference cylinder at the point with vertical coordinate of index p, angular coordinate of index q, for item of index n ($p = 1, \ldots, 210$, $q = 1, \ldots, 64$ and $n = 1, \ldots, 90$). Matrices $\mathbf{V}_n \in \mathbb{R}^{210\times64}$ represent the surfaces.

As first step of the analysis the average surface deviations from the reference cylinders is computed ($\overline{\mathbf{V}} = \frac{1}{90}\sum_{n=1}^{90} \mathbf{V}_n$). Figure (1) plots the colored parametric surface $\overline{\mathbf{V}}$ in a 3D diagram. In order to visually emphasize the surface deviations, these were plotted in a scale 250:1, superimposed on the nominal cylinder of the case study (an ideal cylinder of diameter 16.8 mm and length 42 mm). By a visual inspection of Figure (1), it appears that a systematic shape is characterizing the process "on-average". The mean deviation increases as the vertical coordinate increases resulting in a conical shape for the mean surface.

Figure 1: Mean surface based on 90 replicates plotted on a 3D colored parametric diagram (darker color refers to a minor deviation, a brighter color to a major deviation). Scale 250:1

Application of MPCA

Let $\mathcal{V}^c \in \mathbb{R}^{210\times64\times90}$ represent the centered data tensor obtained from $\mathcal{V}$ by subtracting the mean surface $\overline{\mathbf{V}}$ from each of the 90 samples.

Firslty, the MPCA algorithm was implemented in order to solve the optimization problem in Equation (7). Subsequently, the iterative SMT procedure (Lu et al., 2008) was implemented to solve the dimensionality reduction problem. Given the results of the SMT procedure, a natural criterion in selecting eigentensors to be included in the projection basis consists in sorting them by the portion of total scatter explained by each one, namely, $\widetilde{y}_{p',q'} = \sum_{n=1}^{N}\left(\widetilde{\mathbf{Y}}_n(p',q')\right)^2$, and then selecting the m most significant ones in order to obtain a given percentage measure of the total scatter in the original data set.

In particular, the sequence of 8 most significant eigentensors obtained on the data set of the reference case study are referred to as follows: $1/1, 1/2, 2/2, 2/1, 1/3, 3/5, 1/4$ and $1/6$ where each label is formed as p'/q' (p' is the index of the eigenvector in 1-mode, and q' is the index of the eigenvector in 2-mode). The portion of total scatter explained by each one is as follows: $\widetilde{y}_{1,1} = 0.5911(22.15\%)$, $\widetilde{y}_{1,2} = 0.3928(14.72\%)$, $\widetilde{y}_{2,2} = 0.2341(8.77\%)$, $\widetilde{y}_{2,1} = 0.2323(8.70\%)$, $\widetilde{y}_{1,3} = 0.0711(2.66\%)$, $\widetilde{y}_{3,5} = 0.0534(2.00\%)$, $\widetilde{y}_{1,4} = 0.0405(1.52\%)$ and $\widetilde{y}_{1,6} = 0.0342(1.21\%)$ where in brackets are reported the percentage portion in the total scatter of original data. Therefore, the total scatter of all projected samples on the first 8 most significant eigentensors is equal to $\psi = 1.6476$, which represents 61.73% of the total scatter in the original data set.

The shape of retained eigenvectors is commonly interpreted by means of a graphical representation. In the reference case study, each eigenvector in each mode can be plotted as a function of a location in the space. Figure (2) reports the diagrams of the first four 1-mode eigenvectors (vectors $\mathbf{u}_p^{(1)}$ in Equation (4) where $p' = 1, \ldots, 4$) as a function of the vertical coordinate index. Similar diagrams are depicted in Figure (3), with reference to the first six 2-mode eigenvectors (vectors $\mathbf{u}_q^{(2)}$ in Equation (4) where $q' = 1, \ldots, 6$), in which each eigenvector is depicted as a function of the angular coordinate index. From Figure (2) it can be observed that one important component of variability in the 1-mode is characterized by an increasing trend along the vertical direction. On the other hand, important components of variability in the 2-mode (i.e., along the angular coordinate) are periodic functions characterized by a frequency of 2 and 3 undulations per revolution (u.p.r.), where the bi-lobed shape (2 u.p.r.) explains the most important portion of variability in 2-mode. In particular, sinusoidal functions represent bi-lobed and three-lobed patterns, which can be observed along the angular direction, while (near) linear and quadratic patterns are observed along the vertical direction. These results are consistent to those observed in the literature for cylindrical surfaces obtained by manufacturing processes (Henke et al., 1999).

By considering the eigenvectors in 1-mode and 2-mode respectively, a basis of eigentensors, for the space $\mathbb{R}^{210\times 64}$, can be computed. Each eigentensor is formed by the outer product $(\mathbf{u}_p^{(1)} \circ \mathbf{u}_q^{(2)})$, where $\mathbf{u}_p^{(1)}$ are the retained eigenvectors in 1-mode and $\mathbf{u}_q^{(2)}$ are the retained eigenvectors in 2-mode. These eigentensors form a basis on which a centered surface (in $\mathbb{R}^{210\times 64}$) can be projected.

The 8 most significant eigentensors, which form a basis for the reference case study, are graphically depicted in Figure (4) in a 3D cylindrical coordinate system, where each label is formed as p'/q' ($p' = 1, 2, 3$ is the index of the eigenvector in 1-mode, and $q' = 1, 2, 3, 4, 5, 6$ is the index of the eigenvector in 2-mode). Form Figure (4) it can be observed that each eigentensor combines the shape of two specific eigenvectors, one in 1-mode and the other in

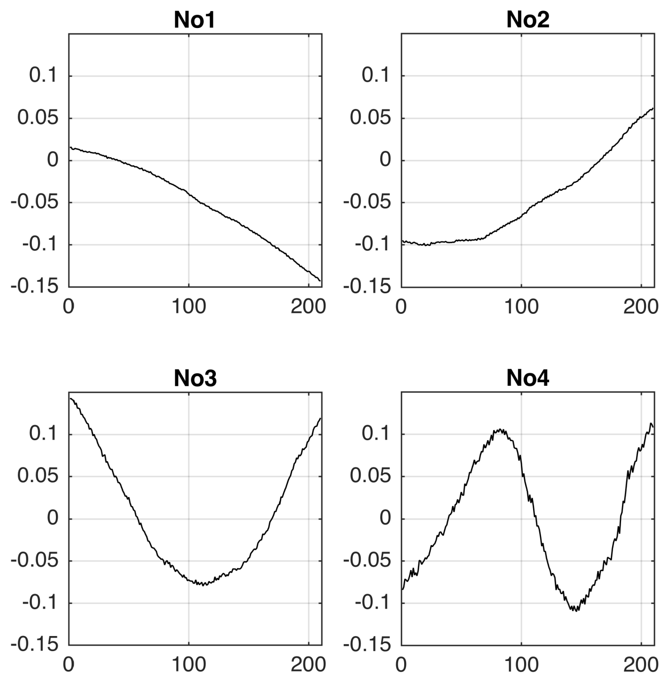

Figure 2: Plots of the first 4 eigenvectors related to 1-mode as a function of the vertical coordinate index $p = 1, \ldots, 210$

2-mode. As for instance, the most important eigentensor, obtained by combining the first eigenvector in 1-mode to the first eigenvector in 2-mode (labelled with 1/1 in Figure (4)) is a conical surface whose superior portion assume a bi-lobed contour. The conical shape along the vertical coordinate is due to the eigenvector in the 1-mode, while the bi-lobed shape along the angular coordinate to the eigenvector in the 2-mode. This result is similar to that reported in the literature where a conical shape along the vertical (referred to as ”taper” error) was defined as ”a dominant” form error of manufactured cylindrical surfaces (Henke et al., 1999; Summerhays et al., 2002).

The values of the 8 scores for each surface in the data set, obtained by projecting the centered surface on the eigentensors depicted in Figure (4), are plotted in the diagrams of Figure (5). Each panel in Figure (5) is related to a specific eigentensor/score (the same labeling convention of Figure (4) is adopted for Figure (5)). The 90 scores in each panel are depicted in a series plot where the abscissa is the index $n = 1, \ldots, 90$. By a rough inspection of Figure (5) it appears that scores can be influenced by the specific experimental condition.

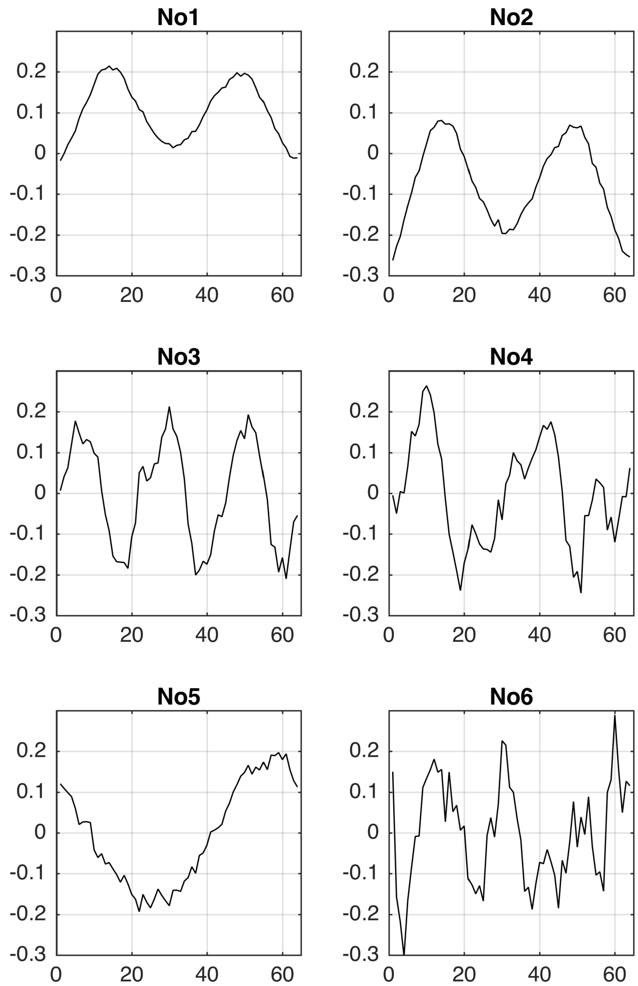

Figure 3: Plots of the first 6 eigenvectors related to 2-mode as a function of the angular coordinate index $q = 1, \ldots, 64$

Statistical monitoring procedure

A statistical monitoring procedure can be implemented by integrating the results of the MPCA with control charting. After defining all the features for each surface in the data set, any multivariate control charts including T^2 Hotelling control chart, Multivariate Cumulative Sum (MCUSUM), Multivariate Exponentially Weighted Moving Average (MEWMA), etc., can be used to monitor the vector of scores obtained by the MPCA. In this paper, we used the T^2 Hotelling control chart for monitoring the vector of 8 scores related to the features

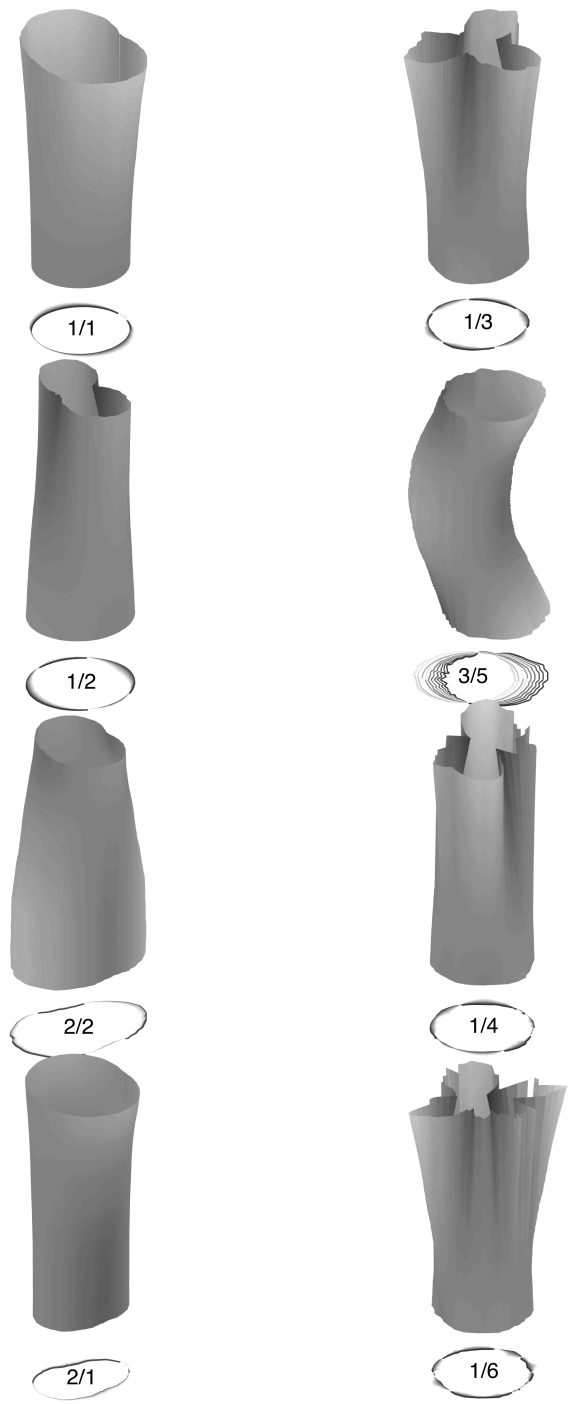

Figure 4: 8 eigentensors for the reference case study. Each label is formed as p'/q' (p' is the index of the eigenvector in 1-mode, and q' is the index of the eigenvector in 2-mode)

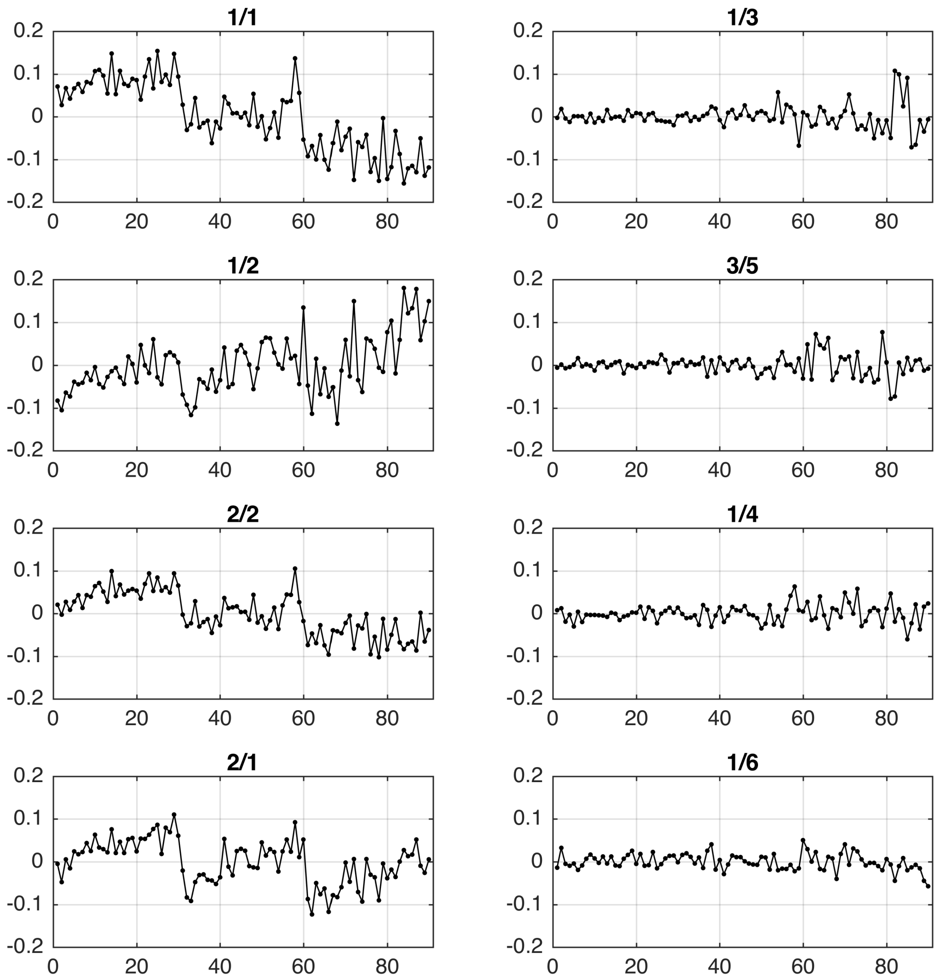

Figure 5: Scores versus sample index $n = 1, \ldots, 90$. Each panel depicts the scores obtained by projecting the (centered) cylindrical surfaces on the eigentensor of label p'/q' (p' is the index of the eigenvector in 1-mode, and q' is the index of the eigenvector in 2-mode)

depicted in Figure (4). In order to implement the control chart, i.e. to estimate the "in-control" process parameters (mean vector and covariance matrix to be used in computing the Hotelling's T^2 statistics), a subset of the 90 score vectors available was used. In particular, we used the first 30 samples in the data set for estimating in-control process parameters, while the remaining 60 was used to check whether the process parameters have changed, i.e. if any "assignable cause" affected the process. As a matter of fact, the first 30 samples are obtained from the lathe-turning process using a depth of cut equal to $0.4mm$, while the remaining subset of samples from the turning process using a greater depth of cut ($0.8mm$ and $1.2mm$).

Assessing the assumption of multivariate normality is required in order to define a given Type-I error control limit for the T^2 Hotelling control chart. Many statistical tests and graphical approaches are available to check the multivariate normality assumption. Among these, the Mardia's multivariate skewness and kurtosis statistics (Mardia, 1970) are commonly adopted. In our study the Mardia's multivariate skewness and kurtosis coefficients as well as their corresponding statistical significance were calculated on the set of 30 in-control scores computed by MPCA. Both the skewness and kurtosis estimates indicated multivariate normality. Therefore, according to Mardia's test, the data set of MPCA scores for the first 30 samples of the case study follows a multivariate normal distribution.

Starting from the set of scores computed by MPCA, the 99.73% upper control limit (UCL) for the T^2 Hotelling control chart resulted equal to $UCL = 23.75$, which is the percentile of the asymptotic distribution: a chi-square with degrees of freedom equal the of dimension of vectors observed over time. The control limit and the T^2 Hotelling statistics are depicted in Figure (6). With reference to the 90 vectors of 8 scores for each surface in the data set, obtained by projecting the centered surface on the eigentensors depicted in Figure (4), the T^2 Hotelling control chart is plotted in Figure (6). From Figure (6), it can be noted that the control chart triggers a total of 20 out-of-control signals for the data set of surfaces obtained by using a cut depth equal to $0.8mm$ (samples from 31 to 60) and a signal for each sample in the case of depth equal to $1.2mm$ (samples from 61 to 90).

Even limited to the data set in the specific case study of this paper, the T^2 Hotelling control chart in Figure (6) shows that integrating MPCA with multivariate control charting can be worth for implementing an effective statistical process control method when the quality characteristic of interest is a 2.5D surface.

One alternative approach to deal with multi-way arrays involves unfolding the multi-dimensional data into vectors in order to apply the regular PCA on the transformed data (Vectorized PCA - VPCA). VPCA is a generalization of PCA to tensor data, which applies the regular PCA to a tensor object reshaped into a vector. The VPCA is the multi-way PCA approach originally proposed in (Nomikos and MacGregor, 1995) for monitoirng batch processes. Although the eigenvectors obtained from VPCA are not as informative as the eigentensors of MPCA, the control chart approach based on the regular PCA (Colosimo and Pacella, 2007) may be applied to the vectorized data. The control chart considered herein is based on the Hotelling's T^2 statistics, used to detect possible deviations along the directions of the first m PCs. In our study, in order to compare the VPCA-based control chart to the MPCA one, we decided to retain a number of $m = 8$ PCs, i.e. the same number of retained features in the MPCA.

The 99.73% UCL and the T^2 Hotelling statistics of the VPCA-based approach are depicted in Figure (7). In this case, the control chart triggers a total of 24 out-of-controls for set of surfaces obtained by using a cut depth equal to $0.8mm$ (samples from 31 to 60) and a signal for each sample in the case of depth equal to $1.2mm$ (samples from 61 to 90). It is worth noting that by retaining of $m = 8$ PCs a percentage of about 70% of the variability in the original data set is explained by VPCA, while by using the MPCA the first 8 most significant eigentensors allow to retain a lesser percentage of variability (61.73%). As a matter of fact, in order to retain a 70% percentage of variability in the data by using MPCA, the number of most significant eigentensors required resulted equal to 22. The observed result implies that VPCA may be more parsimonious than MPCA in terms of the number of features required to capture the majority of variations. This is because, differently from regular PCA, the extracted features by MPCA may be correlated. Despite of this, with reference to the 90 vectors of 22 scores for each surface in the data set, obtained by projecting each centered surface on the 22 most significant eigentensors obtained from the MPCA, the T^2 Hotelling control chart outperforms the VPCA one. In this case, the 99.73% control limit (equal to $UCL = 44.90$) and the T^2 Hotelling statistics are depicted in Figure (8). It can be observed that a signal is triggered for each out-of-control sample both in the case of a cut depth equal to $0.8mm$ (samples from 31 to 60) and in the case of depth equal to $1.2mm$ (samples from 61 to 90).

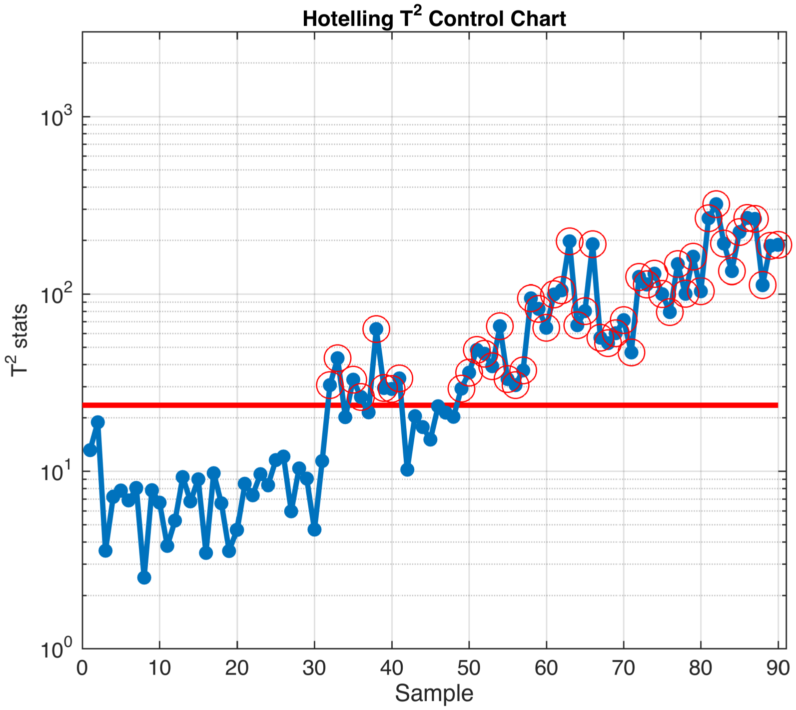

Figure 6: MPCA-based T^2 Hotelling control chart of scores obtained by projecting the (centered) cylindrical surfaces on the 8 most significant eigentensors (61.73% of explained variability). 99.73% UCL, in-control samples $n = 1 \dots 30$, out-of-control samples $n = 31 \dots 90$. Semi-log graph

Conclusions

2.5D surface data are being increasingly considered as quality characteristic of interest for manufacturing processes. One of the main challenges in monitoring this type of data for statistical process control is related to the high-dimensionality and complex spatial correlation structure of such data. Thus, effective dimension reduction and low-rank representation of these data is essential for better process monitoring.

Recent literature has shown that PCA can be implemented for detecting the systematic ways in which manufactured profiles vary around the mean profile. In this paper, it was shown that a generalization of PCA, namely the MPCA, can be used for interpreting and modeling manufactured 2.5D surfaces. The MPCA model can be effectively used to represent surface patterns. As any PCA-based methods the main advantage is its ease of use because no cumbersome activity of regressor selection is required.

A case study related to surfaces of mechanical components obtained by a real lathe-

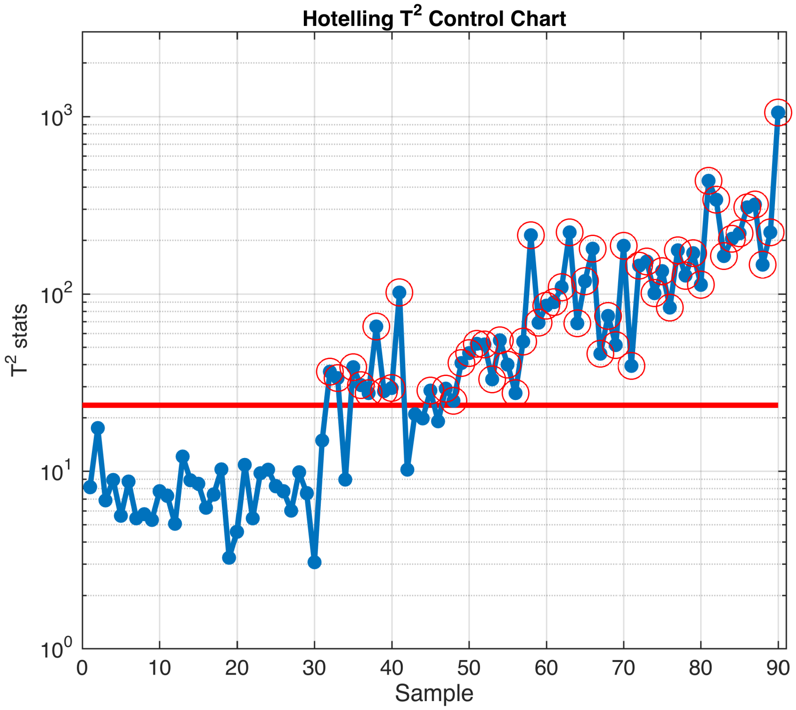

Figure 7: VPCA-based T^2 Hotelling control chart of scores for the 8 most significant PCs (70% of explained variability). 99.73% UCL, in-control samples $n = 1 \dots 30$, out-of-control samples $n = 31 \dots 90$. Semi-log graph

turning process was used as reference. The data set of different cylindrical surfaces was summarized in a tensor object. The MPCA results in a multilinear projection that projects the original tensor objects into a lower-dimensional tensor subspace while preserving the variation in the original data. A basis formed by eigentensors was identified to model the cylindrical surfaces in the reference case study. Plot in a 3D space of the eigenvecotrs in the basis provide some interesting information on the sources of variability behind collected data. The plot on a Cartesian diagram of the projection coefficients (scores) associated to the eigentensors in the basis allows one to gain more insight on the relationship between process parameters and the shape of obtained surfaces. Finally, a statistical monitoring procedure can be implemented by integrating the results of the iterative MPCA with control charting. In particular, any multivariate control charts can be used to monitor the vector of scores obtained by the MPCA.

By comparing the MPCA to regular PCA, e.g. the VPCA consisting of applying PCA to vectors generated by unfolding the original multi-way data-set, the MPCA could be a more

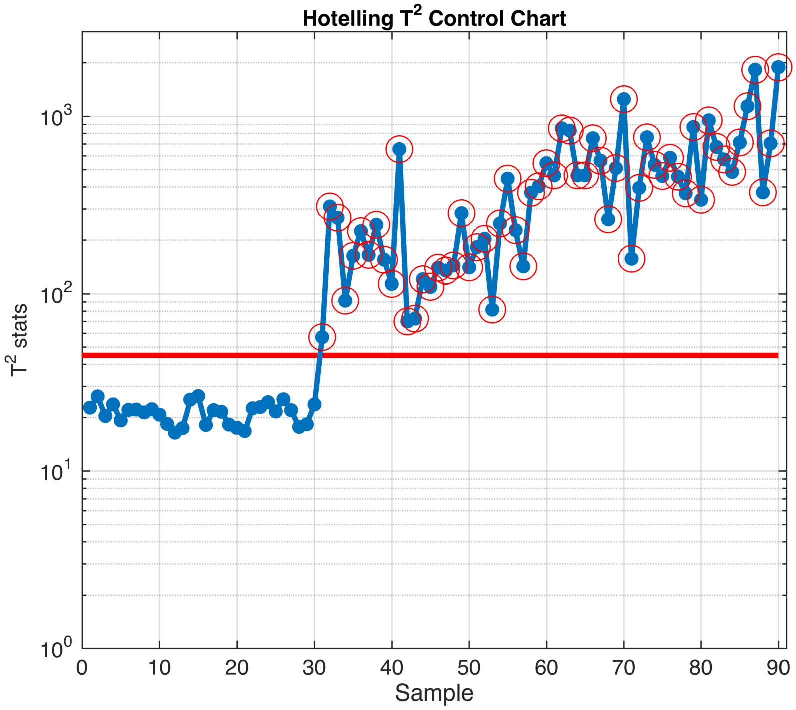

Figure 8: MPCA-based T^2 Hotelling control chart of scores obtained by projecting the (centered) cylindrical surfaces on the 22 most significant eigentensors (70% of explained variability). 99.73% UCL, in-control samples $n = 1 \dots 30$, out-of-control samples $n = 31 \dots 90$. Semi-log graph

efficient method from a computational and memory saving point of view, especially when high numbers of data points are involved or in the presence of more complex data structures, as for example in case of Red, Green, and Blue (RGB)-color images. Furthermore, analysis of out-of-controls released during statistical process monitoring may also benefit of the MPCA method, since MPCA may lead to easier interpretation of retained features than VPCA. As a matter of fact, the eigenvectors obtained from the VPCA are not as informative as the eigentensors of MPCA. Further research and simulation efforts are needed to further clarify the benefits and limitations of the proposed approach in different scenarios of surface monitoring.

As a summary, the topic of surface monitoring via the multilinear dimension reduction techniques applicable to tensor data, such is MPCA, appeared to be a promising field of research for the future. Considering that the approach proposed in this paper is quite general, one direction of future research consists in a revision and extension of the proposed method that could be tested and compared with existing procedures to detect changes in the pattern of manufactured surfaces. Furthermore, as fault diagnosis and root-cause identification is an important task after a change is detected, development of a fault diagnosis methods that can be integrated with the process monitoring is an important, yet challenging research topic that deserves further study.

Acknowledgment

This work has been partially funded by the Ministry of Education, University and Research of Italy (MIUR). The authors thank the Machining Tool and Production Systems Research Laboratory (MUSP) in Piacenza (Italy) for providing the actual measurement data set. The authors also thank the editor and two anonymous referees for their comments that have resulted in improvements in the article.

References

Bader, B. W. and Kolda, T. G. (2006). "Algorithm 862: Matlab Tensor Classes for Fast Algorithm Prototyping". *ACM Transactions on Mathematical Software*, 32(4), pp. 635–653.

Colosimo, B. M.; Cicorella, P.; Pacella, M.; and Blaco, M. (2014). "From profile to surface Monitoring: SPC for cylindrical surfaces via Gaussian Processes". *Journal of Quality Technology*, 46(2), pp. 95–113.

Colosimo, B. M.; Mammarella, F.; and Petro', S. (2010). "Quality Control of Manufactured Surfaces". In Lenz, H. J. and Wilrich, P. T., editors, *Frontiers of Statistical Quality Control*. Springer, London.

Colosimo, B. M. and Pacella, M. (2007). "On the Use of Principal Component Analysis to Identify Systematic Patterns in Roundness Profiles". *Quality and Reliability Engineering International*, 23(6), pp. 707–725.

Colosimo, B. M. and Pacella, M. (2010). "A comparison study of control charts for statistical monitoring of functional data". *International Journal of Production Research*, 48(6), pp. 1575–1601.

Colosimo, B. M. and Pacella, M. (2011). "Analyzing the effect of process parameters on the shape of 3D Profiles". *Journal of Quality Technology*, 43(3), pp. 169–195.

Colosimo, B. M.; Pacella, M.; and Semeraro, Q. (2008). "Statistical process control for geometric specifications: on the monitoring of roundness profiles". *Journal of Quality Technology*, 40(1), pp. 1–18.

Colosimo, B. M. and Senin, N., editors (2011). *Geometric Tolerances. Impact on Product Design, Quality Inspection and Statistical Process Monitoring*. Springer, London.

Dowling, M. M.; Griffin, P. M.; Tsui, K.-L.; and Zhou, C. (1997). "Statistical issues in geometric feature inspection using coordinate measuring machines". *Technometrics*, 39(1), pp. 3–17.

Henke, R. P.; Summerhays, K. D.; Baldwin, J. M.; Cassou, R. M.; and Brown, C. W. (1999). "Methods for evaluation of systematic geometric deviations in machined parts and their relationships to process variables". *Precision Engineering*, 23(4), pp. 273–292.

ISO (2012). *Geometrical Product Specifications (GPS) – Geometrical tolerancing – Tolerances of form, orientation, location and run-out*. ISO 1101:2012, International Standard Organization, Geneva, Switzerland.

Jolliffe, I. T. (2002). *Principal component analysis*. Springer Series in Statistics. Springer, New York.

Kolda, T. G. (2001). "Orthogonal tensor decompositions". *SIAM Journal of Matrix Analysis and Applications*, 23(1), pp. 243–255.

Lathauwer, L. D.; Moor, B. D.; and Vandewalle, J. (2000a). "A multilinear singular value decomposition". *SIAM Journal of Matrix Analysis and Applications*, 21(4), pp. 1253–1278.

Lathauwer, L. D.; Moor, B. D.; and Vandewalle, J. (2000b). "On the best rank-1 and rank-$(R_1, R_2, \ldots, R_N)$ approximation of higher-order tensors". *SIAM Journal of Matrix Analysis and Applications*, 21(4), pp. 1324–1342.

Lathauwer, L. D. and Vandewalle, J. (2004). "Dimensionality reduction in higher-order signal processing and rank-$(R_1, R_2, \ldots, R_N)$ reduction in multilinear algebra". *Linear Algebra and its Applications*, 391(1), pp. 31–55.

Lu, H.; Plataniotis, K. N.; and Venetsanopoulos, A. N. (2008). "MPCA: Multilinear Principal Component Analysis of Tensor Objects". *IEEE Transactions on Neural Networks*, 19(1), pp. 18–39.

Lu, H.; Plataniotis, K. N.; and Venetsanopoulos, A. N. (2009). "Uncorrelated Multilinear Principal Component Analysis for Unsuper- vised Multilinear Subspace Learning". *IEEE Transactions on Neural Networks*, 20(11), pp. 1820–1836.

Mardia, K. V. (1970). "Measures of multivariate skewness and kurtosis with applications". *Biometrika*, 57(3), pp. 519–530.

Nomikos, P. and MacGregor, J. F. (1995). "Multivariate SPC Charts for Monitoring Batch Processes". *Technometrics*, 37(1), pp. 41–59.

Noorossana, R.; Saghaei, A.; and Amiri, A., editors (2011). *Statistical Analysis of Profile Monitoring*. Wiley, New York.

Paynabar, K.; Jin, J.; and Pacella, M. (2013). "Monitoring and diagnosis of multichannel nonlinear profile variations using uncorrelated multilinear principal component analysis". *IIE Transactions*, 45(11), pp. 1235–1247.

Paynabar, K.; Qiu, P.; and Zou, C. (2015). "A Change Point Approach for Phase-I Analysis in Multivariate Profile Monitoring and Diagnosis". *Technometrics*. in press.

Summerhays, K. D.; Henke, R. P.; Baldwin, J. M.; Cassou, R. M.; and Brown, C. W. (2002). "Optimizing discrete point sample patterns and measurement data analysis on internal cylindrical surfaces with systematic form deviations". *Precision Engineering*, 26(1), pp. 105–121.

Yan, H.; Paynabar, K.; and Shi, J. (2015). "Image-based process monitoring using low-rank tensor decomposition". *IEEE Transactions on Automation Science and Engineering*, 12(1), pp. 216–227.