

Data Processing Techniques for 3D Surface Morphology

¹Ramji Singh , ²Dr. Amit Kumar , ³Vatsalya

¹M. Tech, Mechanical Engineering, O.S.D. (Examination), Aryabhatta Knowledge University, Patna – 800001, Bihar (India)

²Assistant Professor, Department of Mechanical Engineering, National Institute of Technology Patna, Patna – 800005, Bihar (India)

³Research Scholar, Department of Computer Science, Birla Institute of Technology Patna Campus, Patna – 800014, Bihar (India)

-----ABSTRACT-----

It is known from experimental trial and error that certain surfaces outperform others in certain applications, specifically; if the film thickness in a lubricated contact is of the same order as the surface roughness then the surface topography plays an important role in determining the tribological response of a precision finished surface. Various 3D parameters have been defined based on different properties like amplitude properties and spatial properties required for describing the surface. These parameters have been successful in analysing various problems in industrial and research field, related to surface characterization. This data is acquired using stylus based contact methods or interferometry based optical methods. The computation of these parameters require considerable amount of calculations. Therefore a tool is required which can calculate these 3D parameters for various precision finished surfaces. The objectives identified for the current work include processing the surface topography data and converting it into an array of surface heights from the mean plane, developing a hole filling algorithm to fill gaps in the surface data and development of codes in MATLAB for computation of 3-D amplitude, spatial and hybrid parameters.

KEYWORDS: surface, topography, texture, parameters

Date of Submission: 31 December 2014



Date of Accepted: 20 January 2015

I. INTRODUCTION

It is known from experimental trial and error that some surfaces perform better than others in wide range of applications due to certain surface characteristics. This is more conspicuous if the film thicknesses are of the same order as the surface roughness in lubricated rolling/sliding point and line contacts then the surface topography plays an important role in determining the tribological response of a surface. It has been observed by Singh et al. [5-6] that surface roughness alone cannot capture the effect of friction in precision finished surfaces. The surface texture of the bio-medical implant affects the biocompatibility, hemodynamics, tissue integration when implanted in the human body. In case of cardio-vascular implant the platelet activation depends on the shear stresses which are a function of surface topography. Recently, there has been a push from both industry and academia towards the use of 3-D roughness parameters due to the fact that surfaces interact in three dimensions. 3-D surface roughness standards have been developed by ANSI and ISO but these only capture the amplitude parameters. The amplitude parameters capture only the deviation from the mean plane and/or line. In order to capture the surface geometry we need spatial parameters as mentioned by Dong et al. [3-4]. This work aims at developing a tool and characterizing select precision finished surfaces in 3-D. Obtaining the data involves extracting the data of surface in (x,y,z) form by stylus based or optical based experiments or by generating the data. This data is then properly arranged to obtain the 3D information about the surface. The x and y data need to be sampled at regular interval along the respective direction within the data range. If any point is missing then there can be a possibility of a hole in our data which could be due to errors in the measuring instrument or surface topography. This hole needs to be filled in a suitable manner to maintain the consistency in the data. One of the methods could be filling the gap point by assuming it to be an average of the 8 neighbouring points. As seen in Figure 1.1, some of the surface data is missing which needs to be filled or eliminated. In addition, the surface has a form or curvature which needs to be filtered and roughness data needs to be extracted.

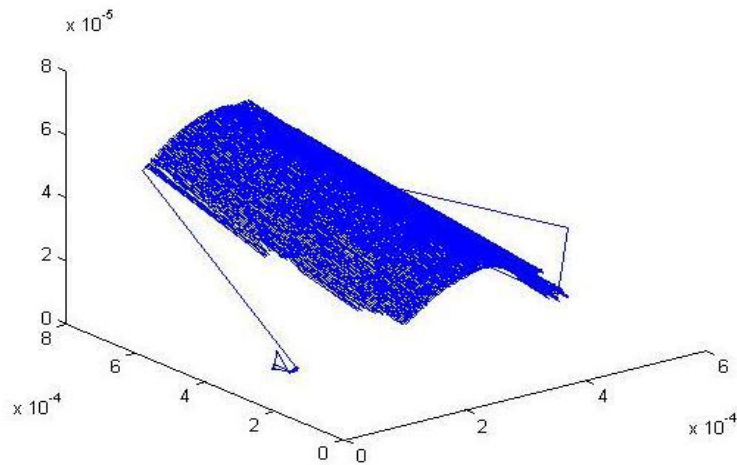


Figure 1.1. The plot of surface data obtained through experiment in 3D form on Matlab®.

II. DATA ACQUISITION AND MEAN PLANE FITTING

Data Acquisition : The data has been acquired from Zygo NewView 200 at Georgia Tech and Veeco WYKO NT910 0 white light interferometer at IIT Bombay. Figure 2.1 shows an image of the white light interferometer.

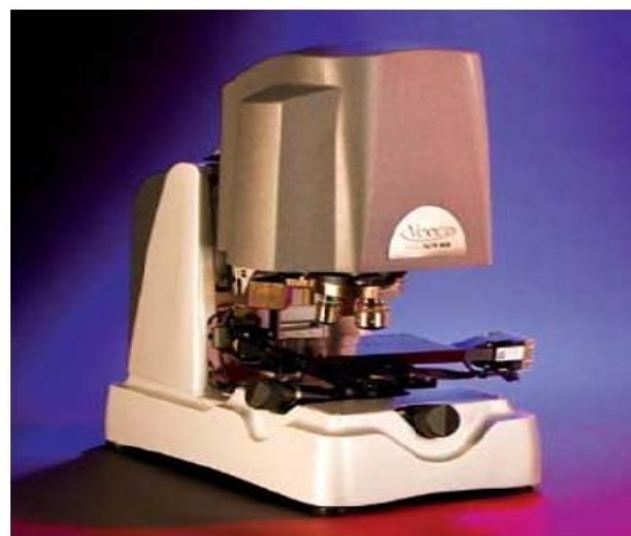


Figure 2.1. White light interferometer Veeco WYKO NT910 00.

The data obtained from the experiments can sometimes lack data points at certain locations called holes. Note that the data is obtained in the form of 3-column X-Y-Z ASCII file. For computation of 3-D parameters this data needs to be first filled in a matrix and then holes need to be filled and if the surface has a form it needs to be filtered. The step-by-step methodology is as follows:

- [1] Read the x, y, z data.
- [2] Generate a matrix the elements of which contain height of points so that the acquired data can be treated as a surface.
- [3] Figure out holes in the data. (Holes are data points which are missing from the data set)
- [4] Remove boundary holes and extract the most meaningful of the surface data
- [5] Fill in the remaining holes as accurately representing the original surface as is possible
- [6] Fit a surface to the data that represents its form. Examples of surfaces are planes, bilinear surfaces, second order surfaces etc.
- [7] Evaluate the residual surface. Hence compute all the parameters

The surface data that has been obtained by white light interferometry is in the form of $n \times 3$ matrix in an ASCII file that contains the x,y,z coordinates of the machined surface, where n is the number of data points. Now, typically this is a huge number hence holes in the data if any are hard to detect at this stage. Hence this data is loaded into a matrix as the first step of the code. The x and y vectors are scanned for their minimum values (as this value is the starting point of all measurements). Using these values and the difference between two adjacent coordinates (x, y), the x and y vectors are normalized. Thus, each row of the loaded matrix is now a triplet with its first two elements giving integer position of a point whose height is given by the third element. Now, a data matrix is created which stores heights of all measured points. While defining the matrix, it is initialized to a zero matrix of number of rows and columns equal to the range of values of the normalized vectors x and y. Then each recorded data point is fed into this zero matrix exactly at the index given by x and y value of the respective point. This data matrix thus accurately represents the measured state of the surface. Different heights are represented by different colors. Blue color corresponds to zero height. Thus all holes correspond to blue colored portion of the surface. There are two different types of holes. Those that exist at the boundaries of the acquired data image and those that exist in the middle of the image. The number of boundary holes is large and they occur quite frequently. Figure 2.2 shows a typical image of a surface with a profile.

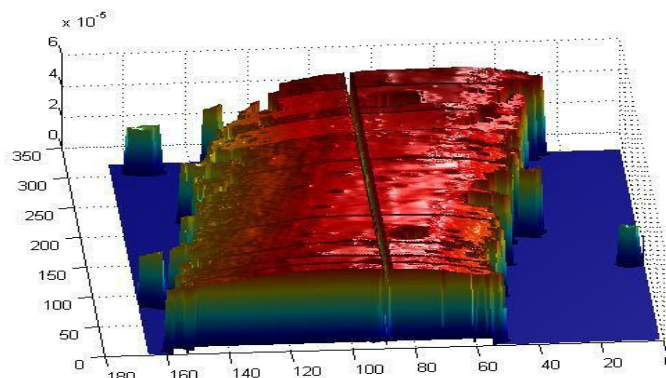


Figure 2.2. Acquired surface data points

It can be seen in Figure 2.2 data acquired near the boundaries is not reliable due to some inadvertent end effect errors encountered in the measurement. A portion of data thus has to be extracted from this surface that has fewer holes that lie in the interior portion of the data. This is discussed in the next section.

Edge Determination : Useful data exists in the form of clusters isolated by a sea of holes. Only at the centre of the image can sufficiently large number of useful points can be gleaned in order to proceed.

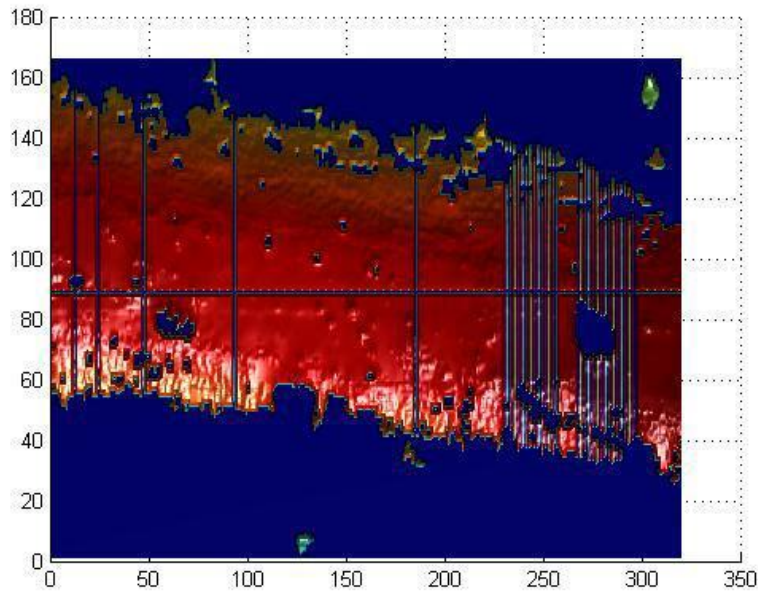


Figure 2.3. Edge determination algorithm

The *bwtraceboundary* feature in Matlab can be used to trace the boundaries of data pools. The algorithm to implement the function involves conversion of surface image to a binary image (a matrix that contains only 0s and 1s). This is done by assigning a value of 1 to any entry that is not zero. This binary image is then scanned from the bottom left corner for a 1. Whenever a 1 is encountered, the function searches connecting elements that are also 1s but contain at least one 0 as a connecting element. If in the process of determining the current boundary the function manages to obtain a closed loop, it stores the coordinates of the boundary in a matrix called contour. The scanning process goes on for the entire image. When plotted, all possible boundaries appear as shown in Figure 2.4.

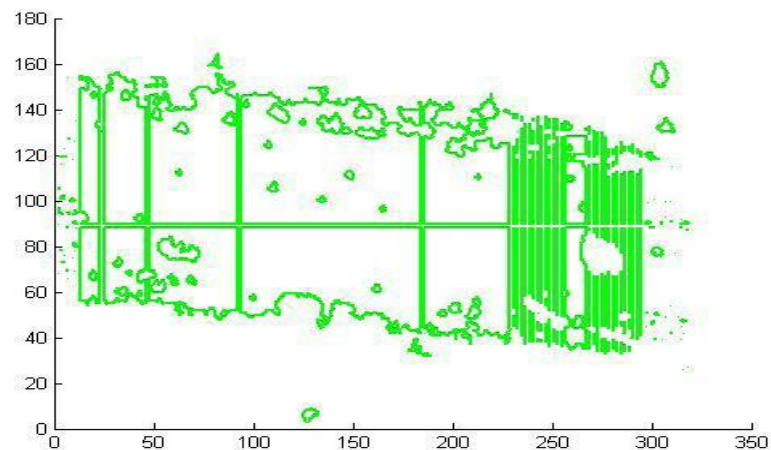


Figure 2.4. Boundaries of data clusters

The green closed loops are the data boundaries. This data set has a large number of holes and as a result data is cluttered. The main purpose of edge determination is to get rid of the boundary holes. Thus, it is desired to cut o the data boundaries so as to keep maximum data for computation. There are some features in the image processing toolbox of Matlab such as the ROI (region of interest) tools that can be useful here but considering the cluttered nature of the data and the large number of boundaries, the following algorithm is used.

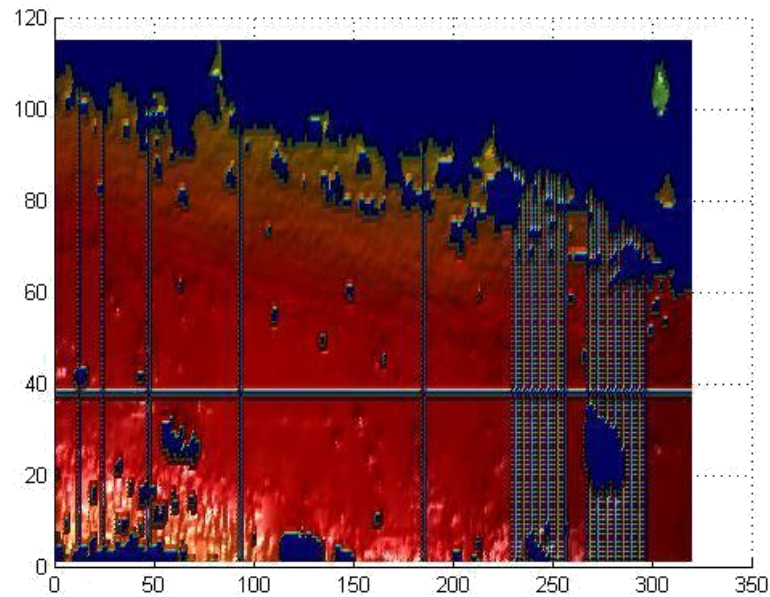


Figure 2.5: Holes removed from one side

Beginning with the bottommost row, every row is scanned and the number of holes is recorded for each row. This number is then compared with the maximum size of the row. The number of holes in each row typically decreases steadily towards the interior. Suppose the number of zeros in the i^{th} row falls to less than $1/4$ the size of the row (or any other suitable tolerance threshold), all rows to the top of the i^{th} row are retained as useful data. Similarly the matrix is inverted and the same procedure is repeated for the other side. The result of the first iteration can be seen in Figure 3.5 and the result of two reductions can be seen in Figure 2.6.

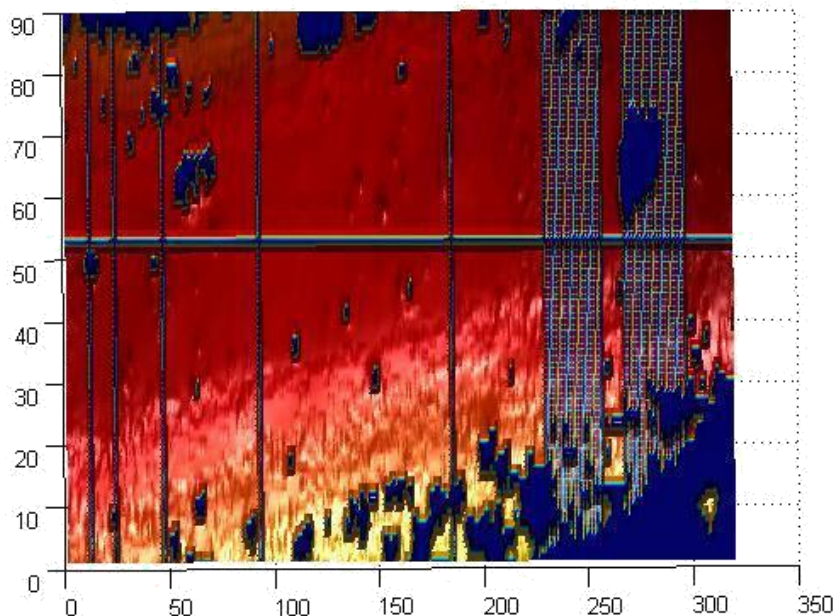
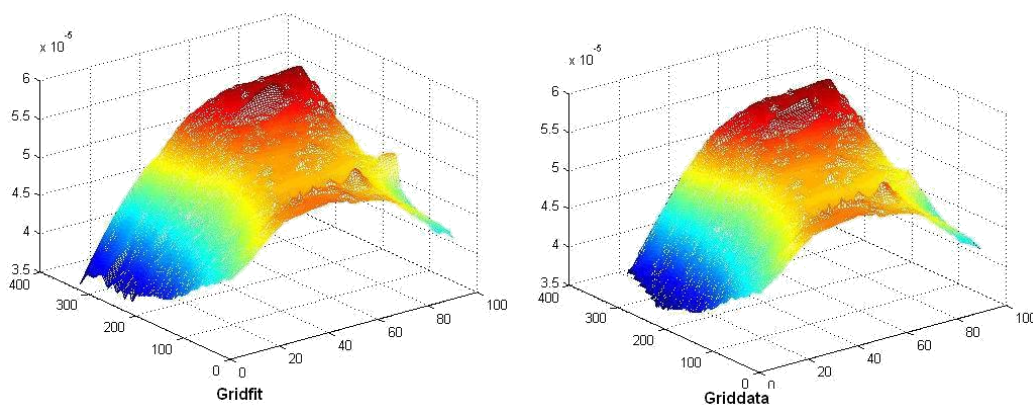


Figure 2.6. Holes removed from both sides

Hole Filling : At this stage useful data has been cut along boundaries such that almost all the boundary holes have been eliminated. However, local holes still remain. However sophisticated the stylus or optical measurements be, there is always a possibility of interior holes, which have to be dealt with. As seen in Figure 2.6, there are horizontal and vertical blue lines that cut across the data. These are extended lengths wherein data is absent. It is not possible to remove this data altogether as it comes at the cost of losing continuity of the

surface. Thus, the best that edge determination algorithm can do is remove undesirable boundary points. There are two ways to proceed. The remaining holes can either be kept as they are or they may be called using fitting techniques. In both alternatives, it is possible to fit a second order surface as form for the data. However, in the former case, data points that read 0 should have actually been something else. If this anomaly is unaccounted for, the processed data surface would not be a faithful representation of the true surface. All fitting functions use interpolation or extrapolation techniques. Hence it is important to plug these holes before fitting any kind of form in order to avoid a 0 value getting interpolated or extrapolated. The functions Gridfit and Griddata can be used for this purpose. The Griddata function is an interpolation function while the Gridfit function is an extrapolation function. The basic difference is that Gridfit can smoothly extrapolate beyond the convex hull of the data. Convex hull is the surface that just encloses all the data points within the given boundary. In other words an interpolation function like Griddata takes the value of the function at a given grid exactly which may lead to discontinuities whereas Gridfit captures the true nature of the surface and is smooth approximation. The appearance of the surface after hole filling is done separately by the two function is given in Figure 2.7

Figure 2.7. Surface with holes called using Gridfit and Griddata



There is not much difference visually between the two functions however as demonstrated Figure 2.7, the difference between Gridfit and Griddata is a tad positive. This is because because of the nature of Gridfit, it can go beyond the values of the four nearest neighbours. Griddata relies on interpolation techniques such as four nearest neighbours which give average value to an intermediate point . However, Gridfit can capture a peak if there is a peak in between four nearest neighbours. Hence it tends to overestimate the surface as compared to Griddata. Another thing worth noticing is the boundary distortion effect. The data matrix after hole removal also has some holes at the boundary invariably.

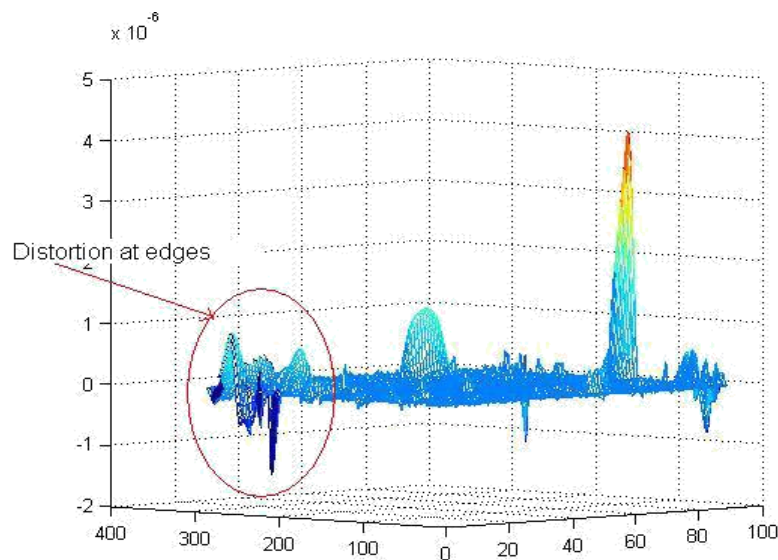


Figure 2.8. Surface with holes filled using Gridfit

These zeroes distort Griddata to some extent but do not affect the Gridfit function appreciably. Considering all the aforementioned factors, Gridfit is found to be a reliable tool for hole filling and the residual surface is calculated using this function.

Obtaining the form : Now that the data matrix is devoid of erroneous data, form of the surface can be obtained. Form is nothing but the least squares plane through the data. In this case, the curvature of the data as seen in earlier plots makes it necessary to fit a second order curve to it. Thus a least squares second order surface should be a good estimate of a form. This surface is typically of the general second order polynomial given by,

$$f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2 \quad (2.1)$$

The two suffixes at the end of constants are the exponent of x and y respectively with which the constant is associated with. The form could be found by estimating the coefficients: a_{00}, \dots, a_{02} such that the sum of squares of height of the surface from this second order surface is minimized. This is done by a MATLAB function Polyfitn. Polyfitn can fit polynomials of any index n on 1D or 2D data. Thus, for a surface data, it fits a least squares surface. If $n = 2$ is specified, it fits a second order surface. The higher order surfaces are computationally intensive to implement. Consequently, it is convenient to work with order 2 as long as it is accurate enough. Calling this function with the help of a variable stores the values of coefficients in a vector. The order of storage is configurable. The form surface is obtained by actually calculating the value of $f(x, y)$ for all (x, y) . The surface plot of this form is shown in Figure 2.9. Polyfitn also lists the standard deviation and RMS error in the calculated values of constants. Thus if these values are found to be too large to be neglected, a different order of surface can be chosen.

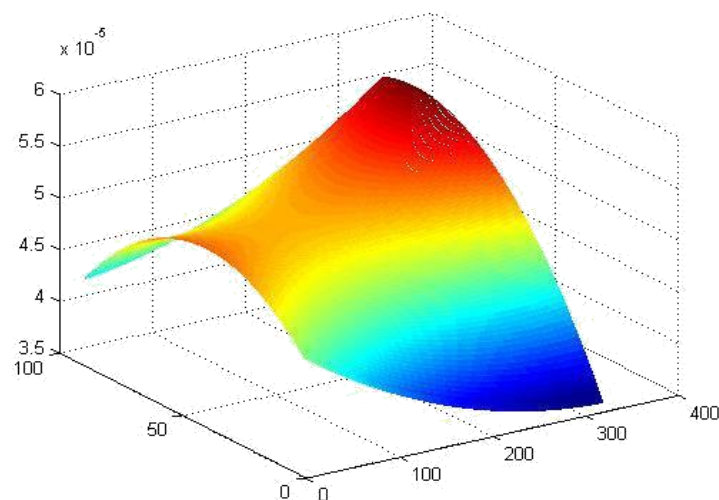


Figure 2.9. Form removed from data points using polyfitn

Residual Surface : As the final piece of data refinement before the actual evaluation begins, the residual surface is determined. Residual surface by definition is the remainder with form is removed from surface data. All the 14 parameters proposed by Dong et al. [7] are a function of this residual surface. Hence, the accuracy of the whole surface characterization process is heavily dependent on this piece of data. The residual surface gives a qualitative feel of the surface even on visual inspection. Factors such as randomness, symmetry and smoothness can be gauged by looking at it. Mathematically,

$$\eta(x, y) = z(x, y) - f(x, y) \quad (2.2)$$

Where, $\eta(x, y)$ is the residual surface function. $z(x, y)$ and $f(x, y)$ are already known from previous sections. Thus the residual can be easily calculated. Its surface plot is shown in Figure 2.10. It has a distinct lay pattern roughly towards the north eastern direction, north being the axis labelled $(0, 100)$. There are a few pronounced peaks which could be due to actual imperfections in machining or some error during data adjustment. For surface characterization purposes, it is desirable that any abnormal peaks be clipped. This adjustment can be made in the final calculations. As a further safety measure, it could be useful to filter boundary data again using

the *Imageclearborder* function from the Image Processing Toolbox of MATLAB.

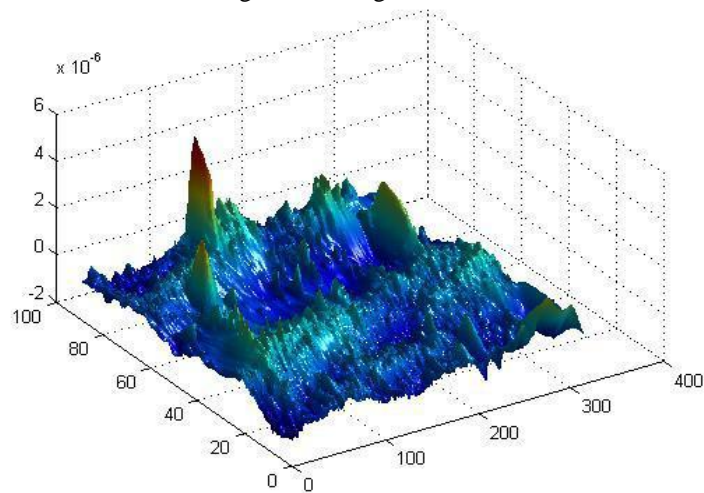


Figure 2.10. The residual surface

The residual surface is used for computation of 3-D parameters presented in the next chapter. Various types of surfaces are explored and the surface characteristics are investigated based on 3-D surface analyses.

III. CONCLUSION

This work is focused on 3-D surface topography of precision finished surfaces. A methodology of getting the surface roughness array and hole filling has been developed. The 3-D root mean square roughness height, S_q , and the density of summits, S_{ds} , can be used to quantify the influence of surface texture on functional response. The density of summits is higher when the surface is dominated by high frequency features or the surface features are spaced closely as opposed to turned surfaces where the feed marks are relatively far apart. AACF and APSD analyses can be effectively used for texture analysis. Efficient form removal and hole filling algorithms can improve the results significantly.

REFERENCES

- [1] ANSI, Surface Texture: Surface Roughness, Waviness and Lay, American Standard ANSI 8.46.1, 1985
- [2] ISO, Surface Roughness - Terminology - Part 1: Surface its Parameters, International Standard ISO 428711, 1984.
- [3] W.P. Dong, P.J. Sullivan and K.J. Stout, Comprehensive study of parameters for characterizing three-dimensional surface topography - III: Parameters for characterizing amplitude and some functional properties, **Wear**, 178 (1994) 29-43.
- [4] W.P. Dong, P.J. Sullivan and K.J. Stout, Comprehensive study of parameters for characterizing three-dimensional surface topography - IV: Parameters for characterizing spatial and hybrid properties, **Wear**, 178 (1994) 45-60.
- [5] R. Singh, S. N. Melkote, and F. Hashimoto, Frictional Response of Precision Finished Surfaces in Pure Sliding, **Wear**, 258 (2005) 1500-1509.
- [6] R. Singh, R. C. Kalil, S. N. Melkote, and F. Hashimoto, Correlation of 3-D Precision Machined Surface Topography with Frictional Response, **Proceedings of IMECE 04: 2004 ASME International Mechanical Engineering Congress and Exposition**, IMECE2004-59534, Anaheim, California, USA, MED, v 15, 2004, 871-879.
- [7] W.P. Dong, E. Mainsah, K.J. Stout and P.J. Sullivan, Three-dimensional surface topography - Review of present and future trends, in K.J. Stout (ed.), **Three-dimensional Surface Topography: Measurement, Interpretation and Applications, A Survey and Bibliography**, Penton, 1994, pp. 65-85.
- [8] W.P. Dong and K.J. Stout, Two-dimensional fast Fourier transform and power spectrum for surface roughness in three dimensions, **Proc. I. Mech. E.**, Vol. 209, pp. 381-391, 1995.
- [9] W.P. Dong, E. Mainsah, K.J. Stout, Reference planes for the assessment of surface roughness in three dimensions, **Int. J. Mach. Tools Manufact.** Vol. 35, No.2, pp. 263-271, 1995.
- [10] D.J. Whitehouse and J.F. Archard, The properties of random surfaces of significance in their contact, **Proc. R. Soc. London, Ser. A**, 316 (1970) 97-121.
- [11] T. R. Thomas, **Rough Surfaces**, Imperial College Press, 1999