Applied Mathematics and Computation 242 (2014) 562–568

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A statistical method for geometry inspection from point clouds

Francisco de Asís López [a], Celestino Ordóñez [a,*], Javier Roca-Pardiñas [b], Silverio García-Cortés [a]

[a] Department of Mining Exploitation and Prospecting, University of Oviedo, 33600 Oviedo, Spain
[b] Department of Statistics and Operations Research, University of Vigo, 36208 Vigo, Spain

ARTICLE INFO

Keywords:
Bootstrap
Kernel smoothing
Inspection
Point cloud

ABSTRACT

This paper introduces a statistical methodology for geometry inspection from point clouds obtained with a laser scanner or other measuring systems. The null hypothesis of interest is that the real surface of an object fits the theoretical shape and dimensions of the object. An algorithm based on bivariate kernel smoothers is used to nonparametrically estimate the surface of the object and bootstrap-based procedures are proposed for testing the null hypothesis. In order to validate the methodology a simulated study was conducted. Finally, the proposed methodology was applied to the inspection of a parabolic dish antenna.

1. Introduction

Inspection is the process of determining if a product deviates from a given set of specifications [1]. The term surface quality inspection is normally applied to inspection works devoted to find small local defects on a freeform surface. On the other hand, geometry inspection is used to make reference to an inspection process whose goal is to find differences in the shape [2,3]. Geometry inspection can be carried out by registering a point cloud of the object and comparing it with the design model to determine whether the geometry of that object is out of tolerance. This process is called point cloud inspection. The point cloud can be obtained by means of laser scanning systems, photogrammetry, structured lighting systems or coordinate-measuring machines, among others [4,5].

The inspection process has attracted increasing interest among engineers and, consequently, there is a growing amount of literature on this subject, as well as commercial software for point cloud processing that incorporates a tool for inspection. Some applications of quality inspection can be seen in [6–8] or [9], among many others.

The comparison between the measurement data and the design model is achieved by calculating the distance between them. Previously, a common coordinate system for both, the measurement data and the design models, had to be calculated by means of registration- The ICP (iterative closest point) algorithm is normally used for registration [10,11]. Principally, two approaches have been proposed to calculate such distance. One consists of calculating the distance between measurement points and the nearest point on the model design [12,13]. The other approach evaluates the distance between vertices of the design model and their closest point on a reconstructed model obtained from the measured data [14,15]. Different methods to construct 3D models from point clouds have been proposed: parametric models such as B-Splines or NURBS [16,17]; non-parametric methods such as radial basis functions [18] or kernel smoothing techniques [19]; and triangular meshes [20,21]. However, it is very unusual that these approaches provide information in statistical terms about the significance level of the discrepancy between the measured data and the model, although the coordinates of the measured points are subjected to uncertainty [22].

* Corresponding author.
E-mail address: ordonezcelestino@uniovi.es (C. Ordóñez).

http://dx.doi.org/10.1016/j.amc.2014.05.130
0096-3003/

In this work we propose a methodology to determine, with a specific level of significance, if a point cloud of an object fits its theoretical shape and dimensions. The paper is structured as follows: first, the proposed methodology is explained; second, a simulated experiment is carried out in order to test the behavior of the method; then the method is applied to a real inspection problem; finally, the results and the conclusions of our research are reported.

2. Statistical methodology

Let us consider (X, Y, Z) the spatial co-ordinates of each point on the object surface and assume that the third co-ordinate Z can be obtained from (X, Y) using an unknown function $m(X, Y)$, which represents a kind of 2.5D *smooth* surface, so that $Z = m(X, Y)$. The aim of the statistical procedure proposed here is to determine if m follows a specific parametric shape m_0. This problem can considered as a hypothesis test where the null hypothesis is

$$H_0 : m(x, y) = m_0(x, y) \quad \text{for all } (x, y) \tag{1}$$

against the alternative hypothesis that m is an unknown function different to m_0.

2.1. Nonparametric estimation

Given that, in the alternative hypothesis, m is not known, we need to estimate this function using the point cloud $\{P_1^\bullet, \ldots, P_n^\bullet\}$ with $P_i^\bullet = (X_i^\bullet, Y_i^\bullet, Z_i^\bullet)$ being n the number of observations. Each of these points $P_i^\bullet$ can be understood as a measure of the real point $P_i = (X_i, Y_i, Z_i)$ of the object surface, so that

$$P_i^\bullet = P_i + \epsilon_i \tag{2}$$

where $\epsilon_i = (\varepsilon_i^X, \varepsilon_i^Y, \varepsilon_i^Z)$ represents the measurement error on the ith point. Moreover, as we have no information regarding the distribution of the error term, it was assumed that error components $\varepsilon_i^X, \varepsilon_i^Y$ and ε_i^Z are random points located on an sphere.

For any given point (x_0, y_0) , a smoothed version of the orthogonal regression (or Deming regression [23]) is proposed. In this type of regression the estimated surface is obtained from the observed sample points $\{P_1^\bullet, \ldots, P_n^\bullet\}$ by minimizing the perpendicular distances of those points to the surface. The proposed estimators are based on the fact that surface $(x, y, m(x, y))$ can be approximated by the plane

$$ax + by + cz + d = 0$$

in values (x, y) near (x_0, y_0). Thus, the estimated coefficients $(\hat{a}, \hat{b}, \hat{c})$ can be obtained as the third principal component of the locally principal component analysis [24,25]. The steps of the proposed procedure are as follows:

- For each point $i = 1, \ldots, n$ a weighting function W_i is computed

$$W_i = W_h(X_i^\bullet - x_0, Y_i^\bullet - y_0) = \exp\left\{-\frac{(X_i^\bullet - x_0)^2 + (Y_i^\bullet - y_0)^2}{h}\right\} \tag{3}$$

- Compute the sample 3×3 covariance matrix $\hat{\Sigma}$ obtained from the point cloud $(P_1^\bullet, \ldots, P_n^\bullet)$ weighted by $(W_1, \ldots, W_n)$, and obtain $(\hat{a}, \hat{b}, \hat{c})$ as the third principal component of the PCA using $\hat{\Sigma}$ as the input covariance matrix
- Calculate a plane normal to this vector and passing through $(\bar{X}^\bullet, \bar{Y}^\bullet, \bar{Z}^\bullet)$:

$$\hat{a}(x - \bar{X}^\bullet) + \hat{b}(y - \bar{Y}^\bullet) + \hat{c}(z - \bar{Z}^\bullet) = 0$$

Finally, evaluating that plane in (x_0, y_0) and solving in z, an estimation of $m(x_0, y_0)$ is obtained:

$$\hat{m}(x_0, y_0) = \bar{Z}^\bullet - \frac{\hat{a}(x_0 - \bar{X}^\bullet) + \hat{b}(y_0 - \bar{Y}^\bullet)}{\hat{c}}$$

Note that, the kernel weight W_i in (3) depends on the Euclidian distance between (x_0, y_0) and $(X_i^\bullet, Y_i^\bullet)$ and, in addition, contains a smoothing parameter h. Detailed information about kernel smoothing can be found in [26–28]. Moreover, although our focus is on kernel smoothers, there are other types of procedures that allow for non-parametric estimations of the model. In [29] the authors investigated alternative methods based on penalized splines, and [30] used thin plate regression splines.

It is well known that the nonparametric estimates $\hat{m}(x, y)$ heavily depend on the bandwidth h used in the kernel weights W_h in (3). The bandwidth is a trade-off between the bias and the variance of the resulting estimates. According to the definition of W_h, the observations $(X_i^\bullet, Y_i^\bullet, Z_i^\bullet)$ with the first two coordinates $(X_i^\bullet, Y_i^\bullet)$ closed to (x_0, y_0) have more influence on the estimate $\hat{m}(x_0, y_0)$ than those farther away. The amount of relative influence is controlled by the bandwidth h. On the one hand, if h is small the resulting estimate $\hat{m}(x_0, y_0)$ heavily depends on those observations that are closest to (x_0, y_0) and tends to yield a more wiggly estimate; and if h becomes closer to zero the estimate tends to adjust too much to the data and, as a consequence, have a very high variance. On the other hand, if the bandwidth is too large the estimated surface will be a plane and will not adjust. That shows the significance of arranging a tool for the automatic choice of the most appropriate