Surface measurement fidelity

David J. Whitehouse
University of Warwick, UK
Metrology Consultant to Taylor Hobson Ltd, Leicester, UK
E-mail: djw@dircon.co.uk

Abstract

There have for many years been advocates for two ways of measuring surfaces; the stylus method and optical methods. This paper examines the advantages and disadvantages of both and indicates the nature and the magnitude of the differences to be expected.

1 Introduction

The most common method for measuring surfaces is the stylus technique in which a sharp diamond stylus is made to traverse the workpiece. The stylus penetrates the valleys and rides the peaks thereby revealing a profile of the surface geometry (Ref. 1). There are many optical methods ranging through interferometry, diffraction and optical follower techniques (Ref. 2). In all these traditional forms the basic underlying quality is determined by the numerical aperture of the optics as will be seen later. The usual criticisms are that stylus methods scratch and optical methods distort! In what follows these criticisms will be examined.

2 Stylus Methods

2.1 Background

From time to time pictures appear in the scientific press purporting to demonstrate the damage that the stylus tip of a measuring instrument can cause on a surface. Usually these pictures are used as an excuse for a functional failure of the workpiece, a premonition that any mark on the surface could result in failure is shown simply to promote the use of an alternative surface measurement, usually optical.

Often such pictures are over exaggerated and sometimes quite simply are misleading. The fact of the matter is that evidence of functional failure due to
any surface marking is rarely, if ever, available. The only concrete fact is that sometimes a mark can be seen. The direct effect of this is that a workpiece could be rejected on the grounds of cosmetic appearance only. The fact that a mark can be seen is, in itself, not of significance.

2.2 True Nature of Mark

2.2.1 Scale of Size
One cause for misrepresentation is due to the way in which such marks are shown - always perpendicular to the direction in which the stylus has moved across the surface. This in itself is confusing. The extent of the so called damage is never shown in the direction of measurement. The effect is similar to that of a walker on a hillside Fig. (1). The depth of footprints as a fraction of the height of the hill is infinitesimal yet when viewed in a plan view can be seen. The scar should really be shown together with the profile in a side view then its true significance (or lack of it) can really be seen. This misrepresentation is the same situation as complaining about the lack of height information on a standard optical micrograph. To be realistic the metrology should mimic the direction of the tool used in the machining of the surface. Comparing like with like is axiomatic in surface metrology (Ref. 3).

![Figure 1: Quantitative side effect](image-url)
2.2.2 Assessment of Tracking Mark

![Figure 2: Integration of mark by eye and the magnified view](image)

Seeing a mark is not the same as measuring it. Very often even a clearly visible mark disappears when magnified by a microscope. This is because of the inherent integrating effect of the eye. This integration can be caused by poor visual acuity i.e. poor resolution. What happens is that the stylus slightly deforms some of the high peaks in its path. This deformation takes the form of a burnishing effect which causes the very small detail on some peaks to reflect light rather more that the surrounding texture. These burnished peaks are not necessarily in an exactly straight line: neither is the path of the burnished peaks continuous. The effect is shown in Fig. 2. So two of the prerequisites traditionally thought necessary to indicate the presence of damage (a) measurability and (b) continuity, are as often as not missing. The reason why the eye appears to see a line or scratch is because it automatically tends to link dispersed-like objects together - it correlates or links together similar objects optically. Another reason is that very small changes in reflectivity are artificially enhanced. This is a result of Fekners physiological law which instruments do not obey. Most stylus marks in effect cannot be seen at high magnification and certainly cannot be measured.

2.2.3 Effect of Skid

There is another issue has to be considered. This concerns use of a skid. This is often used to ease the setting up procedure prior to measuring the surface. Instead of an external reference surface (shown in Figure 3a) situated close to the surface but within the instrument, an intrinsic reference is used (Fig. 3b). In this the calliper of two arms forming the measurement loop has one arm with a blunt foot and the other with a sharp stylus. Both touch the surface to be measured. The blunt foot or skid provides the reference and the sharp stylus the probe searching out the detail on the surface.
The point about this latter configuration is that it enables tilted specimens to be easily measured, but notice that usually the whole of the weight of the pick up arm is supported by the skid. This can be considerable and produce pressures comparable with those produced by the stylus. A more likely source of damage produced by the skid occurs when a speck of dust or debris gets trapped between the skid and the work surface. If the debris is composed of an oxide or similar abrasive material and is small in size the pressures can be enormous, far larger than those pressures produced by the stylus itself. The message here is clear. If the surface is soft and/or liable to be damaged do not use a skid, instead use the stylus instrument in its true mode i.e. having a separate reference surface. Skid effect can usually be recognized by the relatively wide scatter of burnished peaks.

2.3 Physical Considerations

2.3.1 Material Effect

Whilst it makes sense to be cautious about using a stylus instrument for measuring soft surfaces it is nonsense to dismiss all stylus measurement out of hand. This is because of misconceptions in the meaning of the word hardness. Usually the hardness value as specified for any material is that of bulk material. This is taken to relate to the yield strength of the material which is the hardness measured by a conventional hardness tester which uses such measures as Vickers, Rockwell, Brinell etc. In these a large load is used to press the indentor into the surface. Depth of penetration of sub-millimetres are common. This type of measurement senses the yield stress well into the surface - the so called bulk hardness. This measure - the bulk hardness - is not the correct value of hardness upon which to judge the suitability of the stylus method for assessing the surface texture. What is needed is the hardness value in the region of the surface skin (which just encompasses the surface texture).

The elementary calculation for determining whether or not a stylus will damage the surface - at least with a static load is as follows. Assume that the stylus is a cone of radius $R_1$ then using Hertz equations the region under the
A stylus which elastically gives on a flat surface is \( \pi a^2 \left( \frac{4 W R'}{3 E'} \right)^2 \) where \( W \) is the load, \( E' \) is the composite Young's Modulus. Letting 1 stand for the stylus and 2 for the contacting surface – here a flat – then

\[
\frac{1}{E} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}
\]

where \( \nu \) is Poisson's ratio and \( \frac{1}{R'} = \frac{1}{R_1} + \frac{1}{R_2} \).

Obviously the elastic pressure \( P_e \) is

\[
P_e = \frac{W}{\pi a^2} = \frac{1}{1} \left( \frac{\sqrt[3]{W E'}}{R'} \right)^2
\]

Letting the hardness of the flat surface – here the workpiece – be \( H \) and the damage index \( \Psi \) then

\[
\Psi H = \frac{1}{\pi} \left( \frac{\sqrt[3]{W E'}}{R'} \right)^2
\]

or

\[
\Psi = \frac{1}{\pi} \left( \frac{\sqrt[3]{W E'}}{R' H^2} \right)^{\frac{2}{3}} = \frac{1}{\pi} \left( \frac{W}{H^3} \right)^{\frac{1}{3}} \left( \frac{E'}{R'} \right)^{\frac{2}{3}}
\]

If \( \Psi \) is greater than unity then \( P_e > H \) and the diamond scratches. If \( \Psi < 1 \) it does not. There are two points to notice in the equation above. Both are concerned with the parameters in the first bracket. \( W \) has two static components, one due to the weight of the stylus assembly and the force exerted onto the surface by the restoring spring – which is a maximum at surface peaks. There is a dynamic force \( m w \) due to inertia which is a maximum in the valleys. The general reaction is given by \( R \) [5–7].

It is at the surface where energy transfer takes place. It is in this ‘skin’ region where the properties which govern functional behaviour need to be measured. If the ‘hardness’ is measured with two widely differing loads as shown in Figure 4(b) and 4(c), then the bulk hardness value is determined over the large circle of the indentation, shown in Fig. 4(c). This area encloses many dislocations and faults so that the material can flow rather easily. On the other hand if the indentation is nanometric or sub micron the area is small. In this case the indentation dimensions are small compared with the average distance between
faults and the material is reluctant to yield. The effect of this subtle change of scale is to make the ‘skin’ appear to be much harder than the ‘bulk’. Factors of five or even more can be found (Ref.4). The effect of this “extra hard” surface is to make the surface very much less prone to stylus damage than conventional thoughts would indicate.

2.3.2 Instrumental issues

This ‘skin’ effect of materials taken with the considerable advantage of the stylus method in moving aside debris, oils and grease demonstrate that the average stylus instrument achieves a rather delicate balance between measuring what is the real geometry of the workpiece and ignoring what amounts to irrelevant and misleading debris and ‘noise’.

This balance is rarely if ever met with other techniques. Also recent developments in instrument design suggest that the conventional criticisms used because of physical contact are becoming less important. Most important amongst these reasons are the effects of the various dynamic forces of the pick up when traversing. These forces are inertial, damping and elastic forces. Taken together these can produce a surface reaction to the stylus which is not constant as the stylus traverses the surface. This variation in force has in the past been responsible for a number of criticisms of the stylus method.

It is in principle responsible for both the stylus lifting off the surface at the top of a peak and the cause of ploughing in valleys. However, such criticisms are no longer as valid. Two factors have influenced this change of emphasis in recent instrument design. One is the fact that the component forces can now be balanced right up to the maximum frequency range by special design of the pick up in which the damping factor plays an especially important role. This results in a more consistent reaction. The other aspect concerns the inertial term. New pick-up designs are made much smaller than previously were possible. The result of
this is that the inertial term plays a less important role than previously thought possible. The result of these and similar advances have meant that the pick up dynamics are extremely controllable.

Also the effects of the stylus tip shape are now quantifiable (Ref. 8) and specific knowledge of the stylus angle can be used to measure, indirectly, such phenomena as surface free energy (Ref. 9).

3 Alternative Methods

The obvious competitor to the stylus technique is the optical probe sometimes called the optical 'follower' (Ref. 10). This replaces the mechanical stylus with a focused spot on the surface. Such methods are non contacting and as such cannot be criticised for potential surface damage. This is one advantage of the optical method which cannot be denied. Another advantage is the fact that because it is a focused spot it cannot be damaged unlike a stylus which can be abused by mishandling. Obviously it also does not wear. Beneficial as these points may be, the optical technique should be used with caution. The very fact that it does not contact the surface makes it more susceptible to extraneous matter on the surface; it cannot just brush it aside when traversing, it has to measure it.

Also, the optical technique; whatever variant, measures optical path length in one form or another and not geometry as understood in engineering. In many cases there may be closely related but in others they may not. For example the skin of a surface in which there are crystal grain boundaries may well have different refractive indices from point to point. This affects the position of the focus spot because the light penetrates the surface to different depths dependent on the optical inhomogeneity and electrical conductivity of the surface (Fig. 5(a)) (Ref. 11).

![Figure 5: Optical problems](image-url)
Another annoying and potentially misleading physical factor inherent in optical methods is the tendency of sharp points on the workpiece surface to act as secondary sources of light. Unfortunately, if the radius of curvature of a part of the surface is smaller than about 10 μm diffraction effects appear which appear to enhance or brighten the image of the curve. This effect is not serious at large values of surface roughness but is crucial for fine surfaces. It is one reason why optical methods should not necessarily agree with stylus methods for very fine surfaces. In fact they often give bigger answers than the stylus method because of this effect. Another disappointing aspect of optical technique as a follower is the indisputable link between the resolution spot size and the angle of the convergent cone of light focussed on the surface. Both depend upon the numerical aperture (NA) of the system; the spot size is inversely proportional to the NA whereas the angle of the cone of rays is directly proportional to the NA.

Here the numerical aperture is taken to be \( \frac{a}{2f} \) where ‘a’ is the objective lens diameter and ‘f’ is its focal length. Also the depth sensitivity is inversely proportional to NA square in optics and completely independent of the stylus geometry in tactile methods.

These links are not present with the stylus method. So if with an optical follower it is required that the resolution spot should be finer, then the numerical aperture has to be increased. This in turn gives a larger subtended angle as seen in Figure (6 (a)) rather than (b). Hence the ability of the cone of light to see into steep valleys becomes poorer. In a stylus technique this is achieved by making a more acute angled stylus; not easy to do but not impossible. There is no physical impediment as there is in optics. Also another physical link is between the depth of focus and the angle of the cone i.e. the numerical aperture. For acute angles i.e. low NA the technique may be able to measure into steeper valleys but the

\[ d_1 < d_2 \quad \text{and} \quad \alpha_1 > \alpha_2 \]

**Figure 6: Depth of focus problems – slope measurement**
depth sensitivity is reduced at the same time. There is no such restriction with the stylus method. Hence many of the so-called advantages of the optical method are suspect and have to be taken account of when using this technique. The relationship between cone angle, spot resolution, and depth are all connected explicitly in optical methods. This is not the case in stylus methods where all features of the probe can be varied independently which makes the stylus method much more flexible. It is not accidental that the proportion of instrument makers using optical methods is reducing relative to those standardizing on the traditional stylus method.

4 Conclusions

At one time it was thought that stylus methods were on the decline because of problems with contact. In fact the opposite has happened. Most of the new generation of scanning probe microscopes (SPM) including the scanning tunnelling microscope (STM) and the Atomic force microscope (AFM) rely heavily on a mechanical sensor of minute dimension positioned atomically near to the surface (Ref 12). At these levels it is questionable as to the meaning of contact! The ability of the mechanical probe to act as a positive exciter makes it unique in modern sensor technology. The probe manipulates the atomic force and electronic states in the surface in a much more positive way than electromagnetic waves i.e. optics or x rays which tend to get into confusing interactions with the atomic scale surfaces. e.g. Figure 7b. Hence even at the highest levels the stylus method, albeit in disguise, is flourishing. It is paving the way forward because of its positive role in measuring much more than the geometry requirement of the past. Today active consideration is being given to using stylus methods for measuring geometry (or topography) as before but including also physical properties of crucial importance in modern industry such as friction, elasticity, nano-hardness, surface energy and so on. These additional
measurements are almost uniquely possible with stylus type techniques. The stylus technique is not dead nor even dying. It has virtually just been born!

References

(11) Beckman P. and Spizzichino The Scattering of Electromagnetic Waves from Rough Surfaces (Oxford, Permagon) (1963)