Statistica Sinica **8**(1998), 135-149

GEOMETRIC QUALITY INSPECTION

Gemai Chen and Jiahua Chen

University of Regina and University of Waterloo

Abstract: In modern industrial settings, more and more new products are created in the form of computer images via Computer Aided Design (CAD). On the other hand, CAD files are sometimes created to catalog some well known products whose technical drawings may or may not exist. In either case, the geometric features of the products can be measured with a Coordinate Measuring Machine (CMM). In this paper, we discuss how to judge the geometric quality of a product when its CMM data and CAD file are available. We first introduce a probabilistic model to represent the relationship between the CMM data and the CAD file. Then we define a measure of geometric quality and use it to connect the geometric quality inspection problem to a statistical hypothesis testing problem. The measure we propose is easy to calculate and sensitive in detecting manufacture errors. An application to automobile part inspection is included.

Key words and phrases: Computer assisted design, coordinate measuring machine, hypothesis testing, probabilistic model.

1. Introduction

Manufactured products are all designed to have certain geometric shapes. However, due to various uncontrollable factors, there are always geometric departures between the products and the design specifications. One important issue here is how to judge and control the geometric quality of a product. In general, the geometric quality of a product is judged by how close the product is in compliance with the geometric tolerances specified in the design. This involves verifying (1) the presence of each expected feature, (2) the dimensions of these features (e.g. radius and length of a cylinder), and (3) the feature interrelationships (e.g. distances between centers of gravity and angles between plane normals). (See Marshall and Martin (1992)). For these purposes, procedures of dimensioning and tolerancing have been developed. In these procedures, basic concepts, such as roundness and straightness, are well defined, and detailed steps, such as the ones to specify the position of a hole, are clearly described. The tolerance of a dimension is often defined as an ideal value plus or minus an error, and the errors for all dimensions of a product are specified in such a way that can guarantee smooth assembly in the future, while the cost of meeting these error tolerances is kept as low as possible. (See Liggett (1970), Spotts (1983), Foster (1986), Madsen (1988), Turner (1990), and Guiford, Sethi and Turner (1992).)

Traditionally, judgement on the geometric quality of a product is based on the degree of agreement between the measurements taken with gauges, calipers, etc. on particular features of the product and the tolerance limits specified on the technical drawings. A product is rejected if it does not comply with its specified tolerance limit. If several features of a product need to be checked frequently and/or by different manufacture units, a high quality mould is usually made and the product can be placed against the mould to see whether or not the product fits.

To judge the geometric quality of products with sophisticated shapes, more advanced measurement methods are needed. We discuss in this paper the problem of how to judge geometric quality when the coordinates of some points on the surface of a product are available. There are different ways of obtaining coordinate readings. For example, coordinates can be obtained by using a Coordinate Measurement Machine (CMM). CMMs can be used to obtain measurements with high accuracy and have been used widely in industry (Cox and Peggs (1986), Tannock (1992)). We use CMM data to fix ideas in the following discussion, although the data are not required to come from a CMM. We note that with more and more computer technology introduced into industry, more and more products are now first created in the form of computer images via Computer Aided Design (CAD) (Rooney and Steadman (1987)). On the other hand, and to a less extent, CAD files are needed to catalog some well known products whose technical drawings may not exist. In the first case, the specifications of the product are stored as a CAD file which represents the ideal product, while in the second case, the CAD file of a product can serve the purpose of re-creating the product (Stuetzle, (1994)). Using CMM and CAD data to inspect the geometric quality of products has been discussed in the literature and widely applied in practice, especially in the first case (Cox and Peggs (1986), Marshall and Martin (1992), Tannock (1992), Besl and McKay (1992), Goch and Tschudi (1992)). However, most discussions are devoted to checking whether or not a product meets certain tolerance limit in a deterministic manner. Recently, there have been some discussions on probabilistic and statistical approaches (Brou (1984), Horn (1984, 1986), Berman and Culpin (1986), Yang, Li and Li (1988), Gunasena, Lehtihet and Ham (1989), Caskey et al. (1990), Kurfess and Banks (1990), Menq et al. (1990), Chapman and Kim (1992), Chen and Chen (1992), Hulting (1992), (1993), Dowling et al. (1993), Chapman, Chen and Kim (1995)), and it seems that more and more researchers are getting involved.

In this paper, we introduce a probabilistic model in which the features of a product as measured in the CMM file are related to the features in the CAD file plus a random departure that represents the sum of manufacture errors and measurement errors. A product has good geometric quality if the scale of this random departure is small, that is, we judge the geometric quality of a product from a statistical point of view by estimating the scale of random departures. More specifically, we define a measure of geometric quality and use it to connect the geometric quality inspection problem to a statistical hypothesis testing problem. The measure we propose is easy to calculate and sensitive in detecting manufacture errors. Moreover, our method is applicable not only to planar features, but to more sophisticated geometric features such as curves and surfaces as well.

The paper is organized as follows. The problem of geometric quality inspection will be described more precisely in Section 2. The proposed measure of geometric quality and its properties are also discussed there. Simulation results will be presented in Section 3 to validate our proposed procedure, and an application to automobile part inspection will be presented in Section 4. Some remarks are given in Section 5 and short proofs in the appendix.

2. Problem Formulation

2.1. Description of the problem

Although manufactured products in practice may have very complicated shapes, we will use a simplex to explain our idea. Assume that the product has the shape as shown in Figure 1. The surface S of the simplex consists of four planes, denoted as $S = S_1 \cup S_2 \cup S_3 \cup S_4$. In general, we assume that the surface of a product consists of m sub-surfaces and denote them by $S_1, \ldots, S_m$.

Without loss of generality, suppose that the geometric features of a product are measured with a CMM. To fix the idea, we denote the surface of the manufactured simplex as $\tilde{S} = \tilde{S}_1 \cup \tilde{S}_2 \cup \tilde{S}_3 \cup \tilde{S}_4$ and measure, say, 6 points (3×1 vectors) on each of $\tilde{S}_i$, $i = 1, \ldots, 4$. (In practice, we take 3–5 points for a line, 5–8 points for a plane, 4–8 points for a circle, etc., based on experience.) In general, for m sub-surfaces, we denote the jth measurement on $\tilde{S}_i$ by y_{ij}, where $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, and n_i depends on the complexity of $\tilde{S}_i$. Note that for each y_{ij} we usually cannot identify the corresponding points on S_i. However, it is reasonable to assume that

$$R_o y_{ij} + T_o = x_{ij} + e_{ij}, \tag{2.1}$$

where x_{ij} are the points on S_i that correspond to y_{ij}, R_o is a 3×3 rotation matrix, T_o is a 3×1 translation vector and $e_{ij} = e'_{ij} + e''_{ij}$, where e'_{ij} are the manufacture errors and e''_{ij} are the measurement errors. When there are no manufacture errors and measurement errors, R_o and T_o together will transform the coordinate system of the CMM data to that of the CAD file.

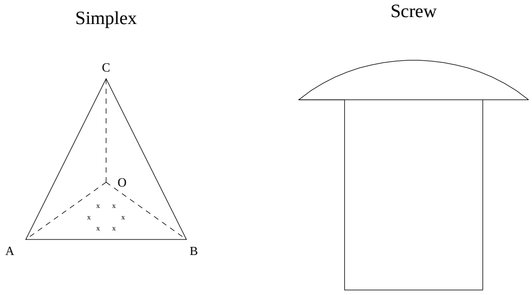

Figure 1. A simplex product and a screw product. Six points are measured on each side of the simplex, as indicated by the crosses marked on the ABO plane.

We have made an implicit assumption. That is, we can identify the subsurfaces both in the CAD file and from the manufactured product. This assumption is reasonable and is not restrictive in practice. To simplify the problem, we further assume that e'_{ij} and e''_{ij} are independent of each other and are normally distributed. In practice, it is not possible to distinguish between e'_{ij} and e''_{ij}; that is why we write $e_{ij} = e'_{ij} + e''_{ij}$ and assume that e_{ij} are independent and identically distributed as trivariate normal $N_3(0, \sigma^2 I_3)$, where I_3 is the 3×3 identity matrix and σ^2 is a positive constant.

We do not discard a product simply because it has a tiny manufacture error. In fact, the real issue in practice is to decide how much manufacture error is acceptable in the presence of measurement error. This leads us to judge the geometric quality of a product by the size of σ^2. That is, whether a product is of satisfactory quality depends on whether $\sigma^2 \leq \sigma_o^2$, where σ_o^2 is determined by the practical requirement on the manufacture precision and the precision of the CMM. In other words, we interpret the geometric quality inspection problem as a statistical hypothesis testing problem; the hypothesis to be tested is H_o: $\sigma^2 \leq \sigma_o^2$. It is noted that the hypothesis testing approach has been used in Menq et al. (1990) and Chapman and Kim (1992).

From model (2.1), one may try to estimate T_o and R_o by using the ordinary linear regression approach. This does not work directly because x_{ij} are usually unidentifiable points on S_i. Instead, we suggest fitting the model by minimizing

$$Q(T, R) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \rho^2(Ry_{ij} + T, S_i) \tag{2.2}$$

over all possible $T = (T_1, T_2, T_3)^T$ $(-\infty < T_k < \infty,\ k = 1, 2, 3)$ and 3 by 3 rotation matrices R, where $\rho(y, S_i) = \min_{x \in S_i} d(y, x)$ and d is a distance which we choose as the Euclidean distance. When all S_i are planes, the expression for $Q(T,R)$ can be simplified as

$$Q(T,R)=\sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle Ry_{ij}+T-a_i,v_i\rangle^2, \tag{2.3}$$

where $\langle\cdot,\cdot\rangle$ denotes the usual inner product in Euclidean space, a_i is an arbitrary point on S_i, and v_i is a unit normal vector to S_i.

2.2. Some properties when all S_i are planes

Let $\hat{T}$ and $\hat{R}$ be a translation and a rotation that produce the global minimum of $Q(T,R)$. Intuitively, $Q(\hat{T},\hat{R})/\sigma^2$ should have a chi-square distribution when the errors are normal. However, this is not true in a strict mathematical sense. Nevertheless, we find the following properties. Let $S=\bigcup_{i=1}^m S_i$ be the surface of a product in the CAD file. Assume that each S_i is a plane with a unit normal vector v_i and e_{ij} are i.i.d. $N_3(0,\sigma^2 I_3)$. Then under model (2.1) and definition (2.3), we have

Property 1. Let $\hat{T}$ and $\hat{R}$ be the minimizer of (2.3). If

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle (R-I_3)x_{ij}+T,v_i\rangle^2>0 \tag{2.4}$$

for any $T\neq 0$ and any $R\neq I_3$, then $\hat{T}$ and $\hat{R}$ are strongly consistent for T_o and R_o, respectively, that is, $\hat{T}$ converges almost surely to T_o, and $\hat{R}$ converges almost surely to R_o, as $\sigma\to 0$. Moreover, as $\sigma\to 0$, the limiting distribution of $Q(\hat{T},\hat{R})/\sigma^2$ exists and is the same as the distribution of

$$\inf_{\alpha,\beta,\gamma,T}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle \tilde{R}x_{ij}+\varepsilon_{ij}+T,v_i\rangle^2,$$

where

$$\tilde{R}=\begin{pmatrix}0&-\gamma&\beta\\ \gamma&0&-\alpha\\ -\beta&\alpha&0\end{pmatrix},$$

α, β and γ are real numbers, and ε_{ij} are $N_3(0,I_3)$.

Property 2. The following is always true,

$$Q(\hat{T},\hat{R})/\sigma^2\leq Q(T_o,R_o)/\sigma^2\sim\chi^2_{\nu_1}\quad \Big(\nu_1=\sum_{i=1}^{m}n_i\Big).$$

Property 3. If the normal directions v_i of S_i do not fall into a single plane, then for any real $q > 0$,

$$P\left\{Q(\hat{T},\hat{R})/\sigma^2 \le q\right\} \ge P\left\{\chi^2_{\nu_2} \le q\right\} \qquad \Big(\nu_2 = \sum_{i=1}^{m} n_i - 3\Big).$$

The proofs of these properties are given in the appendix.

Comments:

1. Although the distribution of $Q(\hat{T},\hat{R})/\sigma^2$ depends on the value of σ in general, according to Property 1, the dependence becomes weak when σ is small compared with the x_{ij}'s. In our formulation of the problem, CMM precision is high compared with the size of the part; therefore, the above dependence can be ignored. Our simulation confirms this comment.

2. The assumption (2.4) has a natural interpretation: when the part is perfect and the CMM has infinite precision, effectively there should be only one transformation set that matches the two coordinate systems. Also, for shapes such as cubes, there can be more than one transformation to match the ideal product perfectly, but condition (2.4) eliminates this possibility.

Property 1 suggests that when σ^2 is small, we can use $Q(\hat{T},\hat{R})/\sigma^2$ as an approximate pivotal to test the hypothesis H_o as follows: Let F_α denote the upper α-percentile of the distribution of $Q(\hat{T},\hat{R})/\sigma_o^2$. We reject H_o: $\sigma^2 \le \sigma_o^2$ when $Q(\hat{T},\hat{R})/\sigma_o^2 \ge F_\alpha$, where $0 \le \alpha \le 1$ is the predetermined significance level.

Although we do not know the exact size of F_α, Properties 2 and 3 can give a good upper bound for it. Moreover, since T and R together have 6 free parameters, we expect $Q(T,R)/\sigma^2$ to lose 6 degrees of freedom in the chi-square distribution after minimization with respect to T and R. This conjecture is supported by our simulation results in Section 3.

2.3. Extension to non-planar geometric components

When some S_i are not planes but curved surfaces, we expect the properties discussed in Section 2.2 to remain approximately true provided σ_o^2 is small. This is because smooth surfaces are locally flat and the size of σ^2 compared to the size of a product is small in practice (otherwise, we do not need very precise devices such as CMM to detect problems in geometric quality). Since for each point y_{ij} we are calculating the perpendicular distance between the point $Ry_{ij}+T$ and the surface S_i, when σ^2 is small, we are essentially calculating the distance between $Ry_{ij} + T$ and the tangent plane to S_i that passes through point x_{ij}. Therefore, when σ^2 is small, the conditions underlying Properties 1 to 3 are nearly satisfied.

When an S_i is one dimensional, such as a curve, each measurement on S_i will contribute 2 degrees of freedom to the chi-square distribution instead of one degree of freedom. Thus, the total degrees of freedom of an approximate chi-square distribution for the most general case is

$$\nu = \sum_{i=1}^{m} n_i \times \{3 - dim(S_i)\} - 6, \tag{2.5}$$

where $dim(S_i) = 2$ if S_i is a surface, and $dim(S_i) = 1$ if S_i is a curve. However, when a product is invariant with respect to certain transformations, the loss of degrees of freedom will be less than 6. For example, if the product is a sphere, then rotation is unnecessary, and the loss of degrees of freedom should be 3. See also Example 2 in Section 3 where only 2 degrees of freedom are lost with respect to rotation. Obviously, when ν is large, which is often the case, this difference is negligible.

3. Simulation

To illustrate the method developed in Section 2, two simple and artificial examples are included in this section. Example 1 below deals with products that consist of planes only, while Example 2 deals with products that consist of various shapes.

Example 1. This is the example we have discussed. The assumed product consists of four planes. In our simulation, we select six points x_{ij} from each plane of the ideal product and add random trivariate normal noises e_{ij} to give $x_{ij} + e_{ij}$, $i = 1, \ldots, 4$ and $j = 1, \ldots, 6$. We then apply an arbitrary but fixed rotation R_o and an arbitrary but fixed translation T_o to obtain simulated CMM measurements $y_{ij} = R_o^T (x_{ij} + \sigma\varepsilon_{ij} - T_o)$ according to model (2.1).

For a range of σ values ($\sigma = 0.0005$ (0.0005) 0.05), we repeat the simulation 10,000 times for each fixed σ value. The subroutine E04JAF from the NAG library is used to find the minimum of (2.3). It is found that the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$ virtually does not change for the σ values considered. Therefore, we concentrate on one σ value in the following discussion to study the distribution property of $Q(\hat{T}, \hat{R})/\sigma^2$. For $\sigma = 0.05$, we compare $Q(\hat{T}, \hat{R})/0.05^2$ with the chi-square distribution on 18 degrees of freedom, because $\nu = 4 \times 6 \times 1 - 6 = 18$ according to (2.5). The following are some selected upper tail probabilities:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{18,\alpha})$	0.5016	0.3981	0.3027	0.2006	0.0989	0.0516	0.0283	0.0117,

where $P(\chi^2_{18} > \chi^2_{18,\alpha}) = \alpha$, and the second row is the proportion of $Q/0.05^2$ that exceed $\chi^2_{18,\alpha}$. Clearly, χ^2_{18} approximates the distribution of $Q/0.05^2$ very well. (See also Figure 2 (a) for a graphical comparison.) Property 3 in Section 2.2 concludes that χ^2_{21} is an upper bound for Q/σ^2 in this example. This is confirmed from the above simulation.

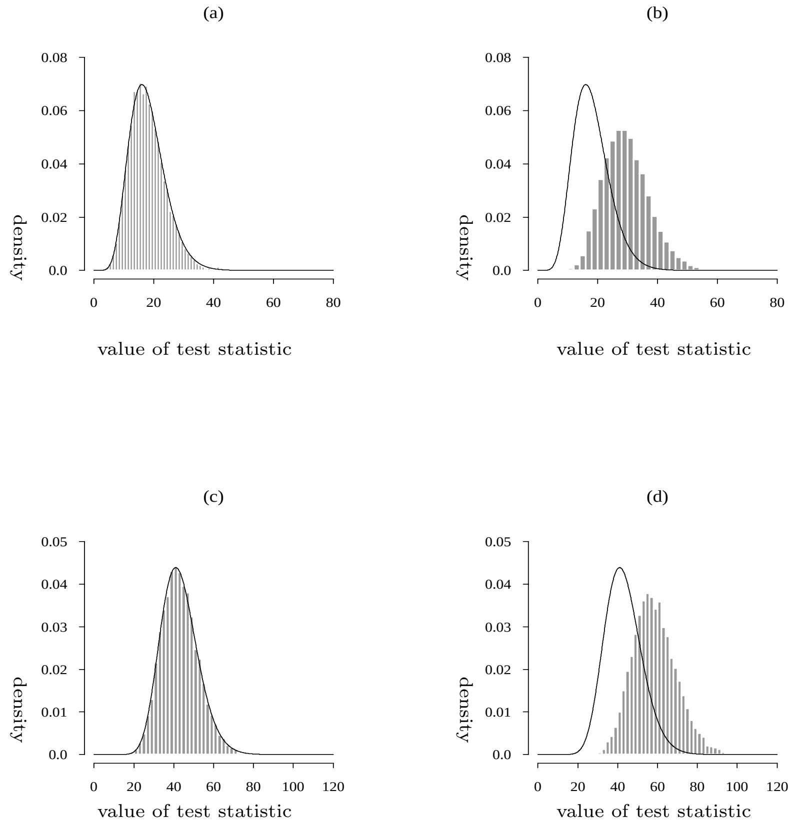

Figure 2. The simulated densities (histograms) and the chi-square approximations (solid lines). (a) when the manufacture errors are equal to the allowable level 0.05, compared to χ^2_{18}. (b) when some systematic manufacture errors are added to the simplex product. (c) when the manufacture errors are equal to the allowable level 0.05, compared to χ^2_{43}. (d) when the radius of the sphere is made 0.1 unit larger than designed.

Instead of assuming that all manufacture errors are random, we now add some systematic errors as well. One way to do this is to assume that in (2.1), x_{ij} are points on new planes S'_i $(i = 1, 2, 3)$ whose normal directions are $v_1 = (0, 0.04471018, -0.999)^t$, $v_2 = (-0.998, 0.06321392, 0)^t$, and $v_3 = (0.07740155, -0.997, 0)^t$. We repeat the above simulation with $\sigma = 0.05$ and find the following results:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{18,\alpha})$	0.9541	0.9267	0.8889	0.8232	0.6973	0.5742	0.4634	0.3309.

See also Figure 2 (b) for a graphical comparison. In terms of testing the hypothesis stated in Section 2.2, H_o: $\sigma^2 \leq 0.05^2$ is more likely to be rejected than accepted. In other words, a product with systematic manufacture errors, or large undesirable manufacture variation ($\sigma > 0.05$), will more likely be judged to have unsatisfactory geometric quality by our method.

Example 2. In this example, the assumed product is like a screw (see Figure 1). The top S_1 is part of a sphere with a radius equal to 1.5 units. The main body S_2 is a cylinder. Since it is possible that the locations and shapes of some curves are also important, we define S_3 as the bottom circle, S_4 as the circle on top of the cylinder and S_5 as the edge of S_1. The last three objects are one dimensional. Six points x_{ij} with 60 degrees apart are taken from each of S_3, S_4 and S_5, and six randomly chosen points are taken from each of S_1 and S_2. Then model (2.1) is used to generate y_{ij}, $i = 1, \ldots, 5$ and $j = 1, \ldots, 6$, in the same way as in Example 1. We simulate 10,000 $Q(\hat{T}, \hat{R})/\sigma^2$ values for each fixed σ. Again, we find that the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$ is essentially independent of σ for the range $\sigma = 0.0005$ (0.0005) 0.05. For $\sigma = 0.05$, we have

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{43,\alpha})$	0.4964	0.3952	0.2968	0.1966	0.0964	0.0478	0.0238	0.0097
$P(Q/0.05^2 > \chi^2_{42,\alpha})$	0.5413	0.4374	0.3367	0.2269	0.1169	0.0591	0.0303	0.0119.

Because $\nu = 6 + 6 + 12 + 12 + 12 - 6 = 42$, one may expect χ^2_{42} to be a suitable reference distribution. However, χ^2_{43} approximates the distribution of $Q/0.05^2$ better than χ^2_{42} does; see Figure 2 (c). This is because the product is invariant to rotation about the z-axis; thus, only 2 rotation parameters need to be estimated and a total of 5 degrees of freedom are lost. This example illustrates that the exact distribution of $Q(\hat{T}, \hat{R})/\sigma^2$ also depends on the physical shape of the product.

Next, we introduce systematic manufacture errors into the product. Let $x_{1j}, j = 1, \ldots, 6$ be on a new sphere S'_1 whose radius is 1.6 units. We repeat the simulation and find the following results:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010,
$P(Q/0.05^2 > \chi^2_{43,\alpha})$	0.9317	0.8958	0.8455	0.7669	0.6253	0.4896	0.3721	0.2515.

(Also see Figure 2 (d).) Again, in terms of testing the hypothesis stated in Section 2.2, products with systematic manufacture errors, or large undesirable manufacture variation, are more likely to be judged to have poor geometric quality.

4. An Application

Figure 3 shows a part in an automatic seat belt system for a car. As a metal bar, its two ends are straight, its middle portion is curved as a circular band with inner quarter circle $\widehat{AB}$ and outer quarter circle $\widehat{CD}$, and there are six rectangular holes on the bar designed for the purpose of fastening. The important geometric features include the location of the centers of the rectangular holes, the orientation of these holes, the roundness of the two quarter circles, and the angle formed by the two straight ends of the bar.

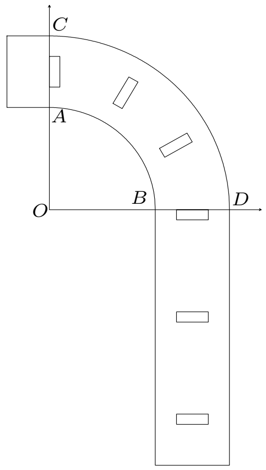

Figure 3. A part in an automatic seat belt system for a car. As a metal bar, its two ends are straight, its middle is curved as a circular band with inner quarter circle $\widehat{AB}$ and outer quarter circle $\widehat{CD}$, and there are six rectangular holes on the bar designed for the purpose of fastening.

For the reason of confidentiality, the name of the company which produces this part, the make of the car and the original design will not be released. Also, only simulated CMM measurements are to be discussed here. Otherwise, the context provided in Figure 3 (a two dimensional application) is completely real.

The sampling plan used by the company is to take eight CMM measurements for each rectangular hole with two points on each side line, and to take seven CMM measurements with 15° apart for each quarter circle. It is known that the precision of the CMM is $\sigma_2 = 0.0005$ unit. Therefore, if the allowable manufacture error is $\sigma_1 = 0.005$ unit, then the total allowable error is $\sigma = (0.0005^2 + 0.005^2)^{1/2} = 0.005025$ unit, and we can set up the following test:

$$H_0: \ \sigma^2 \leq 0.005025^2 \quad \text{vs} \quad H_1: \ \sigma^2 > 0.005025^2.$$

We reject H_0 (claiming the part to be defective) at, say, 1% level, if $Q(\hat{T}, \hat{R}) /0.005025^2 \geq \chi^2_{59,0.01}$, where $59 = \nu = 6 \times 8 + 2 \times 7 - 3$ is obtained by adapting

(2.5) to the two dimensional situation, where only 3 parameters (one for rotation, two for translation) are estimated and each point measured has one degree of freedom.

We simulate a set of CMM measurements according to model (2.1), the assumption that the allowable manufacture error is $\sigma_1 = 0.005$ and the above specification and sampling plan. Then we fit model (2.1) to the simulated data and find $Q(\hat{T}, \hat{R})/0.005025^2 = 64.4132$. Since $64.4132 < \chi^2_{59,0.01} = 87.1657$, we conclude that this simulated part has satisfactory geometric quality. As $\chi^2_{59,0.05} = 77.9305$, a 5% test will also draw the same conclusion.

5. Some Remarks

In both the examples and the application, the CAD files are assumed to exist and represent the ideal products, and the geometric quality of the real products can then be checked against the CAD files. If a product was created without the aid of computer technology and a CAD file is now needed to catalog the product, the problem becomes how to judge the accuracy of a CAD file once it is created. Clearly, our proposed method applies to this 'dual' problem as well.

Because our proposed measure of geometric quality is a sum of squares, where the summands are indexed by the different features we want to inspect, two straightforward diagnostic methods can be developed. The first and simple method is to break the total sum of squares into sums of squares due to $S_i, i = 1, \ldots, m$ and look at their sizes. The second method is to drop the points on each S_i in turn and fit the rest of the data to see which part of the product has the worst quality. If necessary, one can try to drop points from several S_i's simultaneously.

Our simulations reveal that χ^2_ν is indeed a good approximation to the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$. However, when all S_i are planar, it is shown in Section 2.2 that $\chi^2_{\nu+3}$ is an upper bound. When the shape of a product is more complex, we can always think of approximating a small curved surface by its tangent plane because the σ we are dealing with is usually small. In this case, the $\chi^2_{\nu+3}$ upper bound works approximately again. Since m and n_i are large in practice, and in general there are 6 parameters to be estimated, a χ^2 distribution with ν degrees of freedom (see (2.5)) should be a good approximation to the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$.

We should emphasize that our proposed method is applicable in inspecting finely machined products that obey rigid body transformations, such as those produced by casting or extrusion. We have considered mainly some statistical issues related to our proposed method. Some of these issues may need to be readdressed in practice. In particular, the assumption about the error structure perhaps needs to be relaxed from iid $N_3(0, \sigma^2 I_3)$ to more general types, such as elliptical distributions. Also, there are some non-statistical issues that must be addressed before one can implement a CMM based geometric quality inspection method (see Dowling et al. (1993)). For the problem formulated in this paper, a key issue is the need to introduce or develop tolerancing criteria that are appropriate for "soft-gauging" methods, such as the methods using CMM and CAD files. Finally, When the product has a sophisticated surface, the numerical calculation may not be as simple as it appears in our examples and application. We welcome interested collaborators to develop a computer program to implement our proposed method.

Acknowledgements

We want to thank an anonymous referee whose constructive comments have led to improvements in content and exposition. We also want to thank Professors Robert Chapman and Peter Kim for their very helpful comments during the preparation of this paper. A special thank is to Professor Stephen B. Vardeman who brought to our attention a large number of references. The research of both authors is supported by the Natural Sciences and Engineering Council of Canada.

Appendix

Proof of Properties 1, 2 and 3. Let $e_{ij} = \sigma\varepsilon_{ij}$, where ε_{ij} are $N_3(0, I_3)$. Denote the transpose of R_o by R_o^T, then (2.1) can be rewritten as $y_{ij} = R_o^T(x_{ij}+\sigma\varepsilon_{ij}-T_o)$, and for any T and R,

$$\begin{aligned} Q(T,R) &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle Ry_{ij}+T-a_i, v_i\rangle^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle R\{R_o^T(x_{ij}+\sigma\varepsilon_{ij}-T_o)\}+T-a_i, v_i\rangle^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle (RR_o^T-I_3)x_{ij}+\sigma RR_o^T\varepsilon_{ij}+T-RR_o^T T_o, v_i\rangle^2, \end{aligned} \tag{5.6}$$

where we have used the fact that $\langle x_{ij}-a_i, v_i\rangle = 0$. From $Q(T_o, R_o) \geq Q(\hat{T}, \hat{R}) \geq 0$ and

$$Q(T_o, R_o) = \sigma^2 \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle \varepsilon_{ij}, v_i\rangle^2 \to 0$$

almost surely as $\sigma \to 0$, it follows that $Q(\hat{T}, \hat{R})$ converges almost surely to 0 as $\sigma \to 0$. By assumption (2.4) and expression (5.6), this implies that $\hat{T}$ converges almost surely to T_o and $\hat{R}$ converges almost surely to R_o.

Similar to Goch and Tschudi (1992), when R is in a neighborhood of R_o, we can write, as $\sigma \to 0$,

$$(RR_o^T - I_3)/\sigma = \tilde{R} + o(1), \text{ or } RR_o^T = I_3 + \sigma(\tilde{R} + o(1)).$$

Then

$$[(RR_o^T - I_3)/\sigma]x_{ij} = \tilde{R}x_{ij} + o(1)x_{ij},$$
$$RR_o^T\varepsilon_{ij} = \varepsilon_{ij} + \sigma(\tilde{R} + o(1))\varepsilon_{ij} \;=\; \varepsilon_{ij} + o_p(1),$$
$$(T - RR_o^T T_o)/\sigma = (T - T_o)/\sigma + \tilde{R}T_o + o(1)T_o.$$

From expression (5.6), we have

$$\begin{aligned} Q(T,R)/\sigma^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle[(RR_o^T - I_3)/\sigma]x_{ij} + RR_o^T\varepsilon_{ij} + (T - RR_o^T T_o)/\sigma, v_i\rangle^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle\tilde{R}x_{ij} + \varepsilon_{ij} + (T - T_o)/\sigma - \tilde{R}T_o, v_i\rangle^2 + o_p(1), \end{aligned}$$

because m and n_i are fixed. Therefore,

$$Q(\hat{T},\hat{R})/\sigma^2 = \inf_{T,R} Q(T,R)/\sigma^2 = \inf_{\alpha,\beta,\gamma,T'}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle\tilde{R}x_{ij} + \varepsilon_{ij} + T', v_i\rangle^2 + o_p(1).$$

Hence, the limiting distribution is as claimed.

To prove Property 2, we note that $Q(\hat{T},\hat{R})$ has to be bounded by $Q(T_o,R_o)$ because the former is the minimum. The chi-square distribution property of $Q(T_o,R_o)/\sigma^2$ follows, because $\langle\varepsilon_{ij}, v_i\rangle^2$ ($1 \le i \le m$, $1 \le j \le n_i$) are independent and identically distributed, and each $\langle\varepsilon_{ij}, v_i\rangle^2$ has a chi-squares distribution with 1 degree of freedom.

To prove Property 3, we note that for any T

$$\begin{aligned} Q(\hat{T},\hat{R})/\sigma^2 &\le Q(T,R_o)/\sigma^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle\sigma\varepsilon_{ij} + T - T_o, v_i\rangle^2/\sigma^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle\varepsilon_{ij} - \varepsilon_{i\cdot}, v_i\rangle^2 + \sum_{i=1}^{m} n_i\langle\varepsilon_{i\cdot} + (T - T_o)/\sigma, v_i\rangle^2, \end{aligned}$$

where $\varepsilon_{i\cdot} = n_i^{-1}\sum_{j=1}^{n_i}\varepsilon_{ij}$. Therefore,

$$Q(\hat{T},\hat{R})/\sigma^2 \le \sum_{i=1}^{m}\sum_{j=1}^{n_i}\langle\varepsilon_{ij} - \varepsilon_{i\cdot}, v_i\rangle^2 + \inf_{T'}\sum_{i=1}^{m} n_i\langle\varepsilon_{i\cdot} + T', v_i\rangle^2.$$

Because $\varepsilon_{ij} - \varepsilon_{i\cdot}$ is independent of $\varepsilon_{i\cdot}$, the two terms on the right-hand side of the above inequality are independent of each other. A little calculation shows that the first term has a chi-square distribution with degrees of freedom $\sum_{i=1}^m (n_i - 1)$, and the solution for T' in the second term is

$$\hat{T}' = -\Big(\sum_{i=1}^m n_i v_i v_i^t\Big)^{-1}\Big(\sum_{i=1}^m n_i v_i v_i^t \varepsilon_{i\cdot}\Big).$$

It is not difficult to verify that $\sqrt{n_i} v_i^t \varepsilon_{i\cdot}$ are independent and identically distributed standard normal random variables. Let $\mathbf{y}$ be the column vector with $\sqrt{n_i} v_i^t \varepsilon_{i\cdot}$ as its ith component, and let C be the matrix with $\sqrt{n_i n_j} v_i^t (\sum_{l=1}^m n_l v_l v_l^t)^{-1} v_j$ as its (i,j)th element. Then the second term becomes $\mathbf{y}^t(I - C)\mathbf{y}$. Because the matrix $I - C$ is idempotent, it follows that the second term has a chi-square distribution with degrees of freedom equal to

$$\begin{aligned}\mathrm{tr}(I - C) &= m - \sum_{k=1}^m \Big\{ n_k v_k^t (\sum_{i=1}^m n_i v_i v_i^t)^{-1} v_k \Big\} \\ &= m - \mathrm{tr}\Big\{ \sum_{k=1}^m n_k (\sum_{i=1}^m n_i v_i v_i^t)^{-1} v_k v_k^t \Big\} = m - 3.\end{aligned}$$

This proves Property 3.

References

Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, 239–256.

Berman, M. and Culpin, D. (1986). The statistical behavior of some least squares estimators of the center and radius of a circle. *J. Roy. Statist. Soc. Ser. B* **48**, 183–196.

Brou, P. (1984). Using the gaussian image to find the orientation of objects. *Internat. J. Robotics Research* **3**, 89–125.

Caskey, G., Hari, Y., Hocken, R., Palanivelu, D., Raja, J., Wilson, R., Chen, K. and Yang, J. (1990). Sampling techniques for coordinate measuring machines. *Design and Manufacturing Systems Conference* 779–786.

Chapman, G. R. and Kim, P. T. (1992). Spherical regression in quality assurance. The Institute for Improvement in Quality and Productivity, Research Report RR–92–01, University of Waterloo, Ontario, Canada.

Chapman, G. R., Chen, G. and Kim, P. T. (1995). Assessing geometric integrity through spherical regression techniques (with discussion). *Statist. Sinica* **5**, 173–220.

Chen, G. and Chen, J. (1992). Geometric quality assurance. The Institute for Improvement in Quality and Productivity, Research Report RR–92–06, University of Waterloo, Ontario, Canada.

Cox, M. G. and Peggs, G. N. (1986). *Software for Co-ordinate Measuring Machines.* Proceedings of a Conference held at National Physical Laboratory on 25 September, Teddington, Middlesex, Britain.

Dowling, M. M., Griffin, P. M., Tsui, K.-L. and Zhou, C. (1993). Statistical issues in geometric feature inspection using coordinate measuring machines. Research Report No. J-93-02, School of Industrial and Systems Engineering, Georgia Institute of Technology.

Foster, L. W. (1986). *Geo-Metrics II—The Application of Geometric Tolerancing Techniques.* Reading, Addison-Wesley, Massachusetts.

Goch, G. and Tschudi, U. (1992). A universal algorithm for the alignment of sculptured surfaces. *Annals of the CIRP* **41**, 597–600.

Guiford, J., Sethi, M. and Turner, J. U. (1992). Worst case and statistical tolerance analysis of the daughter card assembly. *Computers in Engineering. Proceedings of The 1992 ASME International Computers in Engineering Conference and Exposition, August 2–6* **1**, San Francisco, California.

Gunasena, N. U., Lehtihet, El-Amine and Ham, I. (1989). The verification of composite position tolerance. *IIE Transactions* **3**, 290–299.

Horn, B. K. P. (1984). Extended Gaussian Images. *Proceedings of IEEE* **72**, 1656–1678.

Horn. B. K. P. (1986). *Vision.* MIT Press, Cambridge, Massachusetts.

Hulting, F. L.(1992). Methods for the analysis of coordinate measurement data. *Comput. Sci. Statist.* **24**, 160–169.

Hulting, F. L.(1993). Process capability analysis with geometric tolerances. *1993 Proceedings of the Quality and Productivity Section of the American Statistical Association*, American Statistical Association.

Kurfess, T. R. and Banks, D. L. (1990). Statistical verification of conformance to geometric tolerance. Technical Report 24–38–90, Engineering Design Research Center, Carnegie Mellon University, Pittsburgh, PA.

Liggett, J. V. (1970). *Fundamentals of Position Tolerance.* Society of Manufacturing Engineers, Dearborn, Michigan.

Madsen, David A. (1988). *Geometric Dimensioning and Tolerancing.* The Goodheart-Willcox Company, Inc., South Holland, Illinois.

Marshall, A. D. and Martin, R. R. (1992). *Computer Vision, Models and Inspection.* World Scientific Publishing Co. Pte. Ltd., Singapore.

Menq, C. H., Yau, H. T., Lai, G. Y. and Miller, R. A. (1990). Statistical evaluation of form tolerance using discrete measurement data. *Advances in Integrated Product Design and Manufacturing, ASME WAM at Dallas* PED-Vol **47**, 135–149.

The NAG Fortran Library. (1987). Numerical Algorithms Group Inc, 1101 31st Street, Suite 100, Downers Grove, IL 60515, U.S.A.

Rooney, J. and Steadman, P. (1987). *Principles of Computer-aided Design.* Pitman, London.

Spotts, M. F. (1983). *Dimensioning and Tolerancing for Quantity Production.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Stuetzle, W. (1994). Piecewise smooth surface reconstruction. Invited talk at the Annual Meeting of the Statistical Society of Canada.

Tannock, J. D. T. (1992). *Automating Quality Systems.* Chapman and Hall, London.

Turner, J. U. (1990). Exploiting solid models for tolerance computations. In *Geometric Modeling for Product Engineering* (Edited by M. J. Wozny, J. U. Turner and K. Preiss), North-Holland.

Yang, S., Li, Z. and Li, G. (1988). Statistical inference of form error distribution function. *Precision Engineering* 97–99.

Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan S4S 0A2, Canada.

E-mail: gchen@boxcox.math.uregina.ca

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

E-mail: jhchen@setosa.uwaterloo.ca

(Received December 1995; accepted April 1997)