Precision Engineering
K. Narayanasamy and Y.G. Srinivasa (Eds)

Surface Metrology 2000+

D.J.Whitehouse

Metrology Consultant Taylor Hobson
Professor of Engineering
Department of Engineering
University of Warwick
Coventry CV4 7AL, UK

Summary

The role of surface metrology is explained. It is shown how it has changed over the years and how it is likely to develop. Emphasis is placed on the underlying philosophy rather than on description of individual instruments. It is shown how the classification of workpiece performance is related to surface typology and manufacturing control.

Key Words: Surface texture, Surface instrumentation, Tribology, Nanotechnology.

1. Why Measure Surfaces

In recent years surface texture has been recognized as being significant in many fields. In particular the surface roughness is an important factor in determining the satisfactory performance of the workpiece, in tribology for example or in coatings. Also in engineering applications the surface roughness has been found useful in machine tool monitoring. These aspects will be discussed presently. It is, however, pertinent to consider how the importance of surface roughness is changing with the passage of time, and how the importance of roughness depends on the actual scale of size of the work-piece and the process used to make it.

In very general terms the requirements for energy transfer and information transfer and storage have dominated the development of technology. This will no doubt also be true in the future but the factors governing such transfer depend themselves on size. As objects get smaller changes take place which highlight the importance of the surface. The energy and force equations experience a change of balance, Figure 1(a) as the scale of size of the moving object reduces.

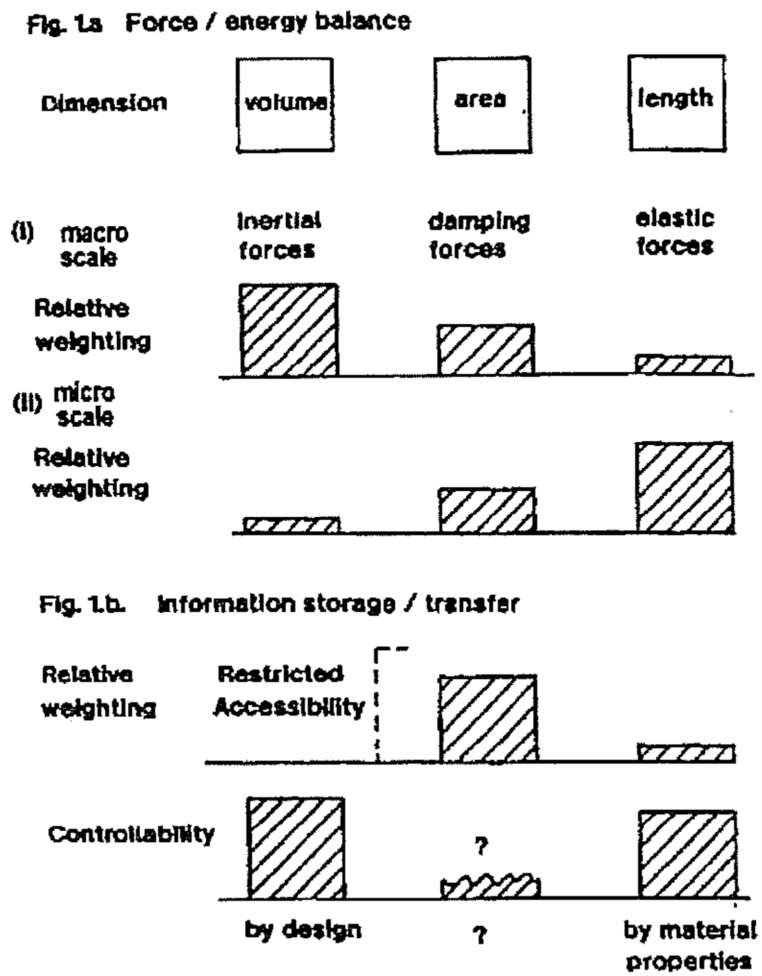


Figure 1 Importance of Surface Properties with Scales of Size.

Information storage in 3D is possible but still very difficult to achieve, also data retrieval is a major problem. On the other hand storage of data on surfaces is still actively being extended. This capability is obviously a function of the surface area. There is no problem with accessibility of the stored data as there is with volume storage. Notice that information storage trends tend to have the opposite sign to the energy equations with respect of the effect of the scale of size, Figure 1(b). In both situations the critical regime is that of area. Consider Figure 1(a). In the force diagram momentum effects quickly decrease as size is reduced, in fact by a factor of L^3. Damping is proportional to area and reduces as L^2. Elastic forces only reduce as a linear factor of L so that they become progressively more important than the others as the scale of size reduces. In rotational situations the dependence on scale is even more pronounced. For example moment of inertia reduces by a factor of L^5.

Of the above factors elastic properties can be controlled by the properties of materials, which are well understood, and inertial forces can be controlled by because of the advent of random process analysis. That is, the use of autocorrelation, power spectra and probability density functions.

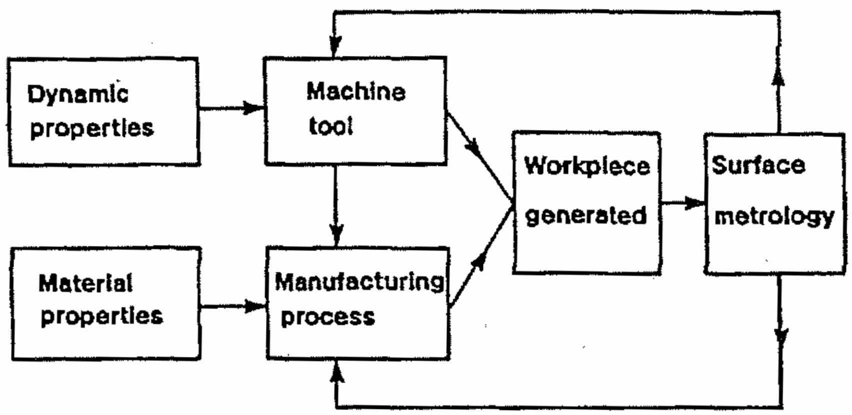


Figure 2 Surface Measurement and Manufacture

These are functions rather than numbers such as the R_a (average Roughness) of the profile and so can reveal much more of the underlying statistics of the surface. They are more reliable because in the case of autocorrelation and power spectral density any random phase shifts between the sinusoidal components making up the profile are eliminated. The autocorrelation function is particularly useful for looking at random surfaces and the power spectrum is more useful for looking at periodic surfaces, as will be seen shortly. Neither are particularly difficult to generate. The autocorrelation function is simply a plot of the correlation coefficient between the surface profile and the same profile shifted in space by a set amount. The Power Spectrum is the Fourier Transform of this.

2.1 Autocorrelation and Manufacture

The reason why these statistical parameters are effective is because they provide a large enhancement of the signal over the noise introduced into the system. For example each point on the autocorrelation function of a profile taken off a surface is a result of a great deal of averaging. Small changes between surfaces became significant. As a general rule the autocorrelation function can best be used to reveal changes in random processes such as grinding whereas power spectral analysis can be used to best advantage in processes which are fundamentally periodic or repetitive as in turning or milling. Both the autocorrelation function and the power spectrum hunt for the unit machining event. In the case of grinding the unit event is the impression left on the surface by an average grain on the grinding wheel. In power spectral analysis it is the periodic signal left on the surface by a clean cutting tool on a perfect machine.

2.2 Power Spectral Density in Manufacture

Another example shows how the power spectrum can be used to identify problems in turning. As the tool wears and the machine tool deteriorates significant changes occur in the spectrum of the surface as shown in Figures 4 and 5.

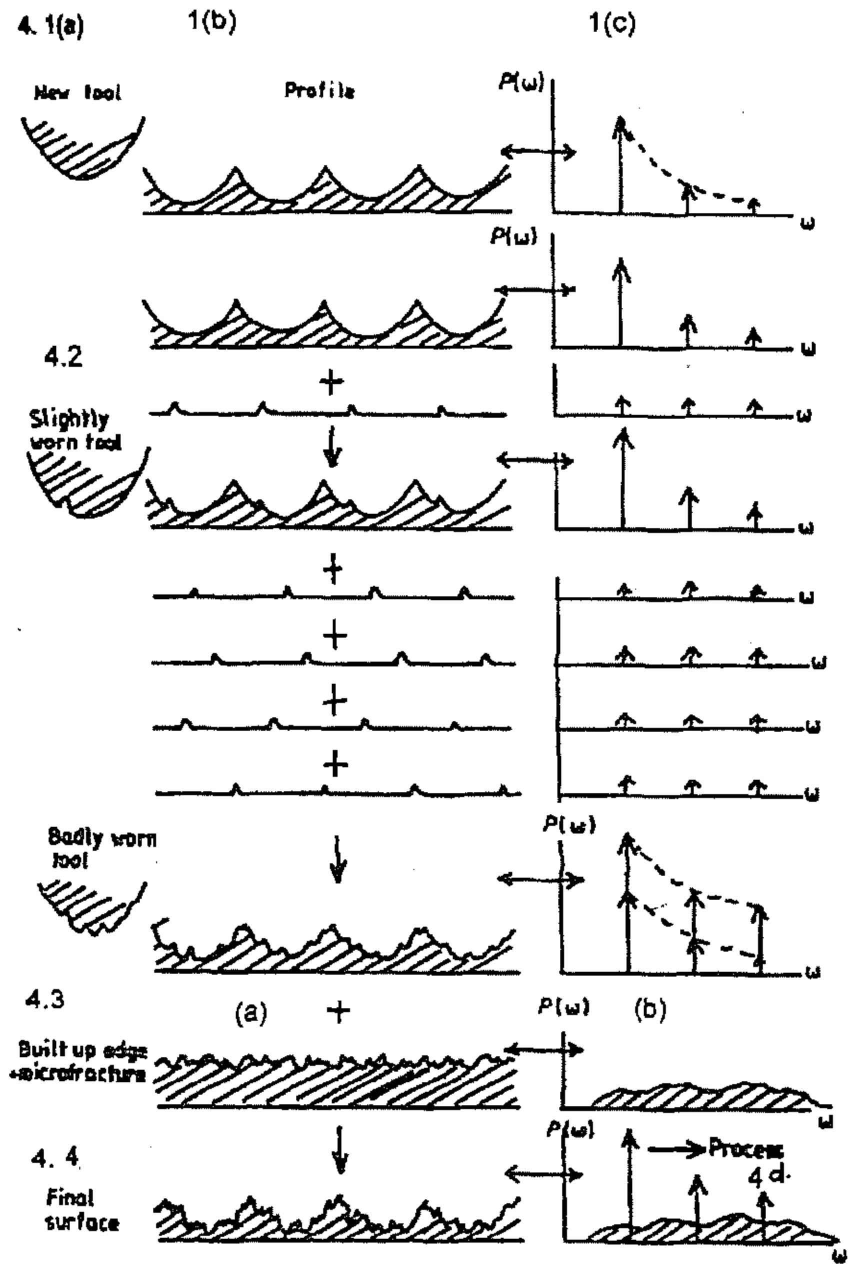


Figure 4 Power Spectral Analysis Harmonics

2.2 Space-Frequency Functions

To the left of the fundamental wavelength periodicities appear whose wavelength is much greater than that of the fundamental. These are due to machine tool problems such as bearing wear, slideway error or even lack of stiffness in the machine tool itself which may cause chatter. Identifying these effects by using the surface texture is an important first step in remedying the problem.

The spectrum can therefore be split up into two parts, one to the right of the fundamental frequency and one to the left, Figure 5.1 - 5.5. On the right appears process problems and on the left, in the sub-harmonic region, machine tool problems. These advances in machine monitoring and diagnostics stem from the realization that the surface generated by the manufacturing process constitutes a very extensive data bank of information. The surface is in effect a fingerprint of manufacture.

From what has been said above it could be argued that autocorrelation and power spectrum are sufficiently comprehensive to be able to cater for most eventualities in manufacture. Unfortunately this is not so. There are instances where subtle changes in the surface geometry can be very important in machine tool monitoring. For example the mode of vibration of a tool column determines the changes in statistics of the surface geometry. Sing Power Spectrum or autocorrelation cannot detect changes in the statistics. Unfortunately changes in the nature of the surface often accompany the presence of defects, flaw etc. which are detrimental to the performance of the workpiece.

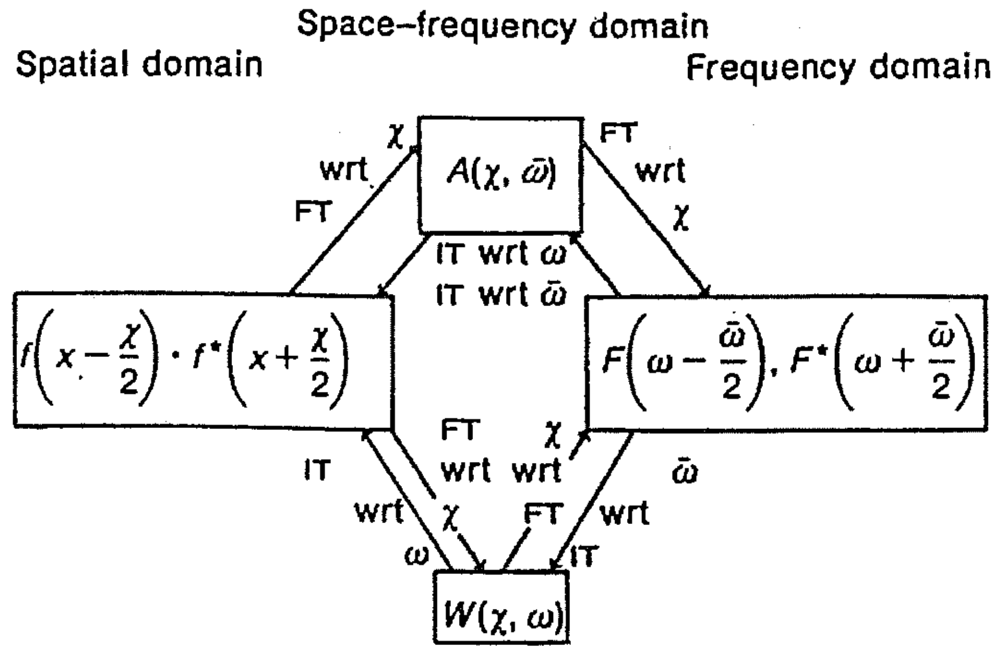
Looking at the formulae for autocorrelation $C(\tau)$ or power spectrum $P(w)$

$$C(\tau) = \frac{1}{2}\int_{-L/2}^{L/2} f(x) f^{*}(x+\tau)\,dx \qquad (1)$$

$$P(w) = \left|\int_{0}^{\infty} f(x)\exp(-jwx)\,dx\right|^{2} \qquad (2)$$

show that the integral limits extends over all the signal $f(x)$. The whole signal is integrated. Any change within these limits in the nature of $f(x)$ will simply be averaged out.

This restrictive behaviour of the random process functions has been recognized for some time and measures have been taken to remedy the situation by modifying the definitions. In effect there has to be introduced a window function in either time or frequency which localizes the extent of signal $f(\cdot)$ which is being examined. The width and shape of the window function together with its



Relationship between functions.

Figure 6 Relationship between Functions

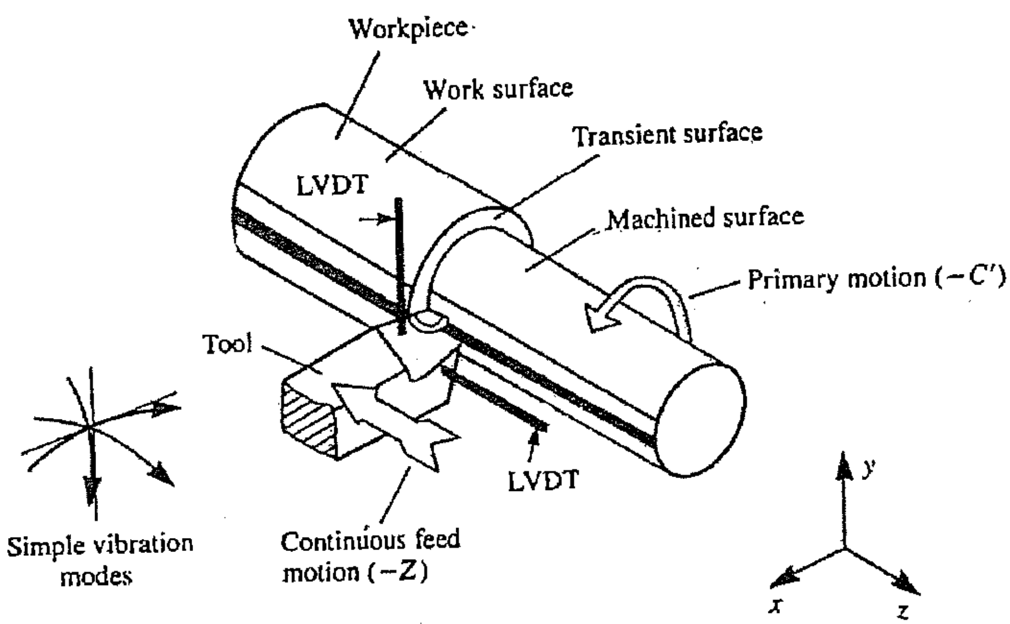


Figure 7 The position of the cutting tool and its vibration modes relative to the workpiece

Figure 8 shows a simplified effect of tool vibration on surface profile.

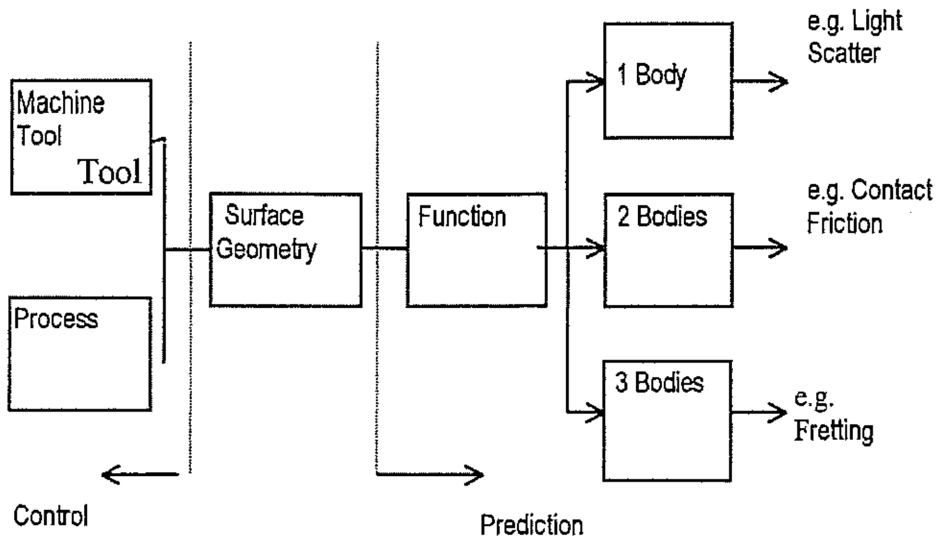


Figure 9 Role of Surface Geometry

Most of the uses above involve contact and as will be seen even the sophisticated methods mentioned above are found lacking. This area is where activity 2000+ should be focussed.

3 The Surface and Function

The surface is obviously important in many practical situations. This has been known for years. The problem is how important? For many years very simple parameters were used to describe the surface and hopefully its properties. These included R_a the average value, R_q the rms value and various peak height estimates. Investigators in Germany and Russia used peak measurements rather than average because they argued that peak measurements correlated more with tribological situations than average values [9]. Also the peak measurements of roughness could be measured equally well with optical and stylus methods. This philosophy proved to be practically unsound.

The notion of being able to predict the performance of a work-piece from the geometry of the surface has been attractive for some time. L Early investigators used simple models of the surface. These usually involved modelling peaks on the surface as hemispherical spheres scattered on a plane. Then these hemispheres or 'bosses' were assumed to be distributed in a random Gaussian way in height [10].

This development was closer to real surfaces than previous ones but it had the disadvantages that two of surface description were needed, one deterministic to conductivity. Neither random process analysis nor space frequency functions help.

Even in the definition of a peak there are problems which have to be addressed when considering functional behaviour. Communication theory simply counts the number of peaks at a given level above the mean line or any other arbitrary height datum. The resulting probability density is derived from the count. Recent work has queried this approach [13]. Using the profile for simplicity and letting $p(y, y', y'')$ be the joint probability density of height y, slope y' and curvature y'' along the profile the standard way of determining the peak height – as in finding the average peak voltage is $\hat{y}_c$ where

$$\hat{y}_c = \frac{\delta y' p'(0) \int_{-\infty}^{\infty} \int_{-\infty}^{0} y . p(y, y'') y'' \, dy'' \, dy}{\delta y' p'(0) \int_{-\infty}^{\infty} \int_{-\infty}^{0} p(y, y'') . y'' \, dy'' \, dy} \qquad (5)$$

Taking the actual curvature of each peak into account gives a similar but different value of average height. Thus, the new height $\hat{y}_p$ is given by equation 6 (with limits removed.)

$$\hat{y}_p = \frac{\delta y' p'(0) \iint y . p(y, y'') dy'' \, dy}{\delta y' p'(0) \iint p(y, y'') dy'' \, dy} \qquad (6)$$

The two equations look similar but the former gives a much higher value than the latter. In the special case where the surface is Gaussian $\hat{y}_c = \frac{\pi}{2} \cdot \hat{y}_p$ which is a completely different height!

Following on from this, other well established parameters such as the material ratio curve *MR(y)* given by

$$MR(y) = \int_{y}^{\infty} p(y) dy$$

are more realistically re-stated in terms of peaks weighted by their curvature (which is a much better predictor of load carrying capacity).

Thus if *PPR* is the peak probability ratio it can be expressed in terms of *MR(y)* area has to be taken into account. These attempt to describe the 'lay' of the surface.

Amplitude Parameters
Root-mean square deviation of the surface (μm)
Ten point height of the surface (μm)
Skewness of the surface
Kurtosis of the surface
Spatial Parameters
Density of summits of the surface (mm^{-2})
Texture aspect ratio of the surface
Fastest decay autocorrelation length (mm)
Texture Direction of the surface (deg)
Hybrid Parameters
Root-mean square slope of the surface (μm/μm)
Arithmetic mean summit curvature (μm^{-1})
Developed surface area ratio (%)
Functional Parameters Characterizing Bearing and Oil Retention Properties
Surface bearing index
Core Oil Retention Index
Valley Oil Retention Index
Material volume ($\mu m^3/mm^2$)
Core valley volume ($\mu m^3/mm^2$)
Deep valley volume ($\mu m^3/mm^2$)

Table 1 Primary set of 3D Surface Roughness Parameters

Texture aspect ratio for example is the ratio of the shortest correlation length on the surface to the longest. Here the correlation length is taken as that length over which the autocorrelation function falls to an agreed low value, usually 1/e but sometimes 10%. This parameter has no equivalent in the profile. The direction in which the correlation length is a minimum is usually taken as zero angle. This has to be referred to some arbitrary angle reference on the workpiece.

The second group are called functional parameters and are unashamedly contrived to satisfy some functional need e.g. oil retention or load carrying capacity.

The actual definitions of these parameters are usually linked to the material ratio parameters They can be obtained from the literature [15] reference to Euro report 15178 EN but it has to be emphasized that these functional parameters and the areal (3D) ones are not standards, they are suggestions.

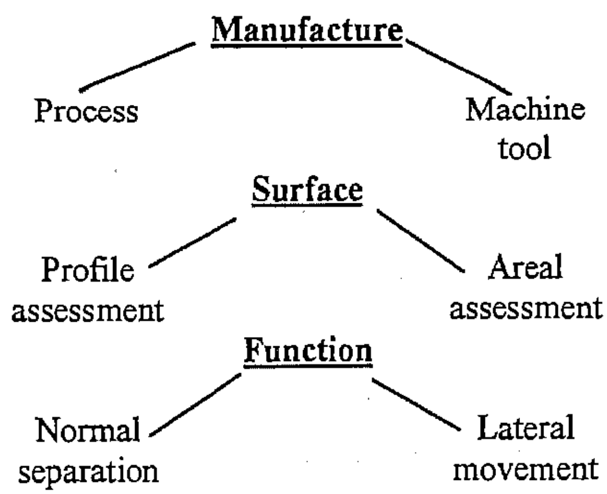


Figure 12 Basic Characterization

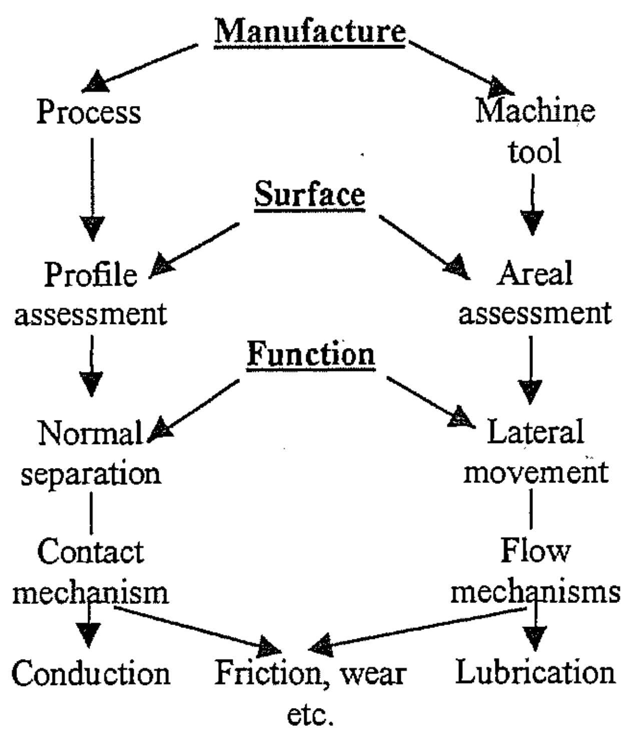


Figure 13 Elementary Links in Characterization

Obviously this is a considerable simplification but it does concentrate emphasis. Also, the surface metrology is fulfilling two different roles as far a function is concerned. One is to ensure good performance and the other is to detect sources of failure. These two jobs require different aspects of the measurement. As a

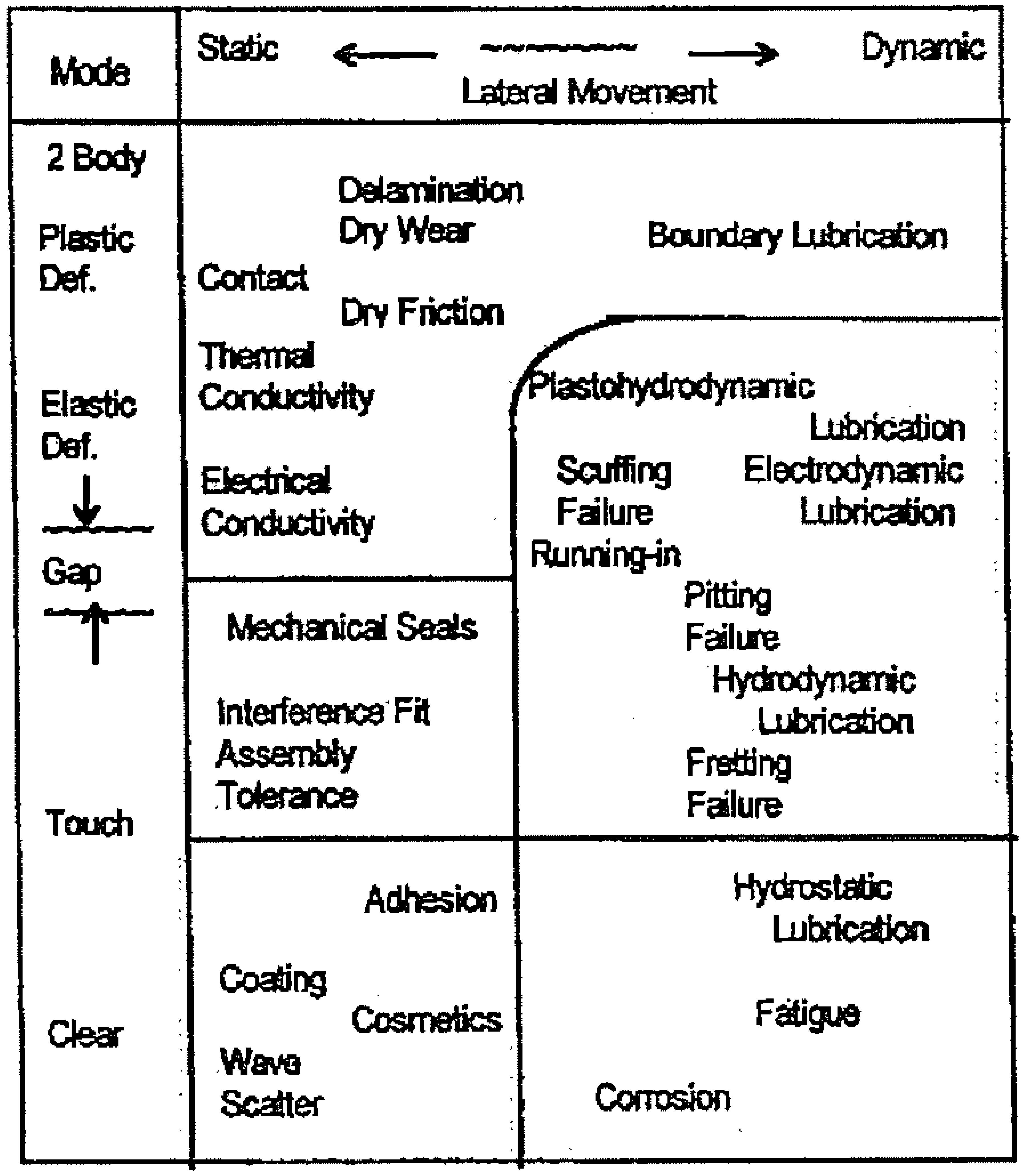


Figure 15 Function Map

which is where the distributions of height of the two surfaces are comparable with the separation.

Summarizing, there is now a possibility of weaving a coherent thread linking manufacture metrology and function. This thread starts with the function map whose format is outlined above. Superimposed on this map is the surface metrology template (or dimensional template) whose axes are static parameters e.g. peaks usually from a profile, against 'dynamic parameters' which involve slopes and curvatures and have to be areal. The manufacturing template has 'process' and machine tool as axes. This is superimposed onto the map and metrology template.

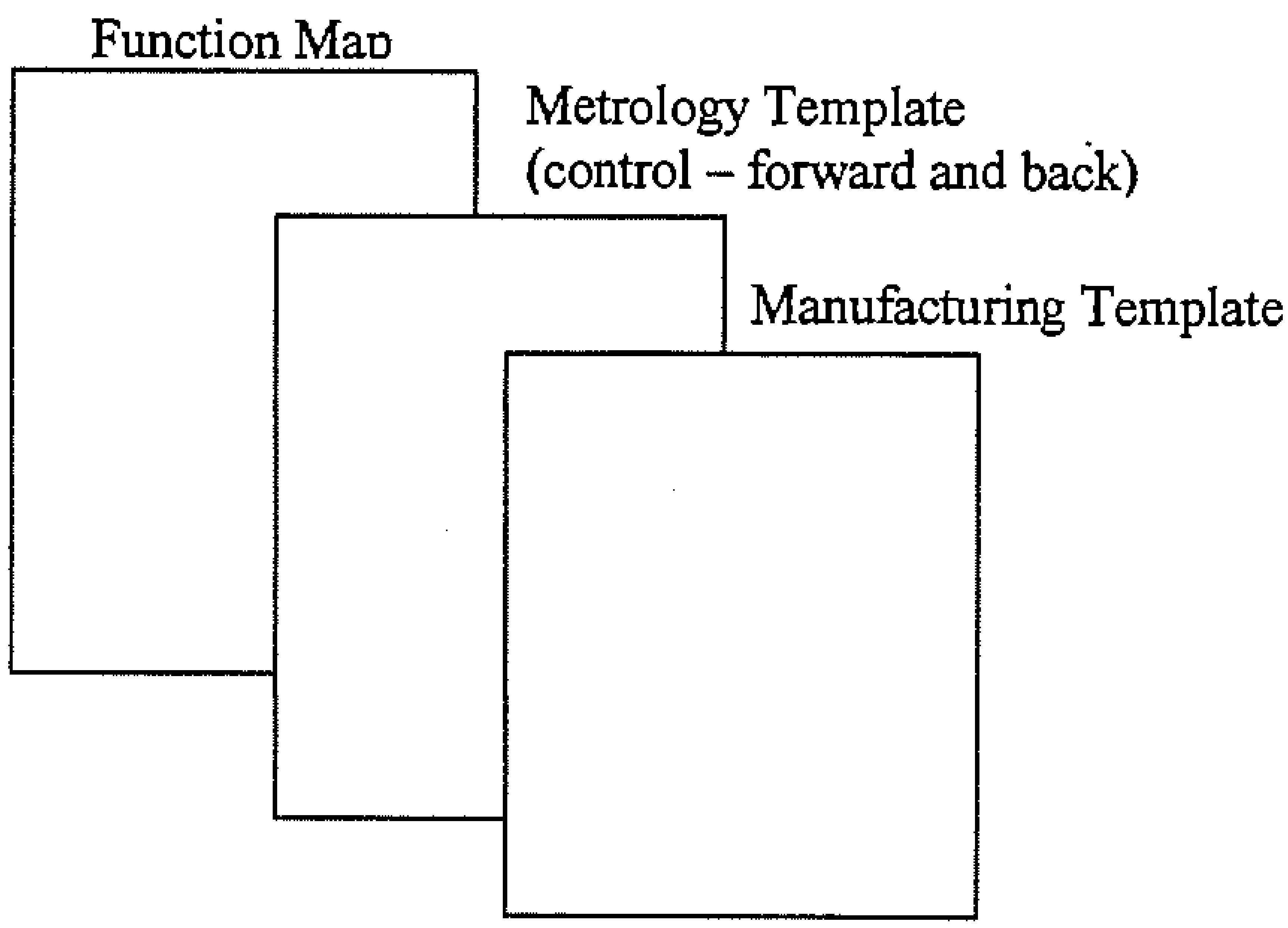


Figure 17 Arrangement of Templates – The Template Stack

Ideally reading through the template stack should give the best surface and manufacturing process for a particular function. This is shown pictorially above but would be carried out from stored computer data bases.

A further step which is even now a possibility is to forget the parameters of the surface and carry out a 'pilot' experiment in the computer to see if the workpieces work together. This involves areal mapping of both surfaces comprehensively using a tactile sensor and them literally making them contact and rub by simulation. It may be that this is the best way forward. Notice that the tactile sensor is suggested as the preferred instrument. The reason for this is that in the tradition of metrology the measurement should mimic as nearly as possible the function. Because most applications involve contact and rubbing the tactile stylus method should be used. For non contact applications obviously optical or other methods are to be preferred.

11. Archard, J.F. *Elastic Deformation and the Laws of Friction* Proceedings Royal Society London A **243** (1957) 190-205

12. Whitehouse, D.J. and Archard, J.F. *Discrete Properties of Random Surfaces of Significance in their Contact* Proceedings Royal Society London A **316** (1970) 97

13. Whitehouse, D.J. *Some Theoretical Aspects of Surface Peak Parameters*. Precision Engineering **23** (1999) 94-102

14. Stour, K.G. *Development of a Basis for 3D Surface Roughness Standards* EC Contract No. SMT4-CT98-2256

15. Euro Report 15178EN

16. Scott, P. *Areal Topological Characterization for a Functional Description of Surfaces*. European Surfaces Workshop Corps la Salette June 1998

17. Radhakrishnan, V. *On an Appropriate Radius for the Enveloping Circle for Roughness measurement in the E System* Annals CIRP **20** No. 1 (1971)