Manufacturing-based sampling strategies for geometric tolerance inspection via CMM

Bianca Maria Colosimo, Ester Gutierrez Moya,
Giovanni Moroni and Stefano Petrò

Abstract

More and more often, geometric specifications (such as straightness, roundness, flatness, cylindricity, free-form profiles and surfaces) appear as requirements for machined items in technical drawings. Typically, Coordinate Measuring Machines (CMM) are used to inspect the machined feature and the cloud of measured points is used to check whether the machine item should be scrapped or not. In fact, the measured points are used to compute the form error with reference to the specific tolerance under study. The estimated error is usually affected by uncertainty which depends on both the sample size (number of points measured on a given profile/surface - which mainly affects the sampling costs) and the sampling strategy (position of the points on the profile/surface). The present research work focuses on the second issue and investigates the performance of several existing sampling strategies and the ones achieved by two newly presented approaches. These new approaches explore advantages arising by including considerations provided by the manufacturing process in the sampling strategy design. Throughout the paper, a real case study concerning flatness tolerance is used as reference.

Keywords: sampling strategy, geometric tolerance, measuring, CMM, variable selection, PCA

1 Introduction

Inspection of geometric tolerances is a complex task because geometric errors are related to three-dimensional features and estimation of this type of errors is usually based on a cloud of points that has to be measured on the machined surface. CMMs (Coordinate Measurement Machines) are the most diffused instruments for 3D measurement in the mechanical field, because of their accuracy and flexibility. Recently, approaches for geometric shape monitoring have been proposed in the SPC (Statistical process control) literature [1, 2]. These approaches assume that, for each specific nominal geometry (circle, cylinder, line, plane, etc.), a given set of points is measured via CMM on the new machined item. These points are then elaborated for inspection or monitoring purposes, where inspection deals with deciding whether or not the machined item is conforming to requirements (by computing the form error and comparing it with the specified tolerance limit), while monitoring is aimed at detecting an out-of-control state of the manufacturing process (using approaches as the one presented in xxxxx).

In order to perform either inspection or monitoring or both, a basic issue consists in selecting the measurement strategy, which consists in solving two main issues. The first consists in selecting the sample size of the cloud of points that has to be measured, while the second issue consists in selecting the exact position of each measured point. This paper deals with this second issue, which will be referred to as sampling strategy in the following. The selection of a sampling strategy becomes more and more relevant as the sample size decreases. In fact, when the cloud of points is particularly dense (i.e., thousands of points), the required tolerance is well estimated despite of the exact position of each sampled points. However, on-line applications of tolerance estimation require reduced sample size (because measurement costs are directly related to the number of points that have to be measured). When the sample size is reduced, the exact position of the measurement points play a relevant role. Most of the approaches proposed in the literature and in the international standard , e.g., ISO 12781 for flatness [3], suggest the use of strategies which are defined a priori, despite of the specific manufacturing process which has been used to produce the feature. Examples are the uniform, random or quasi-random (e.g., Hammersley sequence [4]) distribution of the sample points on the surface that has to be inspected. All these sampling strategies will be referred to as “a-priori strategies”.

A second approach proposed in the literature consists in adopting “adaptive sampling strategies” [5, 6, 7, 8]. Adaptive sampling is a multi-step methodology, which starts with a low density, usually uniformly spaced, sampling of the feature of interest. Given information collected in this starting sample, the adaptive algorithm selects the next sampling points. The procedure is iterated until a required precision of the tolerance estimate is achieved. Although effective, adaptive sampling strategies are hard to implement in traditional CMMs, and this is probably why they did not receive great attention in the literature and in industrial practice.

Given that each feature is obtained with a specific process, a different approach to sampling strategy consists in considering that each particular production process leaves a particular "fingerprint" on the machined surface. This fingerprint will be referred to as the "manufacturing signature". As an example, consider the increase in the diameter of cylindrical parts which can be observed when workpieces are obtained by turning. This particular signature is due to an inflection of the workpiece while the tool is moving far from the spindle. The manufacturing signature can be defined as the systematic pattern which characterizes all the features machined with a given process. This systematic pattern will be obviously "masked" by additional random noise caused by the natural variability of the process (e.g., due to vibrations, dirt, non homogeneous material, etc.).

The knowledge of the manufacturing signature can obviously aid the selection of a proper sampling strategy. In [9, 10], geometrical features are modeled via Discrete Fourier Transform (DFT) and models identified are then used to select a sampling strategy, with reference to an economical objective function. However, these approaches focus mainly on dimensional inspection of geometric features (e.g., the diameter of cylindrical components) and the specific model adopted, namely DFT, induces a uniformly spaced sampling strategy.

A true attempt to link the manufacturing signature to the selection of the sampling strategy has been proposed in [11, 12]. This method, called the "Extended Zone" (EZ), is a two-step approach: in the first step, a parametric model of the signature is identified and estimated by using linear regression. In the second step, sampling point are selected in order to minimize the variance of the regression model coefficients.

In this paper we will further investigate the advantages obtainable by selecting a sampling strategy which is based on information related to the manufacturing signature. In particular, we will explore the use of a technique adopted to select variables in multivariate statistical analysis. This technique, known as Principal Component Variables (PCV) selection will be adapted to the specific problem at hand (i.e., defining a sampling strategy of geometrical tolerances) and described in Section 2.1. In order to define the sampling strategy, a second and new approach will be proposed in this paper. This approach will be based on using the signature to identify the points that mainly influence the geometrical tolerance and will be referred to as "Extreme Points Selection" (EPS). This approach will be presented in Section 2.2. Eventually, Section 2.3 will present a comparison of all the different approaches (a polar grid which represents the ISO standard, Hammersley sequence, PCV, EPS) with reference to a real case study related to flatness of surfaces obtained by face turning.

1.1 Case study

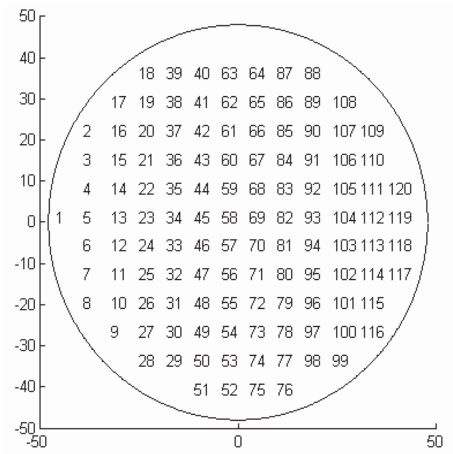

Figure 1: The original sampling strategy for the manufactured disks.

The case study consists of planar surfaces obtained by face turning. The test item are a set of 130 austempering iron disks of 90 mm diameter. 120 points were measured on each disk, as shown in Figure 1: this sampling strategy can be considered as a "rectangular grid". To sample the surface a "Zeiss Prismo VAST HTG" CMM was used; whose main characteristics are:

- Maximum permissible error [13]: $MPE_E = 2 + L/300$ µm
- Maximum permissible error [13]: $MPE_P = 1.8$ µm

For each machined feature, the whole set of 120 points was used to estimate the flatness tolerance by using a Minimum Zone (MZ) algorithm; the standard uncertainty for this reference value is about 0.7 µm (according to previous experimental studies on similar features).

Figure 2 shows the surface obtained by averaging the 130 measured surfaces. The presence of a signature is quite evident: the surface, that nominally should be a flat plane, is concave. Some reasons can be suggested for this behavior: the axis of the lathe used to machine the surface were perhaps not perpendicular, or some strange inflection of the tool may have happened.

2 Sampling Strategy for Flatness: the Standard

The international standard ISO 12781 [3] defines requirements for a sampling strategy that is intended to be used for estimating flatness. Because

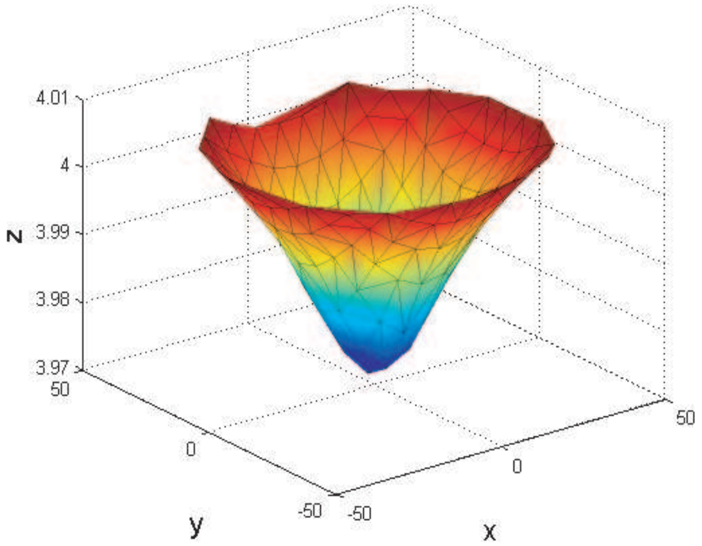

Figure 2: The average surface observed in the real case study (all the dimensions shown are in [mm]).

harmonic content of the surface is considered to be the main contributor to flatness error, the standard proposes some sampling strategy which are constituted of densely sampled profiles, giving rise to polar grid, rectangular grid, union jack, triangular grid and parallel profiles extraction. Depending on the particular geometry, a point extraction strategy is also suggested (for instance, disks should be inspected by means of a polar grid extraction strategy).

Unfortunately, profile measurement can not be easily performed when common CMM with touch trigger probe are used. Moreover, sampling thousands of points with a touch trigger CMM would require a very long time, so sample size has to be reduced. Given the sample size, in order to compare different measuring strategies with the one assumed by the standard, a polar-like positioning of measurement points has been considered.

2.1 A PCA-based selection: the Principal Component variables (PCV)

Principal Components Analysis (PCA) is one of the best known methods to reduce the dimensionality of a multivariate data set, while preserving as much as possible of the variability present in the data set. An exhaustive description of PCA can be found in standard text as [14]. When dealing with geometric tolerances, it has been shown that PCA can aid determining the systematic pattern characterizing a set of machined items, i.e., the manufacturing signature [15]. In this case, the points sampled on a given profile or surface can be modeled as a multivariate random vector of size p. Therefore the p components of the random vector (or p variables) are related to the p location at which the data are collected on profiles and/or surfaces. In particular, with reference to flatness specification, the p-variate random vector shows the values observed along the z direction observed at p different locations in the $x - y$ plane (shown in Figure 2) .

Based on PCA, several variable selection methods have been proposed in the literature [16, 17, 18]. The main idea behind these selection procedures is to retain just few variables (i.e., few points) among the initial set of p variables (i.e., an initial set of points sampled with a dense inspection strategy) while retaining most of the variability observed in this initial set. Among the different approaches for Principal Component variable selection, it has been shown that the *RM* method is able to achieve effective results at low computational costs [19] and this is why we will refer to this approach in this paper. The *RM* method [19, 20] is based on an optimality index aimed at selecting a subset of k variables (i.e., a subset of locations) which preserving most of the information contained in the whole set of p original variables (locations). In particular, the *RM* method is based on computing a similarity index between the subspace spanned by the original set of p variables and the one spanned by the selected subset of dimension k. This similarity index is based on the definition of "angle" between two matrices **A** and **B**, which is defined by the angle cosine, referred to as "matrix correlation" between matrices **A** and **B** and given by:

$$cos\left(\mathbf{A},\mathbf{B}\right) = \frac{< \mathbf{A},\mathbf{B} >}{||\mathbf{A}|| \cdot ||\mathbf{B}||} \tag{1}$$

where $< \mathbf{A},\mathbf{B} >= tr(\mathbf{A}^{\mathrm{T}}\mathbf{B})$ is the inner product of two matrices while $||\mathbf{A}|| = \sqrt{tr(\mathbf{A}^{\mathrm{T}}\mathbf{A})}$ is the norm [20].

With reference to a set of geometric features inspected at p locations, let $\mathbf{X}$ represent the $n \times p$ matrix containing, in the i^{th} row, the set of p data observed on the i^{th} item $(i = 1, ..., n)$. Assume that the matrix $\mathbf{X}$ has been column-centered, as usually done in PCA, i.e., assume that the average path observed over the p locations has been subtracted by the original data. Let $\mathbf{K}$ denote the matrix obtained by selecting a subset K of k columns of $\mathbf{X}$. The *RM* indicator is defined as the cosine of the angle between the original data matrix $\mathbf{X}$ and the matrix whose columns are obtained by regressing each of the p original columns on K (i.e., orthogonally projecting the original columns in the subspace $\mathbf{K}$ of $\mathbb{R}^n$ spanned by K)

$$RM = \cos\left(\mathbf{X}, \mathbf{P_K X}\right) = \sqrt{\frac{\mathrm{tr}\left(\mathbf{X}^{\mathrm{T}}\mathbf{P_K X}\right)}{\mathrm{tr}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)}} \tag{2}$$

where $\mathbf{P_K}$ is the matrix of orthogonal projections on $\mathbf{K}$. A different meaning of RM can be seen introducing $(r_m)_i$ as the multiple correlation coefficient between the i^{th} PC (computed on the original data matrix $\mathbf{X}$) and the k variables in K:

$$RM = \sqrt{\frac{\sum_{i=1}^{p} \lambda_i \left(r_m\right)_i^2}{\sum_{i=1}^{p} \lambda_i}} \tag{3}$$

where λ_i is the eigenvalue of the covariance matrix of $\mathbf{X}$ associated to the i^{th} principal component (note that the eigenvalue λ_i, $\forall i$, does not depend on the choice of K). In this case RM can be interpreted as a weighted average of the squares of the multiple correlations between the PCs and the k selected variables, where the weights $\lambda_i' s$ are the PCs variances.

An exact identification of the best subset of k variables would require an exhaustive enumeration of all the possible subset of size k that can be found starting from the original set of p variables. Unfortunately, this exhaustive enumeration requires a prohibitive computational time. Therefore, different heuristic algorithms have been proposed [20]. In particular, the present paper will investigate performance obtainable by using a simulated annealing algorithm for subset selection (a freely available software can be found among packages running with “R” [21]).

This sampling strategy will be referred to as “Principal Component Variables” (PCV) in the following.

2.2 Extreme Points Selection (EPS)

When estimating form error via the Minimum Zone algorithm (MZ), only few points held all the information needed, namely the “essential subset” [22]. For instance, when the form error concerns circularity, the MZ algorithm requires one to compute the radial distance between two concentric circles containing among them all the observed data and having the least radial separation. In this example, only four points define the inscribed and the circumscribed circles required to characterize the MZ estimate of the form error. These four points constitute the essential subset (ES).

If a manufacturing signature is present, points in the ES will tend to appear approximately in the same position. In fact, the manufacturing signature is defined as the systematic pattern characterizing all the items produced with a specific process. When all the machined items have the same systematic pattern, it is likely that the ES will be constituted by a specific subset of points which are always in the same positions.

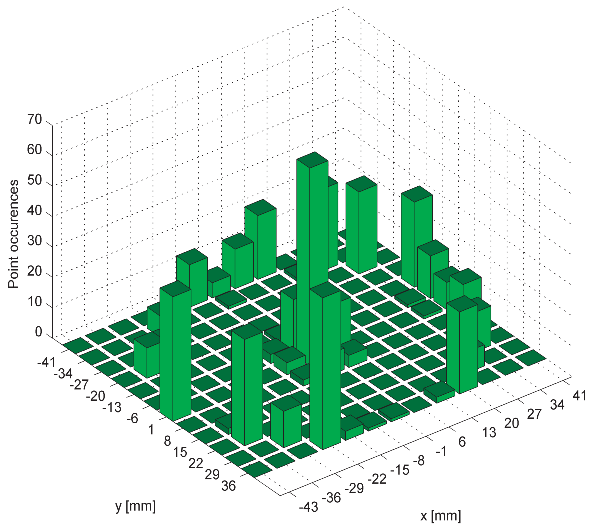

Figure 3: Extreme point histogram.

Therefore, the proposed procedure consists in identifying these subset of relevant points by using a (dense) sampling strategy on a first set of machine items. Assuming n items have been sampled in p locations, define:

$$e_{ij} = \begin{cases} 1 & \text{if the sampling point corresponding to the } j^{\text{th}} \text{ location belongs} \\ & \text{to the essential subsets (i.e., it an extreme point) for the } i^{\text{th}} \text{ part;} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$. Define

$$E_j = \sum_{i=1}^{n} e_{i,j} \qquad (j = 1, 2, \ldots, p) \quad , \tag{5}$$

then E_j is a counter for the number of times that the j^{th} location resulted an extreme point (i.e., affected the form error estimate). Eventually, we can simply rank E_j's, so that $E_{[1]} \leq E_{[2]} \leq \cdots \leq E_{[p]}$, and given $k < p$ sampling locations have to be selected, locations corresponding to $E_{[1]}, E_{[2]}, \ldots, E_{[k]}$ will be selected. Clearly, this point selection methodology corresponds to choosing the points which correspond to the highest values in an EPS histogram, i.e., an histogram showing the occurrences of extreme points at each location.

Figure 3 shows the EPS histogram in the case of our flatness case study, where the ordinate represents the number of times each location corresponded to an "extreme" point (i. e., the point influenced the form error estimate) in the whole set of $n = 130$ disks in the reference sample.

2.3 Experimental results

With reference to the real case study, five measuring strategies are compared. The first two strategies are "a-priori" strategies, namely the Hammersley (H) and the polar grid (which represents the ISO-like procedure for flatness error estimation and will be labeled ISO in the following). The remaining three procedures are all signature-based points extraction strategies: the EZ procedure [11, 12] and the two procedures (PCV and EPS) proposed in this paper. Let i denote the i^{th} flatness surface obtained by face turning an austempering iron disk ($i = 1, \ldots, n = 130$). For each surface, a MZ flatness form error estimate was firstly computed by using the whole set of $p = 120$ points represented in Figure 1. For each item $i \quad (i = 1, \ldots, 130)$ this value of the form error is taken as reference. As a matter of fact, the estimate of the MZ form error for the same item will be different from this reference value when just a subset of $k < p = 120$ locations are considered in the form error computation. In particular, the estimated error associated to this subset of points is usually lower than the reference one. In fact, when only few points out of the whole set of data are selected, it is unlikely that these points are just the extreme ones. This is why the performance indicator of the selection procedure is computed as the difference between the reference form error (computed by using the whole set of data) and the error computed by using just the specific subset chosen by the sampling strategy (H, ISO, EZ, PCV, EPS) for each machined item. Obviously the best procedure will be the one which allows one to achieve the minimum performance indicator.

Figure 5 shows the comparison of errors obtained by comparing the polar strategy (ISO) and Hammersley procedure (H) for different values of the subsample size $k = [9, 13, 19, 25, 33]$. For each value k, the different sampling strategies are applied to all the flat surfaces ($i = 1, 2, \dot{;} n = 130$) under study and a final 95% confidence interval of the median performance indicator is computed (as shown in). Here the median is used because the computed

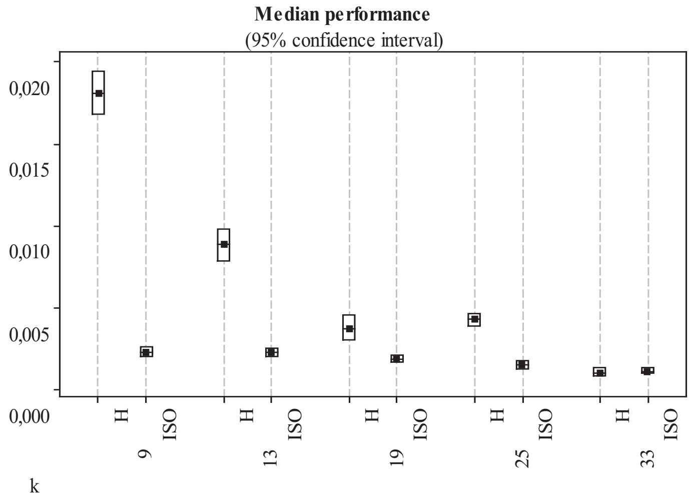

Figure 4: Performance comparison: polar (ISO) vs. Hammersley (H) strategies as a function of the number of selected points k.

performance indicators are not normally distributed. As shown in Figure 5, the Hammersley sequence induces always a median indicator which is greater than the one obtained by using a polar-grid (ISO) approach. Therefore, the ISO procedure should be preferred to the Hammersley's one in the specific case under study.

Figure2.3 shows the comparison of this ISO (polar) strategy with the signature-based sampling methods (EZ, PCV, EPS). Here again, the figure reports the 95% confidence intervals on the median indicator obtained by using the different approaches for subset size ranging from 9 to 33. Figure2.3 shows that, with the exception of samples of size 9 and 13, all the signature-based strategies perform always better than the ISO-based approach. In particular, the EPS approach proposed in this paper outperforms all the other methods, despite of the actual value of the sample size $k = [9, 13, 19, 25, 33]$.

Eventually, Figure 6 shows the position of the points chosen by the different signature-based sampling strategies (for the sample size $k = 33$). As clear from this pictures, all these sampling strategies tend to concentrate sampling points at the center and on the borders of the machined surface, according with the signature pattern shown in Figure 2. This well-behaved pattern is particularly clear for the EPS algorithm, which resulted to be the best approach.

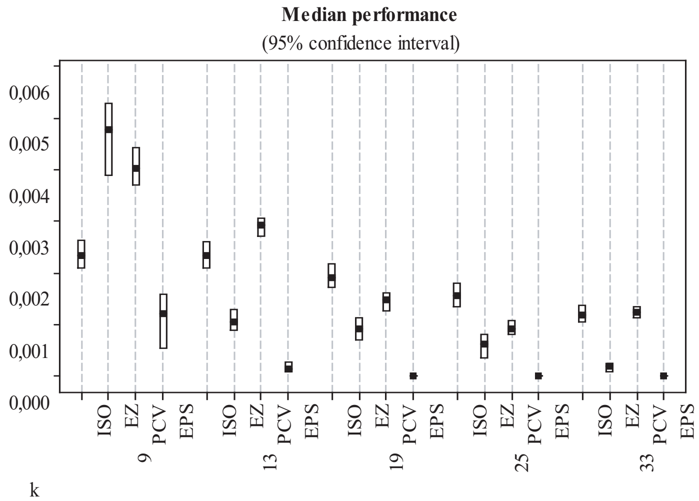

Figure 5: Performance comparison: polar (ISO) vs. signature-based strategies (EZ, PCV, EPS) as a function of the number of selected points k.

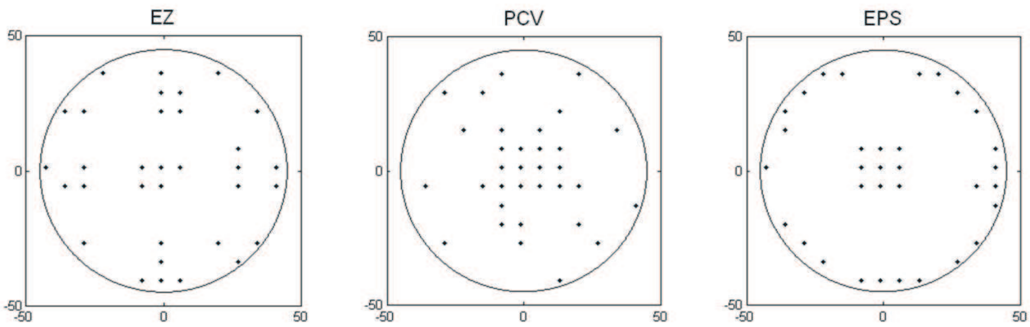

Figure 6: Pattern of the selected locations with different sampling strategies.

3 Conclusions

In the present paper, we showed that sampling strategies which account for the information coming out from the manufacturing process can be very effective for inspecting items when geometrical form tolerance are of interest. In particular, information coming out from the manufacturing process is summarized in what we called the manufacturing signature, i.e. the systematic pattern which characterizes all the features machined with a given process. A new sampling strategy proposed in this paper (the Extreme Points Selection strategy) was shown to outperform different approaches proposed in the literature and in the ISO standard. Since measurement costs are proportional to measurement time, techniques aimed at obtaining small systematic errors in geometrical tolerance estimates while reducing the number of sampled points can be effectively used in industrial practice. The main limitation behind the application of these methods is that if the signature changes without any advice (for instance because of a sudden change in the machine tool) these techniques can become “dangerous”, because they can suggest the measurement of points which are no more relevant to describe the systematic pattern of machined items. Therefore a signature monitoring technique should always to be used when applying one of these methods [23, 2]. These approaches basically consist of a first step in which the signature model for an in-control process is identified. Coefficients characterizing the in-control model are then monitored using multivariate control charting. Therefore, any instability of the manufacturing signature results in unusual values of the model coefficients which, in turn, generate an out of control alarm.

Acknowledgments

This work was carried out with the funding of the Italian Ministry of University and Research.

References

[1] Colosimo, B. M., Mammarella, F. and Petro, S. (2007): Quality Control Of Manufactured Surfaces. In *Proceedings of the IX Intl. Workshop on Intelligent Statistical Quality Control.*

[2] 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, 1–18.

[3] International Organization for Standardization, Geneva, Switzerland (2003): *ISO/TS 12781-2: Geometrical Product Specifications (GPS) - Flatness - Part 2: Specification operators.* 1st edition.

[4] Kim, W. S. and Raman, S. (2000): On the selection of flatness measurement points in coordinate measuring machine inspection. International Journal of Machine Tools and Manufacture 40, 427–443.

[5] Edgeworth, R. and Wilhelm, R. G. (1999): Adaptive sampling for coordinate metrology. Precision Engineering 23, 144–154.

[6] Badar, M. A., Raman, S. and Pulat, P. (2003): Intelligent Search-Based Selection of Sample Points for Straightness and Flatness Estimation. Journal of Manufacturing Science and Engineering 125, 263.

[7] Rossi, A. (2001): A form of deviation-based method for coordinate measuring machine sampling optimization in an assessment of roundness. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 215, 1505–1518.

[8] Badar, M. A., Raman, S. and Pulat, P. S. (2005): Experimental verification of manufacturing error pattern and its utilization in form tolerance sampling. International Journal of Machine Tools and Manufacture 45, 63–73.

[9] Capello, E. and Semeraro, Q. (2001): The harmonic fitting method for the assessment of the substitute geometry estimate error. Part I: 2D and 3D theory. International Journal of Machine Tools and Manufacture 41, 1071–1102.

[10] Capello, E. and Semeraro, Q. (2001): The harmonic fitting method for the assessment of the substitute geometry estimate error. Part II: statistical approach, machining process analysis and inspection plan optimisation. International Journal of Machine Tools and Manufacture 41, 1103–1129.

[11] 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, 273–292.

[12] 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, 105–121.

[13] International Organization for Standardization, Geneva, Switzerland (2001): *ISO/TS 10360-2: Geometrical Product Specifications (GPS) – Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 2: CMMs used for measuring size.* 2^{nd} edition.

[14] Jackson, J. E. (1991): *A User's Guide to Principal Components.* Wiley-InterScience, New York.

[15] Colosimo, B. M. and Pacella, M. (2007): On the use of Principal Component Analysis to Identify Systematic Patterns in Geometric Profiles. Quality & Reliability Engineering International 23, 707–725.

[16] Beale, E. M. L., Kendall, M. G. and Mann, D. W. (1967): The Discarding of Variables in Multivariate Analysis. Biometrika 54, 357–366.

[17] Jolliffe, I. T. (1972): Discarding Variables in a Principal Component Analysis. I: Artificial Data. Applied Statistics 21, 160–173.

[18] McCabe, G. P. (1984): Principal Variables. Technometrics 26, 137–144.

[19] Cadima, J. F. C. L. and Jolliffe, I. T. (2001): Variable Selection and the Interpretation of Principal Subspaces. Journal of Agricultural, Biological, and Environmental Statistics 6, 62–79.

[20] Cadima, J., Cerdeira, J. O. and Minhoto, M. (2004): Computational aspects of algorithms for variable selection in the context of principal components. Computational Statistics and Data Analysis 47, 225–236.

[21] Cerdeira, J. O., Silva, P. D. and Cadima, M., J.and Minhoto (2007): *subselect: Selecting variable subsets.* R package version 0.9-9992.

[22] Anthony, G. T., Anthony, H. M., Bittner, B., Butler, B. P., Cox, R., M. G.and Drieschner, Elligsen, R., Forbes, A. B., Gross, H., Hannaby, S. A., Harris, P. M. and Kok, J. (1996): Reference software for finding Chebyshev best-fit geometric elements. Precision Engineering 19, 28–36.

[23] 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, 707–725.

Bianca Maria Colosimo
Dipartimento di Meccanica
Politecnico di Milano
Via la Masa 1
20156 Milano, Italy
e-mail: biancamaria.colosimo@polimi.it
Ester Gutierrez Moya
University of Seville
C/ Enrquez de Ribera 1
41092 Sevilla, Spain
e-mail: egm@esi.us.es
Giovanni Moroni
Dipartimento di Meccanica
Politecnico di Milano
Via la Masa 1
20156 Milano, Italy
e-mail: giovanni.moroni@mecc.polimi.it
Stefano Petrò
Dipartimento di Meccanica
Politecnico di Milano
Via la Masa 1
20156 Milano, Italy
e-mail: stefano.petro@polimi.it