University of Huddersfield Repository

Lou, Shan, Jiang, Xiangqian, Bills, Paul J. and Scott, Paul J.

Defining True Tribological Contact Through Application of the Morphological Method to Surface Topography

Original Citation


This version is available at http://eprints.hud.ac.uk/16778/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/
Defining True Tribological Contact through Application of the Morphological Method to Surface Topography

Shan Lou, Xiangqian Jiang*, Paul J. Bills, Paul J. Scott

EPSRC Centre for Innovative Manufacturing in Advanced Metrology, University of Huddersfield, Huddersfield, HD1 3DH, UK

*Corresponding author: Tel: +44 1484 473634; Fax: +44 1484 472161; E-mail: x.jiang@hud.ac.uk

Abstract In tribological functions high peaks (summits) in the surface topography play a dominant role in that they determine the position of first contact and how the contact will occur. Both statistic based methods and feature based methods address the characterization of a single surface while neglecting the interacting surface. A morphological method is proposed to simulate the contact of two mating surfaces. The surface under evaluation is rolled by a ball with radius meant to simulate the largest reasonable peak curvature at a contact. In such a situation the contact points of the rolling ball may serve as an identification of those surface portions that are in real contact. The morphological closing operation could then be applied to detect the contact points of the rolling ball, however, the traditional computation method does not lead to an accurate result. To overcome this deficiency, a geometrical computation approach has been developed to capture the contact points based on four searching procedures. The resulting method has been verified through experimentation and then applied to a case study in which the underlying form of the surface of a hip replacement taper junction is analyzed to remove the effect of the dominant threaded structure.

Keywords: surface topography; rough surface; contact point; morphological operations

1. Introduction

The surface of a component is an interface which limits the body of the component and separates it from the surrounding medium [1]. It governs the functional behaviour of the product, whether that be a mechanical, thermal, chemical or biological property, all of which are of tremendous importance in the tribology of any system. Within a tribological system, contact mechanisms are fundamental and the geometry of contact depends strongly on the surface topography [2].

In many engineering applications, the contact between two surfaces is non-conforming, i.e., the contact area is very small when compared with the geometry of the bodies in contact. Even in situations between conforming contacts, the contact between the asperities that compose the
surface topography is known to be non-conforming [3]. Historically, statistical models have been used to predict contact parameters which have contained many assumptions about asperity geometry and height distributions, and in which the asperities on a rough surface were modeled as an array of hemisphere or paraboloid bumps [4-6]. These models are highly conceptual and thus a geometrically complicated surface is usually represented by a few statistical parameters. Although the statistical modes result in simple relationships and are able to predict important trends in the effect of surface properties on the real area of contact, they are limited because of over simplifying assumptions about asperity geometry and height distributions, the difficulty in determination of statistical roughness parameters [7, 8].

In the 1970s, Nayak [9] and Sayles & Thomas [10] initially used the five nearest-neighbour ordinates in areal surface data to define a peak or pit. In order to investigate contact phenomena of random surfaces, Whitehouse & Phillips [11] also initially defined three areal parameters: summit density, summit height and summit curvature. These definitions, however, depended on sampling density, and the results could be distorted by measurement noise. In the 1990s an integrated method for the areal characterization of surfaces was explored by Stout et al [12], which led to the so-called “Birmingham 14 parameters”. Although these parameters could provide a general description of the rough surface in a statistical sense, they neglect the local complexity of surface geometries, which determines the actual contact areas. The last decade saw more novel methods focusing on topographical features [13, 14]. The surface is treated as a collection of Maxwellian features, such as hills, dales, saddle points, ridge lines and course lines and it is partitioned into a series of regions that contains individual topographically significant features. The segmented geometrical features can then be analyzed individually or statistically.

The segmentation method is a major progress in characterizing surface geometry and a useful tool in analyzing contact phenomena. However this method, in common with statistical methods, only addresses a single independent surface, whereas contact clearly involves the interaction of two mating surfaces. It therefore follows that the influence of the opposite surface should be taken into account while evaluating the master surface. This paper proposes a novel morphological method to simulate the contact of two interacting surfaces. The interaction is simulated by rolling a ball with a given radius, which is sized to simulate the largest reasonable radius at a contact e.g. peak curvature, upon the underlying surface. The contact points of the rolling ball against the rolled surface are then captured. This serves as an indication of surface summits and surface portions which are in real contact.

2. Morphological operations

2.1 Morphological operations in image processing

Mathematical morphology is a mathematical discipline which aims at extracting the shape and form of objects in an image by probing the image with the structuring element [15].

Four basic morphological operations, namely dilation, erosion, opening and closing, form the foundation of mathematical morphology. Dilation combines two sets using the vector addition of set elements. The dilation of $A$ by $B$ is:
\[ D(A, B) = A \oplus \bar{B} . \]  

(1)

where \( \bar{B} \) is the reflection of \( B \) through the origin of \( B \).

Erosion is the morphological dual to dilation. It combines two sets using the vector subtraction of set elements:

\[ E(A, B) = A \ominus \bar{B} . \]  

(2)

where

\[ A \ominus B = \overline{A + B} . \]  

(3)

Opening and closing are dilation and erosion combined pairs in sequence. The closing of \( A \) by \( B \) is obtained by applying the dilation followed by the erosion,

\[ C(A, B) = E(D(A, B), \bar{B}) . \]  

(4)

The reverse order generates the opening:

\[ O(A, B) = D(E(A, B), \bar{B}) . \]  

(5)

2.2 Morphological operations in surface metrology

Morphological operations are universal in the field of surface metrology. The scanning of a workpiece surface using a tactile probe is a very common practice in geometrical measurement and a hardware implementation of morphological dilation operations [16]. The workpiece surface as the input set is dilated by the structuring element, in this case the spherical probe tip, to generate the morphological output, the measured surface, see Figure 1.

![Figure 1](image.png)

**Figure 1.** The dilation of the workpiece surface by a spherical tip.

It could be noticed that the measured data are not the real surface data, but the tip centre data. To reconstruct the real surface, an ideal sphere with the same size to the probe tip is employed to roll over the dilated profile, the locus of the sphere centre is treated as the real mechanical surface. See Figure 2. This treatment is in essence a morphological erosion operation. The two operations,
a dilation followed by an erosion, are combined to yield a single effect, i.e. the closing operation. The resulted closing envelope overlaps with the profile peaks while it hovers over the profile valleys where the surface local curvature is smaller than that of the tip.

![Figure 2](image)

**Figure 2.** The closing of the workpiece surface by a spherical tip.

3. **Contact points and their search procedures**

3.1 Contact points

In physics, the contact points are those points on the surface which are in contact with the rolling ball. These points give an indication that their neighborhood surface portions are most likely to be active in contact phenomenon. By identifying the contact points, those areas of a surface that may be especially susceptible to wear at process start-up can be readily identified and remedial action taken if necessary. From a point of view of mathematical morphology, the contact points are those points on the surface that remain constant with the morphological closing operation, see Figure 2. Therefore these points might be captured by computing the closing envelope and comparing it with the original surface. The overlapping portions are the contact points. This solution, however, is impractical due to the fact that the numerical comparison is sensitive to round off errors in calculation and this situation is even compounded by sampling the structuring element discretely. As a consequence, a capable algorithm is required in order to detect contact points in a robust fashion.

3.2 Searching procedures

In stead of computing the dilation and erosion to yield the closing, the developed method computes the contact points in a geometrical manner. A valuable property which could be of paramount value in searching for contact points is the containment relationship of the contact points in response to diverse ball radii [17]. Given a surface, two balls with different sizes are rolled over the underlying surface, one with a large radius and the other one with a small radius. The contact points of the large ball are contained within the boundaries of those of the small ball. This property establishes the searching sequence for the contact points. The large ball radius is
utilized first for ease of computation. The searching routine then downscales the rolling ball until it reaches the given radius settled for simulating the largest curvature of the interacting surface. The following procedures demonstrate the details of searching operations. In the context of the statements below, \(a\) and \(b\) are two known contact points, e.g. two end points of a surface profile, \(r\) is the given radius of the ball (disk).

**Procedure 1:** If there are sample points lying above \(\overline{ab}\) (left/positive side of \(\overline{ab}\)), then the contact point is the furthest sample point orthogonal to \(\overline{ab}\).

As illustrated in Figure 3, there are a number of sample points between \(a\) and \(b\). The furthest point from \(\overline{ab}\) is \(p_5\), which in fact is a point on the convex hull of the point set \(\{a, b, p_1, \ldots, p_6\}\) above \(\overline{ab}\). This searching procedure is suggested by the computation of the convex hull [18] and corresponds to rolling a disk with an infinitely large radius over the surface profile.

![Figure 3. Search the furthest point orthogonal to \(\overline{ab}\).](image)

**Procedure 2:** If there are no sample points lying above \(\overline{ab}\) and sample points \(\{p_i\}\) exist in the circular section \(\overline{ab}\) of the ball with radius \(\alpha = \max\{r, \sqrt{ab}\}\), then the contact point is that of \(\{p_i\}\) in \(\overline{ab}\), which has the largest radius among the circumscribed circles of the simplices \(\{\sigma_{abp}\}\).

This procedure includes two different cases, \(|ab| \leq 2r\) and \(|ab| > 2r\). For the first case, see Figure 4(a), the disk with the given radius \(r\) is not empty because it encloses two sample points, in this case \(p_4\) and \(c\). From this, it follows that some sample points between \(\overline{ab}\) must be contact points. Next the circumscribed circle of \(\{\sigma_{abp}\}\) is calculated. It is obvious that the circumcircle of \(\sigma_{abc}\) is empty, while other circumcircles are not, and its radius is the largest so far calculated. The second case where \(|ab| > 2r\) is presented in Figure 4(b), and from this we can find a similar result.
Procedure 3: If $|ab| > 2r$, no sample points lie above $\sigma_{ab}$ and no points lie within the circular section $\overline{ab}$ of the ball which has a radius $\alpha = \frac{1}{2}|ab|$, then the contact point is identified as that which has the smallest radius among the circumscribed circles of $\{\sigma_{ab_p}\}$.

See Figure 5. Among the calculated circumcircles, only $\sigma_{abc}$ contains no sample points and furthermore has the smallest radius.

Both Procedure 3 and Procedure 4 calculate the circumscribed circles. The difference is that the former takes the largest circumcircle and the later takes the smallest one. However the two situations can be unified by the use of the signed circumcircle radius.

If there are sample points $\{p_i\}$ lying below $\sigma_{ab}$ (right/negative side of $\overline{ab}$) and no point above, the simplex $\sigma_{ab_p}$ has an unique circumscribed circle with radius $\alpha$. If the centre of the circumscribed circle is on the positive side of $\sigma_{ab}$, then the circle has the positive radius $+\alpha$. 
otherwise the negative radius \(-\alpha\). In Figure 6, \(\sigma_{abp}\) has its circumcircle centre \(o_1\) above \(\sigma_{ab}\), thus it has a positive radius. Conversely, the centre of the circumcircle of \(\sigma_{abp}\) lies below \(\sigma_{ab}\), therefore its radius is negative. The critical case is that of \(\sigma_{abp}\) which has its circumcircle centre \(o\) at the centrepoint of \(\sigma_{ab}\). In this case it is taken that the radius is positive. With the signed circumscribed circle radius, Procedure 2 and Procedure 3 are unified and both take the largest circumcircle radius.

**Figure 6.** Signed circumscribed circle radius.

**Procedure 4:** If \(|ab| \leq 2r\), no sample points lie above \(\sigma_{ab}\) and no sample points lie within the circular section \(\overline{ab}\) of the given disk with radius \(r\), meaning that the disk is empty, then there are no contact points between \(\sigma_{ab}\) and the searching procedure exits.

**Figure 7.** Sampled points distribution below \(ab\).
To determine the termination of the searching procedure, Procedure 4 suggests three conditions. These conditions can be calculated by examining the sampled point distribution below $\sigma_{ab}$, see Figure 7. First of all, $|ab|$ has to be equal or smaller than $2r$, which means the given disk is larger than the smallest circumcircle of $\sigma_{ab}$. In Figure 7, there are 7 sample points between $a$ and $b$. They fall into three categories:

1. Points lying within the circular section $\widehat{ab}$ of the ball with the given radius, such as $p_3$ and $p_7$, and may be contact points. They are featured by the positive radius $\{\rho_{i}\}$ of the circumcircle of $\{\sigma_{abp}\}$ and $\rho_{i} \geq r$.

2. Points lying in the circular section $\widehat{ab}$ of the smallest circumcircle of $\sigma_{ab}$, but that are not in category (1), such as $p_1$, $p_4$, and $p_5$. These points cannot be contact points as they have positive radii $\{\rho_{i}\}$ of the circumcircle of $\{\sigma_{abp}\}$, but $\rho_{i} < r$. Thus in this case $0 \leq \rho_{i} < r$.

3. Points not contained in categories (1) and (2), such as $p_2$ and $p_6$, and cannot be contact points. These points have negative radii $\{\rho_{i}\}$ of the circumcircle of $\{\sigma_{abp}\}$, i.e., $\rho_{i} < 0$.

To sum up, the searching procedure terminates when no points lie above $\sigma_{ab}$, $|ab| \leq 2r$ and $\rho_{i} < r$.

3.3 Searching algorithm

A practical recursive algorithm has been constructed based on the above searching procedures. The pseudocode of the algorithm to compute contact points on the surface profile is presented in Figure 8. The algorithm starts with the left end sample point $a$ and the right end sample point $b$ of the measured profile, which are guaranteed to be the initial contact points as they are on the convex hull. The algorithm then starts to search the contact points between $a$ and $b$ in sequence by applying procedures 1 to 4. Once a contact point is found, such as a point $c$, it can be treated as a partition point and the profile $(a,b)$ is partitioned into two segments $(a,c)$ and $(c,b)$. The algorithm keeps partitioning the segments into smaller ones until the segment being evaluated satisfies the condition specified by Procedure 4 and that segment is accepted as the boundary facet. The vertices of all final boundary facets are then defined as the contact points.
Figure 8. Recursive algorithm for searching contact points on the surface profile.

```
Algorithm ContactPoints(S, r)
{Given a profile S and the chosen disk radius r,
} computes the contact points Contacts.

a ← the left end point of S.
b ← the right end point of S.
Partition(a, b);

Procedure Partition (a, b)
if \{p_i\} above ab
    c ← the furthest point from ab in \{p_i\};
else
    calculate the signed circumcircles radii R of
    \{ab p_i\};
```

The searching procedures for profile data also hold for areal data if the disk is replaced by a ball and the circumcircle is replaced by a circumsphere. In such a case instead of starting with the left and right profile ends in the scenario of profiles, it is easier to start with the convex hull faces for areal data and thereafter repeat the partition for each convex hull face. Figure 9(a) illustrates an example surface with 50 $\times$ 50 points on which the convex hull faces are presented as the triangular meshes. The algorithm searches for contact points by computing signed circumsphere radii and partitions each convex hull face. For instance, starting with a convex hull face $\sigma_{abc}$, a contact point $d$ is found by seeking the largest circumsphere radius of $\{\sigma_{abpc}\}$, where $\{p_i\}$ are the sample points inside the circumsphere of $\sigma_{abc}$. Then $\sigma_{abc}$ is partitioned into three new simplices $\sigma_{abd}$, $\sigma_{bcd}$ and $\sigma_{cad}$. The partition process is repeated on each new generated simplex until it can hold an empty circumsphere. The highlighted triangle in Figure 9(a) denotes one of convex hull faces. The resulting boundary facets are highlighted in Figure 9(b). When the searching procedure is completed, the vertices of the obtained boundary facets are the desired contact points, as shown in Figure 9(c).
Figure 9. Partition procedure on an areal data. (a) Convex hull faces of the areal data; (b) Boundary facets generated by partitioning one of the convex hull face; (c) Final boundary facets.

3.4 Verification

For the purpose of verifying the proposed method, it is applied to an experimental profile and a surface respectively. The profile is 1.25 mm in length with sampling interval of 5 μm and applied by the morphological method using disks with radius 5 mm and 0.5 mm respectively. Figure 10 illustrates the contact points of the profile. It is clearly shown in the figure that the contact points of the 5 mm disk identify significant profile peaks while those of the 0.5 mm disk involve less...
significant peaks. Furthermore, the contact points of the 5 mm disk are contained in those of the 0.5 mm disk. The experimental surface illustrated in Figure 11 is 0.495*0.495 mm$^2$ in size with sampling interval of 5 µm in both X direction and Y direction. A similar result can also be found on the surface, using balls with radius 5 mm and 1 mm respectively.

![Graph](image)

**Figure 10.** The contact points of an experimental profile. (a) Disk radius 5 mm; (b) Disk radius 0.5 mm.

![Graph](image)

**Figure 11.** The contact points of an experimental surface. (a) Ball radius 5 mm; (b) Ball radius 1 mm.
4. Case study

In hip replacement, the introduction of modular large head metal-on-metal (LHMoM) hips promised low wear rates and reduce chances of dislocation couple with an increase range of motion compared to the conventional metal-on-metal hips. The clinical experience of the use of LHMoM hip replacements shows that they exhibit a significantly higher revision rate compared to other types of implant, at 5 years the revision rate is 7.8% compared to 6.3% for hip resurfacings and 2% for conventional cemented implants [19]. The difference in revision rate between resurfacings and LHMoM hips has been attributed to the neck/taper junction [20, 21], thus the specification and measurement of this area of the component is key to the understanding of the operation of the implant and the failure mechanisms at this interface.

The interlocking male taper surface that mates with the femoral head female counterpart has a structured micro-threaded surface, see Figure 12. The specification of such surfaces is not well understood but has been shown to be important as possible corrosion and wear at this interface have been identified as a possible source of debris that could cause tissue reaction and progress to implant failure. Analysis of this structured conical surface requires the extraction and examination of the conical form and contact. In this example vertical measurements are performed axially relative to the aligned component axis such that the outputted value of profile straightness can then be used as a measure of conical form.

![Figure 12. Total hip replacement femoral stem with highlighted micro-threaded taper surface.](image)

The combined effect of form and roughness has long been recognized in the measurement of the form of machined rough surfaces [22]. The effect of the surface structure on the resulting form value can be disproportionate, thus the size of the probe relative to texture spacing has to be large [23, 24]. Current industry practice in the measurement of hip stem tapers is to attempt to perform this task through use of mechanical filtering, by using a large diameter ruby stylus on a CMM. However, this is largely performed on a trial and error basis and makes no account of how much useful data is being discounted or erroneous data included through the bridging of surface contact points. Furthermore the use of such a large stylus method is suboptimal as the required component accuracy is on the limit of that of the CMM (< 1 µm). This coupled with the difficulties in locating data points when using a prohibitively large measurement stylus means that this method is far from ideal.

To overcome the deficits of mechanically traversing the stylus on CMM, the morphological method is employed to improve measurement accuracy and extract the contact points. To achieve this, a number of new hip replacement femoral stems were measured, a series of linear
measurements are performed on each component, axially along the neck taper. The measurements were performed using a Talyrond 365 roundness machine (Taylor Hobson, UK) with a 5 µm diamond stylus. Figure 13(a) presents such an example profile with sampling length 8 mm and sampling interval 0.25 µm. For convenience of visualization, the profile was translated and rotated, see Figure 13(b). The contact points with disk radius 5 mm are then extracted from the profile texture. Finally the form error of the straightness of the profile was calculated by applying the minimum zone method to the contact point set. The obtained straightness is 1.434 µm in the example.

![Figure 13(a)](image)

![Figure 13(b)](image)

**Figure 13.** Form evaluation of an experimental taper junction: (a) Surface profile measured along the neck taper; (b) The extracted contact points and the minimum zone for straightness.

The morphological method searching for contact points on the surface allows for the optimisation of mechanical traversing process by the determination of what equivalent stylus size would be required to perform this task. The use of a roundness machine and the proposed morphological method enable the data to be captured at a higher density and accuracy (Gauge resolution ~30 nm) with a greater level of control in the extraction of the true envelope profile.

5. **Conclusion**

In the contact phenomenon of interacting surfaces, peak features on surfaces play a dominate role in that they determine the position of first contact and how the contact will occur. Both statistics based methods and feature based methods address the characterization of a single independent
surface while neglecting the influence of the mating surface. This paper proposes a method of using the morphological method to simulate the interaction of two mating surfaces, which is carried out by rolling a ball with radius equal to the largest curvature of the slave surface at a contact upon the underlying surface.

The contact points of the rolling ball may serve as an identification of those surface areas that are in real contact and the morphological closing operation can be used to detect contact points. However an accurate solution cannot be reach by using the traditional computation method. A geometrical algorithm is developed for both profile and areal data based on four searching procedures. This method can accurately capture the contact points and has been verified by application to an experimental profile and a surface.

The contact points are employed to evaluate the underlying form of the textured surface of hip replacement taper junction. The use of surface texture instrument and the proposed morphological method guarantees the precision of measurement and accuracy of evaluation and allows for more accurate specification of component form which has been shown to be of prime importance to component performance.

Acknowledgements The authors gratefully acknowledge the UK’s Engineering and Physical Sciences Research Council (EPSRC) funding of the EPSRC Centre for Innovative Manufacturing in Advanced Metrology (Grant Ref: EP/I033424/1) and the European Research Council under its programme ERC-2008-AdG 228117-Surfund.

Conflict of interest The authors have no conflict of interest for this manuscript.

References