Gaussian process machine learning-based surface extrapolation method for improvement of the edge effect in surface filtering

Ming Yu Liu[a,b], Chi Fai Cheung[a*], Xiaobing Feng[b,c], Lai Ting Ho[a], Shu Ming Yang[d]

[a]State Key Laboratory of Ultra-precision Machining Technology, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong
[b]Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom
[c]School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China
[d]State Key Laboratory of Mechanical Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an，China

[*]corresponding author: benny.cheung@polyu.edu.hk

Abstract

Filtering for signal and data is an important technology to reduce and/or remove noise signal for further extraction of desired information. However, it is well known that significant distortions may occur in the boundary areas of the filtered data because there is no sufficient data to be processed. This drawback largely affects the accuracy of topographic measurements and characterizations of precision freeform surfaces, such as freeform optics. To address this issue, a Gaussian process machine learning-based method is presented for extrapolation of the measured surface to an extended measurement area with high accuracy prior to filtering the surface. With the extrapolated data, the edge distortion can be effectively reduced. The effectiveness of this method was evaluated using both simulated and experimental data. Successful implementation of the proposed method provides a promising solution for numerous applications involving filtering processes.

Keywords: freeform surface; surface filtering; edge distortion; extrapolation; precision metrology; machine learning; ultra-precision machining

1. Introduction

Spatial filtering is a widely used technique in surface metrology to extract useful information from two-dimensional (2D) profile and three-dimensional (3D) areal topography measurement data, commonly achieved by separating components with different spatial frequencies (e.g. noise, roughness, waviness, form) [1] [2]. The low-pass filter is the most commonly used filter and is typically used to remove measurement noise and to smoothen surfaces [3]; a high-pass filter is more commonly used in image processing for edge detection [4]; narrow band-pass filtering can be implemented to extract signals with a specific wavelength [5]. Spatial filtering of surfaces was first implemented in hardware such as resistor-capacitor (RC) filter [6] and two-resistor-capacitor (2RC) filter [7] to remove measurement noise. Nowadays, filtering techniques are mostly implemented using software algorithms due to their flexibility and superior performance [7]. Many different means of filtering techniques have been developed such as 2RC filter [7], Gaussian filter [8], B-spline filter [9], morphological filter [10], wavelets filter [11] and Gaussian regression filter [12]. Many of these filtering techniques have been included in international standards such as ISO 25178 [13], ASME B46.1 [14] and ISO 16610 [15, 16].

The ordinary Gaussian filter is the most commonly used filter due to its simplicity [17]. However, the transmission characteristic of the ordinary Gaussian filter limits its performance in the presence of outliers [18]. Moreover, the filtering result can be significantly distorted near the boundary [19], which is known as the edge effect. Many advanced filtering techniques such as robust Gaussian filtering [16, 20], Gaussian regression filtering [12, 21] and the combined robust Gaussian regression filtering [22, 23] have been developed to address the edge effect. Most of these filtering techniques are implemented using convolution approaches with different filtering operators such as Gaussian and B-spline functions. As the convolution process requires data

outside the boundary, which has to be extrapolated, there is an inherent edge effect [24] due to extrapolation error. Solutions have been proposed to address this issue. Janecki [25] proposed a method of recursive Gaussian filters by selecting appropriate initial values for the filter difference equations. Moreover, Gaussian regression filter [21] and robust Gaussian filter [26] were developed using a modified spatially varying Gaussian weighting function, i.e. applying narrower weighting functions near the boundary.

However, the modified Gaussian weighting function is computationally intensive and computational speed is very low. Researchers have since developed accelerated algorithms to improve the computational speed utilising Graphics Processing Units (GPU) [27, 28], however implementation of these algorithm requires knowledge in GPU-optimised programming and costly GPU. An alternative solution to manage the edge effect is to remove from the filtered surface area near the boundary known to be distorted. However, this method discards part of the data, the amount of which can be significant when the cut-off length (i.e. the window size of the convolution) is large. This can be undesirable when the entire measured surface is needed for evaluation, or when there is insufficient area left to be statistically meaningful.

In order to avoid loss of measured data while managing the edge effect, researchers have attempted to extrapolate the surface so that the fast ordinary Gaussian filtering can be applied. For example, Dai and Yang [29] proposed a Fourier transform-based method for extrapolation of the fringe pattern for interferogram analysis. However, the method relied on a strong periodical pattern, which is available in fringe images but can hardly be applied in freeform surfaces. Lundström and Unsbo [30] proposed a B-spline-based method for unwrapping Hartmann-Shack images used in the measurement of wavefront aberrations. A B-spline function was fitted using a least-squares estimate and then the function was extrapolated to find expected spot patterns for unconnected

lenslets. Foracchia et al. [31] proposed a parametric model-based extrapolation method for the detection of optic discs in retinal images even when the target is outside of the images. These are essentially model-based methods which are well established where the images have a known mathematic model. Cigizoglu [32] proposed an artificial neural network (ANN)-based method to estimate, forecast and extrapolate river flows. Compared with conventional models, the method could provide a better fit to the data. However, the estimation error was still relatively large at 20%. Janecki [33] proposed an extrapolation method to reduce the edge effect in the profile filtering. The original profile was extrapolated at both ends using appropriate polynomial functions. Most existing methods are focused on specific tasks and there is relatively little research for precision surface measurement, especially for freeform surface measurement with an unknown mathematic model. Moreover, the sub-micron level accuracy requirement for precision surfaces is difficult to achieve. Recently, Gaussian process has gained research interest [34, 35] for data modelling regarding the measurement of precision freeform surfaces with high accuracy. The measurement process contains measurement noise which is governed by Gaussian distribution and hence the measurement process is essentially a Gaussian process. With its powerful prediction function, the Gaussian process data modelling method can not only interpolate unsampled data within the measured area but can also extrapolate data outside the measured area with high accuracy. With accurate extrapolation, the enlarged surface can be filtered with the ordinary Gaussian filter, followed by removal of the extrapolated area from the filtered surface. As a result, the fast ordinary Gaussian filtering can be performed on the measured surface with minimal edge effect, and without discarding valuable measurement data.

This paper proposes a Gaussian process machine learning-based surface extrapolation and filtering (GPEF) method to address the edge effect issue. The method is designed to accurately

extrapolate the measured surface topography outside its boundary using the Gaussian process machine learning method. The effectiveness of the proposed method is demonstrated by simulations using three different freeform surfaces and verified with measurement experiments. Results show the GPEF method is able to manage the edge effect better than zero-order Gaussian regression filtering and robust Gaussian filtering techniques included in international standards. Successful implementation of this method not only contributes to the measurement of precision surfaces, but also provides a new filtering approach in other research areas such as signal processing and image processing.

2. Problem statement - edge effect in Gaussian filtering of a 2D profile and a 3D surface

Filtering of a profile or a surface can be determined as a convolution operation of the profile or surface with a specific shape of the convolution window [19]. For a Gaussian type filter, the window shape is a Gaussian function. Gaussian filtering for a 2D profile is the convolution calculation of the profile using a 1D Gaussian function:

$$S(x) = \frac{1}{\alpha\lambda_c}\exp\left[-\pi\left(\frac{x}{\alpha\lambda_c}\right)^2\right], \tag{1}$$

where $\alpha = \sqrt{\frac{\ln 2}{\pi}}$, x is the location of the centre of the weighting function and λ_c is the cut-off length.

Gaussian filtering for a 3D surface is the convolution of the surface using a weighting of a 2D Gaussian function which is the product of two 1D Gaussian functions:

$$S(x,y)=\frac{1}{\alpha^2\lambda_{xc}\lambda_{yc}}\exp\left[-\left[\pi\left(\frac{x}{\alpha\lambda_{xc}}\right)^2+\pi\left(\frac{y}{\alpha\lambda_{yc}}\right)^2\right]\right], \quad (2)$$

where $\alpha=\sqrt{\frac{\ln 2}{\pi}}$, and x and y are the location coordinates of the centre of the weighting function, respectively. λ_{xc} and λ_{yc} are the cut-off length in x and y directions, respectively.

The Gaussian filter for a 2D profile and for a 3D surface as a convolution process is illustrated as shown in Figure 1 and Figure 2, respectively. In the case of a 2D profile, the edge effect is caused by a lack of data at the beginning and end of the convolution process. In the case of a 3D surface, the edge effect occurs along the boundary area. For example, the design of the 2D profile as shown in Figure 1 is determined by $z=\cos(x), x\in[-\pi,\pi]$, to which a zero-mean Gaussian noise signal with a variance of 0.1 mm has been added. The cut-off length of the low-pass Gaussian filter is 0.8 mm, so the window size of the weighting function is 0.8×2 = 1.6 mm, as shown in Figure 1(b). Figure 1(b) shows that the filtered profile is distorted within the 0.8 mm band around the boundary. The deviation in the filtered profile from the design profile is shown in Figure 1(c), where the largest distortion in the boundary is approximately 0.5 mm, which is significant compared to the profile height (2 mm).

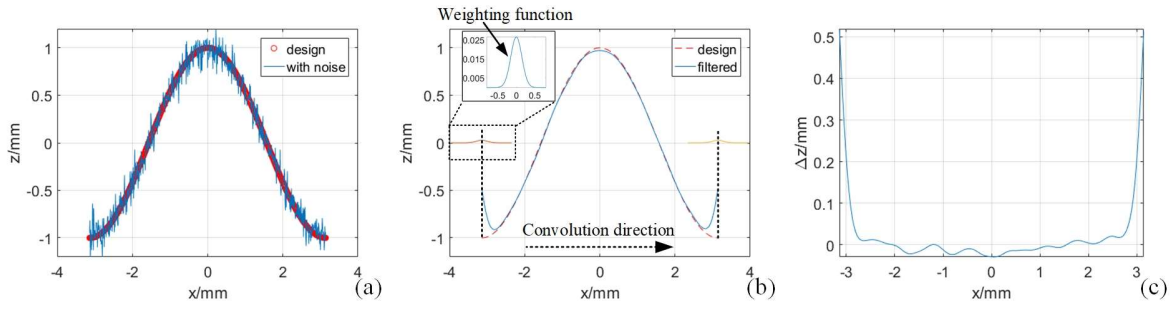

Figure 1 Convolution for a 2D profile, (a) design cosine profile with added noise, (b) filtered profile and (c) deviation in the filtered profile

Another example with the design of a 3D surface as shown in Figure 2 is determined by $z = \cos(x) + \cos(y)$, where $x, y \in [-\pi, \pi]$. The cut-off length of the Gaussian function in the x and y directions are both 0.8 mm, which means the window size is (1.6×1.6) mm. The filtered surface is shown in Figure 2(b), indicating distortion in the 0.8 mm-wide band around the edges of the surface. The deviation in the filtered surface from the design surface is shown in Figure 2(c) where the largest distortion in the boundary is approximately 1.5 mm. It is large compared to the surface height, i.e. 1.5 mm to 3 mm.

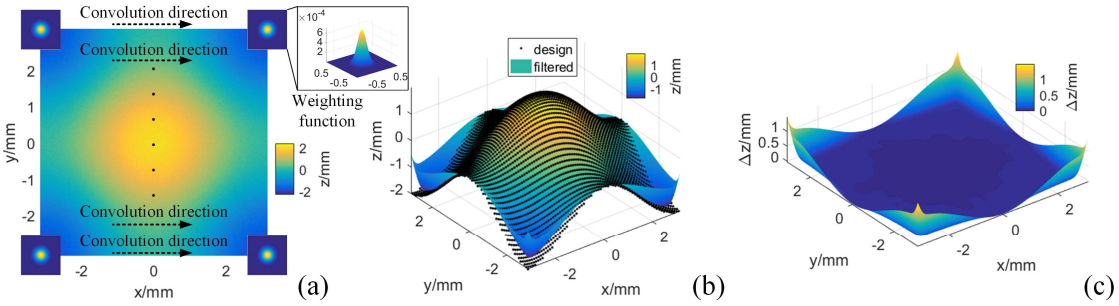

Figure 2 Convolution for a 3D surface, (a) convolution with a Gaussian weighting function over the surface, (b) filtered surface, and (c) deviation in the filtered surface

The Gaussian regression filter, which is designed to improve the edge effect is discussed in this paper and it is determined by minimizing the objective function [12]:

$$E(k) = \sum_{p=1}^{n} (z(p) - w(k))^2 S(k, p) \Delta x \tag{3}$$

where z is the profile data, w is the mean line data, n is the number of points, k is the index for the location of the centre of the weighting function, p is the index of points, Δx is the spacing, and S is determined by:

$$S(k, p) = \frac{1}{\lambda_c \sqrt{\ln 2}} \exp\left(-\frac{\pi^2}{\ln 2} \frac{\left((k - p)\Delta x\right)^2}{\lambda_c^{\,2}} \right) \tag{4}$$

where λ_c is the cut-off length.

The zero-order Gaussian regression filter for a 2D profile has the following weighting function:

$$S_{MOD}(k,p)=\frac{S(k,p)}{\sum_{p=1}^{n}S(k,p)}, \tag{5}$$

where $S(k,p)$ is given by:

$$S(k,p)=\frac{1}{\alpha\lambda_c}\exp\left[-\pi\left(\frac{p-k}{\alpha\lambda_c}\right)^2\right] \tag{6}$$

where $\alpha=\sqrt{\frac{\ln 2}{\pi}}$.

The zero-order Gaussian regression filter for a 3D surface has the following weighting function:

$$S_{MOD}(kx,px,ky,py)=\frac{S(kx,px,ky,py)}{\sum_{ly=1}^{ny}\sum_{lx=1}^{nx}S(kx,px,ky,py)}, \tag{7}$$

where *kx* is the index for the location of the centre of the weighting function in the *x* direction, *px* is the index of points in the *x* direction, *ky* is the index for the location of the centre of the weighting function in the *y* direction, *py* is the index of points in the *y* direction, *lx* is the index of points in the *x* direction and *ly* is the index of points in the *y* direction, *nx* is the number of points in the *x* direction and *ny* is the number of points in the *y* direction, and $S(kx,px,ky,py)$ is given by:

$$S(kx,px,ky,py)=\frac{1}{\alpha^2\lambda_{xc}\lambda_{yc}}\exp\left[-\left[\pi\left(\frac{px-kx}{\alpha\lambda_{xc}}\right)^2+\pi\left(\frac{py-ky}{\alpha\lambda_{yc}}\right)^2\right]\right] \tag{8}$$

where $\alpha=\sqrt{\frac{\ln 2}{\pi}}$, λ_{xc} and λ_{yc} are the cut-off lengths in the *x* and *y* directions, respectively.

It should be noted that the computational complexity of the zero-order Gaussian regression filter is much higher than that of the ordinary Gaussian filter, and the second-order Gaussian regression filter is even more complex [7]. The computational complexity is due to the weighting function for each point in the boundary area having different weighting values while for the inner data they are the same. Robust Gaussian filter included in ISO 16610 [16] also has the computational complexity problem. The proposed method aims to retain the fast computation speed of the ordinary Gaussian filter and eliminate the edge effect by enlarging the surface with accurately extrapolated data.

3. Gaussian process machine learning-based extrapolation and filtering (GPEF) method

The schematic of the proposed Gaussian-process machine learning-based extrapolation and filtering (GPEF) method is shown in Figure 3. The original surface is extrapolated using the Gaussian-process machine learning (GPML) method [36]. With the GPML method, a surface model is first trained using the original surface and then used to extrapolate outside the surface boundary. With the enlarged surface, the problem of having insufficient data in the boundary area is solved. Next, the enlarged surface is filtered with the fast ordinary Gaussian filtering algorithm.

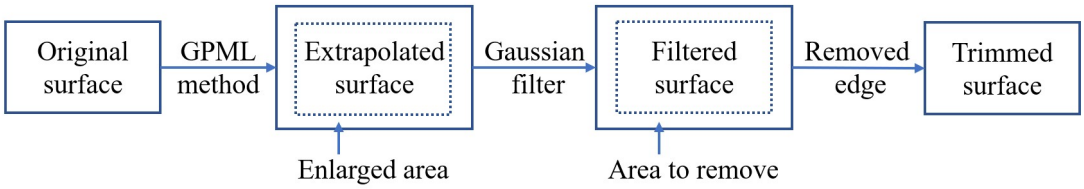

Figure 3 Schematic of the proposed GPEF method

At this point, the filtered surface has the same size as the enlarged surface and the edge effect occurs only within the extrapolated area. In the last step, the filtered surface is trimmed to the same size as the original surface as the final result. With the edge area removed, the influence of the edge effect is eliminated.

Gaussian process machine learning-based surface extrapolation

The measurement process contains noise governed by Gaussian distribution and the original measured surface can be determined by [35, 36],

$$z = f(x) + \varepsilon, \tag{9}$$

where z is the measurement result, $f(x)$ is the true value, x is the measured location, and ε is the measurement noise, which can be determined by:

$$\varepsilon \sim N(0, \sigma^2), \tag{10}$$

where the noise is Gaussian distributed with zero mean and σ^2 variance.

Since the true value of the surface $f(x)$ is unknown and the Gaussian process machine learning method is used for modelling of the measurement data:

$$f(x) \sim GP(m(x), k(x, x')), \tag{11}$$

where $m(x)$ and $k(x, x')$ are the mean function at location x, and covariance function at x and x', respectively, they can be determined by:

$$m(x) = E[f(x)], \tag{12}$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))], \tag{13}$$

Prediction of new data f_* at new location X_* can be determined by:

$$\begin{bmatrix} z \\ \mathrm{f}_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right), \tag{14}$$

where I is the identity matrix, X is the matrix of the measured locations, X_* is the matrix of the predictive locations and σ^2 is the noise variance.

The predictive equation for Gaussian process regression is:

$$\mathrm{f}_* \mid X, z, X_* \sim N(\bar{\mathrm{f}}_*, \mathrm{cov}(\mathrm{f}_*)) \tag{15}$$

where

$$\bar{\mathrm{f}} = E[\mathrm{f}_* \mid X, z, X_*] = K(X_*, X)[K(X, X) + \sigma^2 I]^{-1} z \tag{16}$$

$$\mathrm{cov}(\mathrm{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma^2 I]^{-1} K(X, X_*) \tag{17}$$

In this study, Gaussian process modelling was implemented by using the GPML toolbox [37]. It should be noted that the new location for prediction can be both inside or outside the area of the original dataset. When the new location is outside of the original surface, additional data are added to enlarge the surface. The optimal extrapolation size is critical and is determined by the cut-off length used in the Gaussian filter. Since the affected length of the distortion of the Gaussian filter implemented in the next step equals the length of the cut-off length, and that the extrapolated data will be trimmed after filtering, the enlarged size of the surface is a cut-off length outward from the edges.

After the original surface is enlarged using the GPML method, the enlarged surface is then filtered with the ordinary Gaussian filter. At this point, the edge effect is only observed in the enlarged area and the affected area is then trimmed and a final result with the same size as the original surface is obtained.

4. Experimental verification and discussion

To evaluate the effectiveness of the proposed GPEF method, a series of experiments using simulated and experimental data were conducted and the results and discussions are given in this section. In order to demonstrate the advantage of the proposed method, the results are also compared with those produced by a zero-order Gaussian regression filter and a robust Gaussian filter. All cut-off lengths used in the experiments were chosen according to the international standards of ISO 16610.

4.1. Simulation experiments

1) Sinusoidal surface

A sinusoidal surface was simulated as shown in Figure 4 and it is determined by:

$$z = 0.1\left[\sin(\frac{2\pi x}{5}) + \cos(\frac{2\pi y}{5})\right] + N_m \tag{18}$$

where $x, y \in [-10,10]$ mm, sampling space is 0.2 mm. N_m is the normally distributed measurement noise with zero mean and 0.1 mm standard deviation. Figure 4(a) shows the underlying surface and Figure 4(b) shows the surface with measurement noise. The period of the sinusoidal pattern is 5 mm.

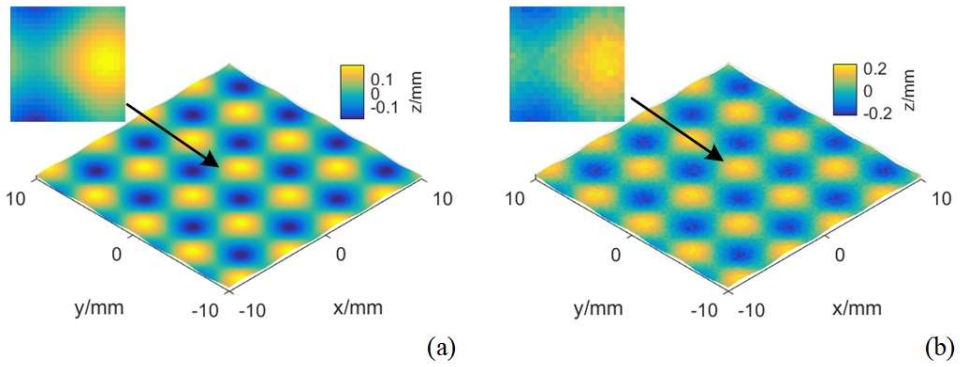

Figure 4 Simulated sinusoidal surface, (a) design surface, and (b) design surface with noise

The result of filtering with the zero-order Gaussian regression filter (ZOGF) with the cut-off length of 0.8 mm is shown in Figure 5. The noisy surface is smoothed and the deviation from the underlying surface is shown in Figure 5(b), which exhibits an obvious edge effect: the deviation in the edge area is larger than the inner area. The root mean square (RMS) of the deviation for the

entire surface is 4.3 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 4.1 μm and 5.1 μm, respectively.

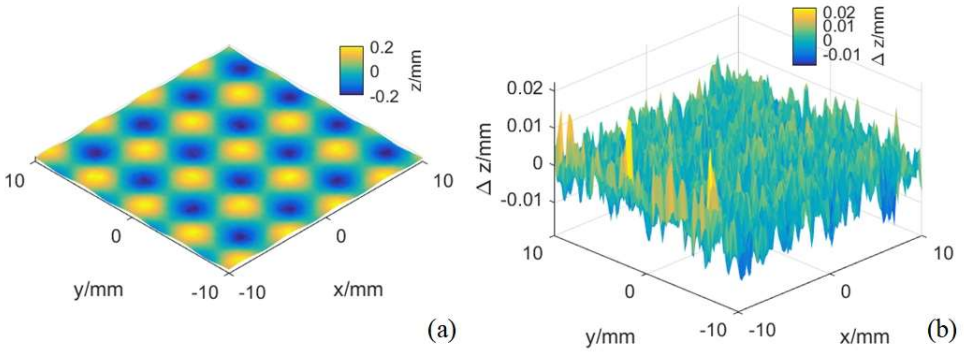

Figure 5 Result of ZOGF (a) filtered surface, and (b) deviation from the design surface

The result of filtering with the robust Gaussian filter (RGF) with the same cut-off length of 0.8 mm, is shown in Figure 6. The RMS of the deviation for the entire surface is 6.9 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 5.6 μm and 16.5 μm, respectively. The result shows that the edge effect is still large.

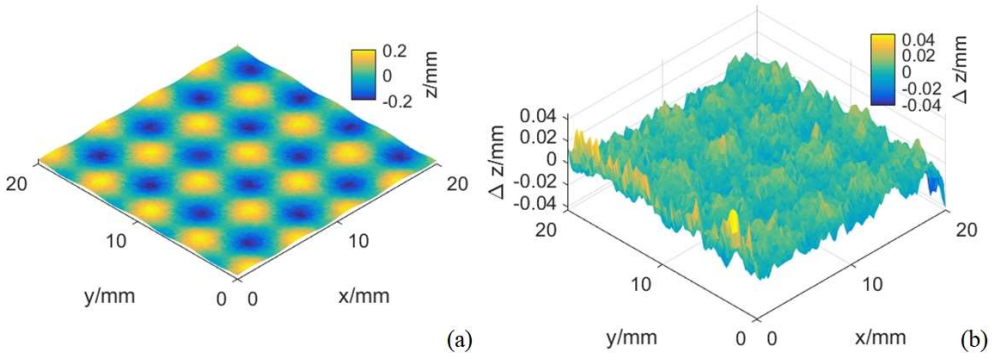

Figure 6 Result of RGF (a) filtered surface, and (b) deviation from the design surface

With the proposed GPEF method, the original surface is first extrapolated and the result is shown in Figure 7(a). The result shows that the pattern of the original surface is successfully learned and the extrapolated surface stitches to the original surface seamlessly. The extrapolated

surface is also compared with the underlying surface (extended definition of x, y in Eq. (18)) and the deviation is shown in Figure 7(b). The result shows that the deviation in the extrapolated area is relatively small compared to the amplitude of the design surface and therefore does not significantly influence the convolution result.

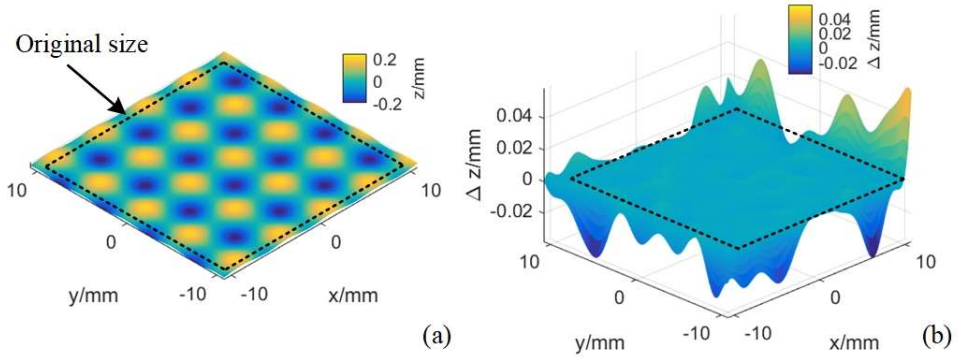

Figure 7 Result of Gaussian process extrapolation (a) extrapolated surface, and (b) deviation from the reference surface

The ordinary Gaussian filter was applied to the enlarged surface and the extrapolated area was subsequently removed, resulting in a filtered surface with the same size as the original surface, as shown in Figure 8(a). The deviation to the underlying surface is shown in Figure 8(b) and the pattern is evenly distributed with a small amplitude, which is significantly lower than those produced by ZOGF and RGF. The RMS of the deviation in the entire surface is 4.2 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 4.1 μm and 4.5 μm, respectively. The results are summarised in Table 1. The improvement in terms of

RMS of deviation on the whole surface of the proposed GPEF method compared to the ZOGF and RGF methods are 3.13% and 39.82%, respectively.

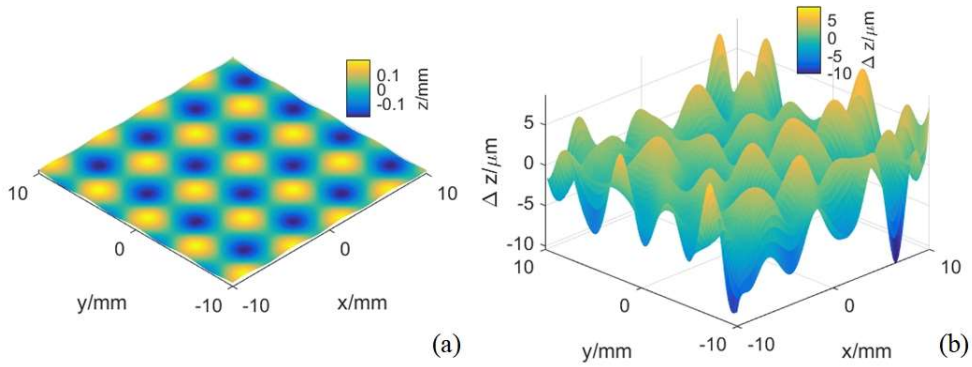

Figure 8 Result of GPEF method (a) final result, and (b) deviation from the reference surface

Table 1 Comparison of deviations in the filtered surfaces using different methods

	ZOGF	**RGF**	**GPEF**
RMS in the interior	4.1	5.6	4.1
RMS in the edge band	5.1	16.5	4.5
RMS on the whole	4.3	6.9	4.2

(unit: μm)

2) F-theta lens surface

To demonstrate the effectiveness of the proposed method with freeform surface without periodic patterns, an f-theta lens surface was simulated for the experiment as shown in Figure 9. Figure 9(a) shows the design surface and Figure 9(b) shows the target surface with added noise and it is determined by:

$$z = ax^2 + bx^4 + cy^2 + N_m \tag{19}$$

where $x \in [-40, 40]$ mm, $y \in [-15, 15]$ mm, $a = -1/250$, $b = 1/92000$ and $c = -1/25$. The sampling space is 0.2 mm. N_m is the measurement noise with zero mean and 0.1 mm standard deviation.

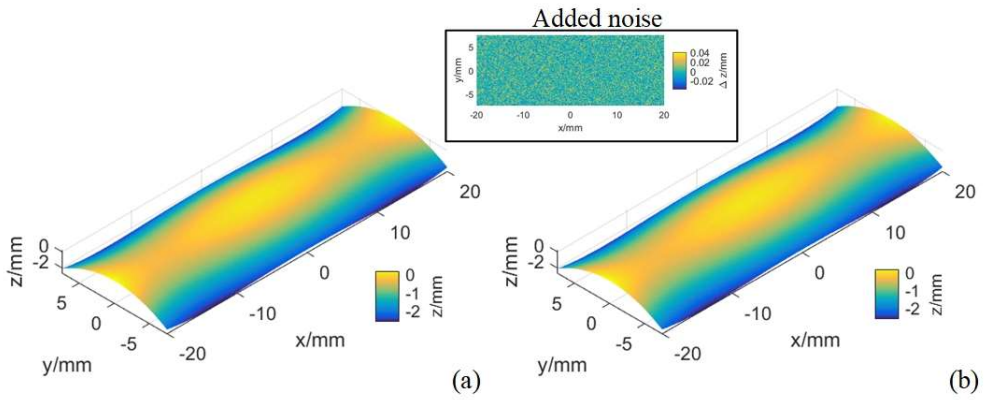

Figure 9 Simulation with f-theta surface (a) design surface, and (b) design surface with noise

The surface was filtered with the ZOGF with cut-off length of 0.8 mm and the result is shown in Figure 10(a). The result was compared to the design surface and the deviation is shown in Figure 10(b). The deviation is significantly large near the edges. The RMS of the deviation for the entire surface is 20.4 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 18.8 μm and 26.7 μm, respectively.

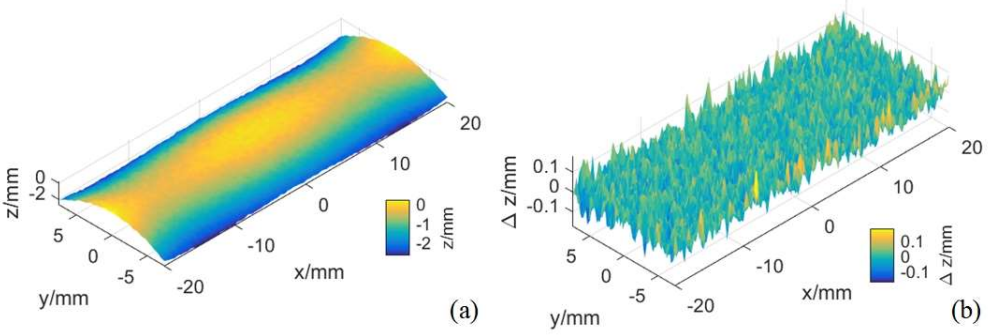

Figure 10 Result of ZOGF (a) filtered surface, and (b) deviation from the design surface

The surface was also filtered with the RGF and the result is shown in Figure 11(a). The RMS of the deviation for the entire surface is 40.6 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 40.1 μm and 41.2 μm, respectively. The result shows that the edge effect is also obvious.

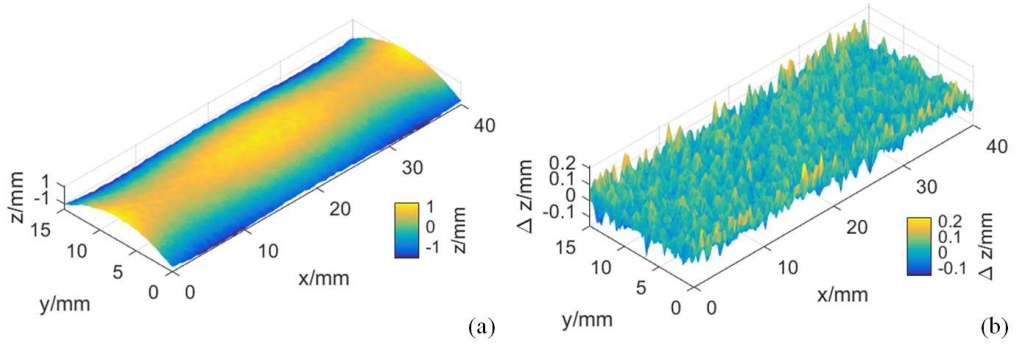

Figure 11 Result of RGF (a) filtered surface, and (b) deviation from the design surface

With the proposed GPEF method, the surface was firstly extrapolated and the result is shown in Figure 12(a) while the deviation to the design surface (with extended definition of *x, y* in Eq. (19)) is shown in Figure 12(b). The result shows that although the edge area of the extrapolated surface has a relatively large deviation from the design surface, it is continuously extended from the original surface with micrometre accuracy.

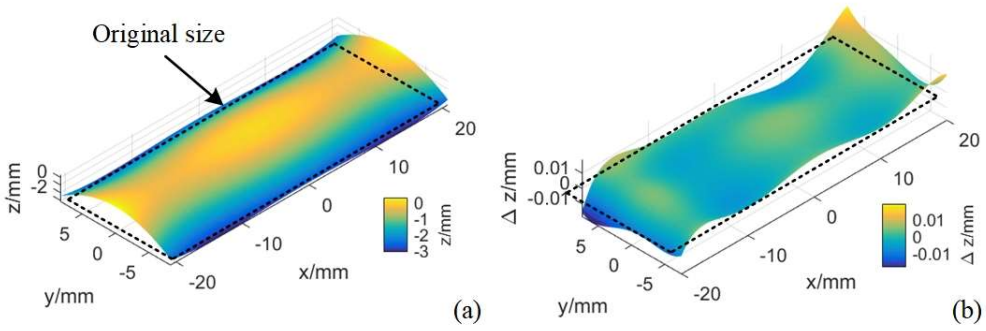

Figure 12 Result of Gaussian process extrapolation (a) extrapolated surface, and (b) deviation from the reference surface

After the surface was extrapolated, it was filtered with the ordinary Gaussian filter and the affected edge was removed and finally the result obtained is shown in Figure 13(a). The deviation of the result to the design surface is shown in Figure 13(b). The result shows that the deviation is smaller compared to both ZOGF and RGF. The RMS of the deviation of the entire surface is 18.9 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 18.7 μm and 21.6 μm, respectively. The results are also summarised in Table 2. The improvement in terms of RMS of deviation on the whole surface of the proposed GPEF method compared to the ZOGF and RGF methods are 8.02% and 53.84%, respectively.

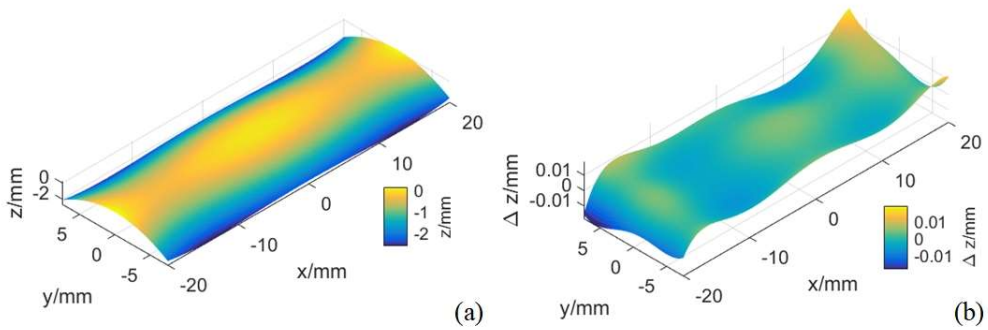

Figure 13 Result of GPEF method (a) filtered result, and (b) deviation from the reference surface

Table 2 Comparison of deviations from filtered results to underlying surface for different methods

	ZOGF	**RGF**	**GPEF**
RMS in the interior	18.8	40.1	18.7
RMS in the edge band	26.7	41.2	21.6
RMS on the whole	20.4	40.6	18.9

(unit: μm)

4.2. Actual measurement experiment

The proposed method was applied to the experimental data to verify the effectiveness of the proposed GPEF method. A sinusoidal surface with only one cycle was used in this experiment and it can be determined by Eq. (20). A workpiece was computer numerical control (CNC) milled and the workpiece was measured with a Werth VideoCheck UA multi-sensor coordinate-measuring machine (CMM) using a touch trigger probe - Renishaw TP200; the probing error is ±0.65 μm while the Maximum Permissible Measuring Error (MPE) of the CMM is $(0.75+L/300)$ μm. The sampling space was 1 mm for the experiment.

$$z=\sin(\frac{2\pi x}{60})+\cos(\frac{2\pi y}{60}) \tag{20}$$

where $x, y \in [-30, 30]$ mm. The design of the workpiece and the measurement process is shown in Figure 14.

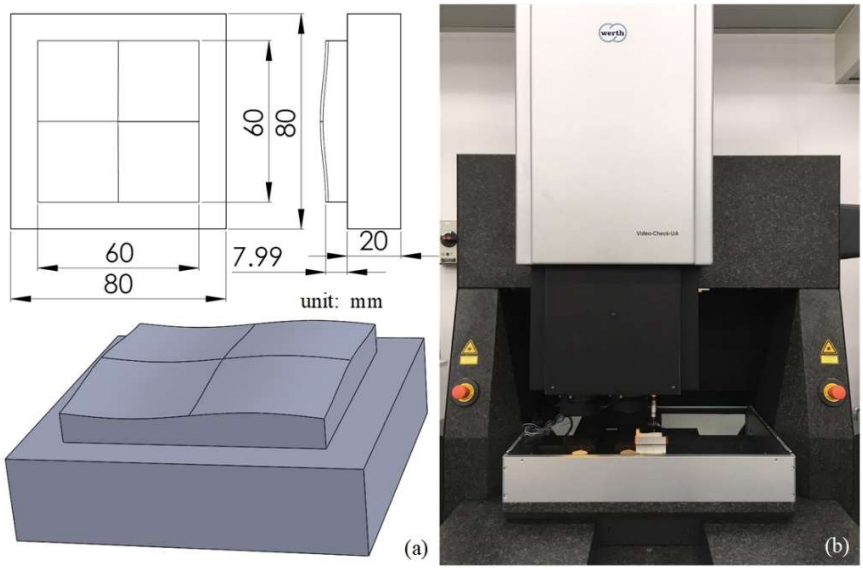

Figure 14 Actual measurement experiment, (a) design of the workpiece, and (b) measurement process

In the actual measurement experiment, the reference surface was chosen to be the design surface, where $x, y \in [-30, 30]$ mm. In order to easily align the data to the design surface, the

processed area was chosen to be slightly smaller than the design surface, where $x, y \in [-27, 27]$ mm, as shown in Figure 15. All the alignment results in this experiment were processed by using the Iterative Closest Point (ICP) algorithm [38].

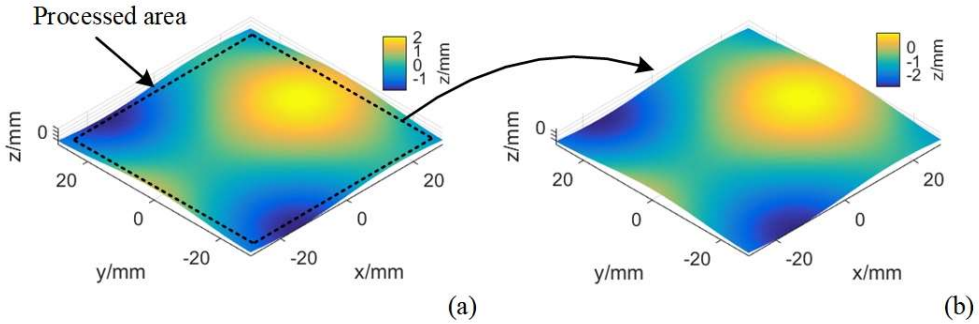

Figure 15 Data for the measurement (a) whole measurement data, and (b) trimmed data for processing

The filtered surface using ZOGF is shown in Figure 16(a) and the deviation from the design surface is shown in Figure 16(b). The cut-off length of the filter is 2.5 mm. The edge distortion is large as compared with the inner surface. The RMS of the deviation in the entire surface is 11.2 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 6.3 μm and 29.1 μm, respectively.

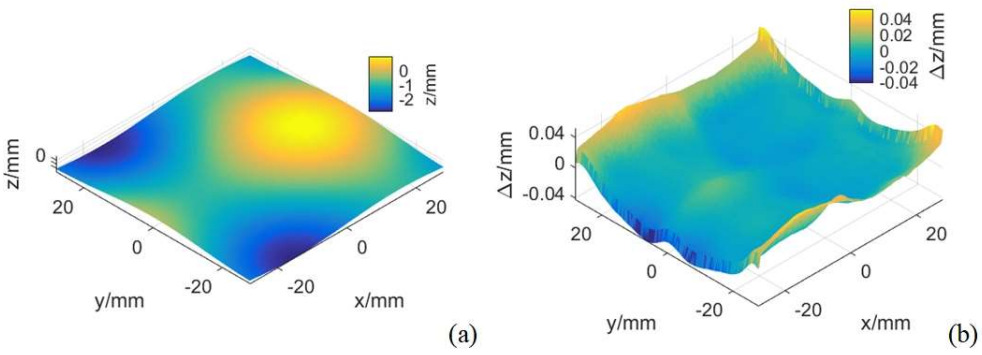

Figure 16 Result of ZOGF (a) filtered surface, and (b) deviation from the design surface

The filtered surface using RGF is shown in Figure 17(a) and the deviation from the design surface is shown in Figure 17(b). The RMS of the deviation in the entire surface is 31.2 μm. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 25.5 μm and 36.8 μm, respectively. The result shows that the edge effect is also obvious.

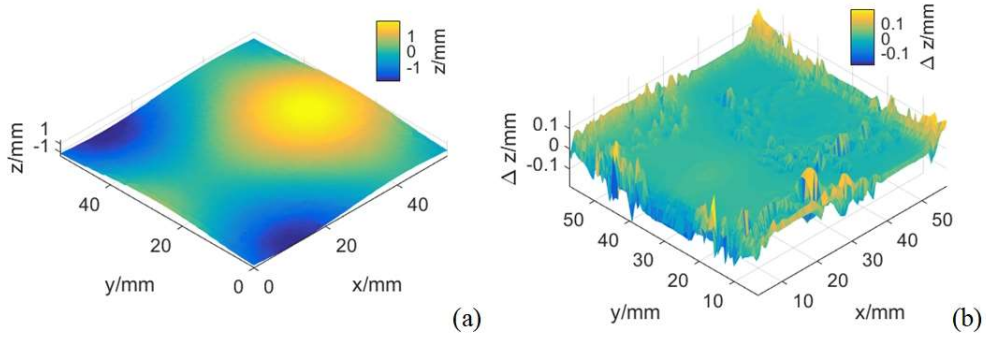

Figure 17 Result of RGF (a) filtered surface, and (b) deviation from the design surface

With the proposed GPEF method, the extrapolated surface is shown in Figure 18(a) and the deviation from the design surface is shown in Figure 18(b). The extrapolated surface is well connected with the original surface and it shows the effectiveness of the surface extrapolation method. The extrapolated surface was then filtered with the ordinary Gaussian filter and trimmed back to the original size, as shown in Figure 19(a). The result was compared with the design surface and the deviation is shown in Figure 19(b). The RMS of the deviation of the entire surface is 10.0 μm, which has improvement over ZOGF and RGF. The RMS values of the deviations in the interior excluding the edge band and in the edge band are 6.2 μm and 25.3 μm, respectively.

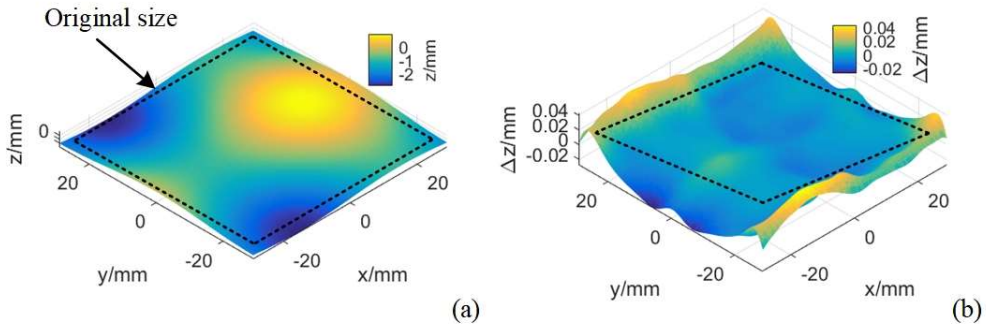

Figure 18 Result of Gaussian process extrapolation (a) extrapolated surface, and (b) deviation from the reference surface

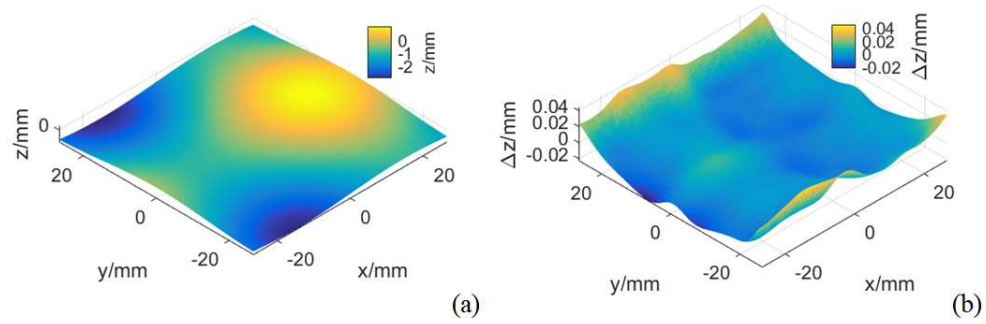

Figure 19 Result of GPEF (a) filtered result, and (b) deviation from the reference surface

The performance of the proposed GPEF method is summarised in Table 3. The improvement in terms of RMS of deviation on the whole surface of the proposed GPEF method compared to the ZOGF and RGF methods are 10.92% and 68.07%, respectively. All the results show that the performance of the proposed GPEF underwent improvement compared to the two other methods. It is interesting to note that the performance of the filters regarding the edge effect differs for different surfaces and this is due to the fact that different surfaces have different deviations from the zero plane which influences the filtering result.

Table 3 Comparison of deviations from filtered results to underlying surface for different methods

	ZOGF	RGF	GPEF
RMS in the interior	6.3	25.5	6.2
RMS in the edge band	29.1	36.8	25.3
RMS on the whole	11.2	31.2	10.0

(unit: μm)

The surfaces used in the simulation experiments are sinusoidal surface and f-theta surface, where the sinusoidal surface has strong periodical pattern while the f-theta has not. The surface used in the measurement experiment is a sinusoidal surface with only one cycle, so it has not periodical pattern either. All these freeform surfaces are common in precision engineering. The Gaussian process machine learning method can successfully predict the surface data at unsampled locations with high accuracy. The accurately extrapolated data in the non-periodical surfaces demonstrated that GPML method can accurately predict the extra data outside the surface and is not limited to periodical surfaces. Hence, this approach is expected to be applicable to other types of surfaces with high confidence.

4.3. Influence of the cut-off length

The cut-off length is an important parameter of the filtering process to filter out unwanted signal and keep the desired wavelength of the signal. It is interesting to note that the cut-off length affects the area of edge distortion as well, with a large cut-off length, and the affected zone is large, and vice versa. The influence of cut-off length for different filtering techniques was studied, i.e. for ZOGF, RGF and the proposed GPEF method with cut-off lengths from 2.5 mm, 8.0 mm to 25 mm (taken from international standard of ISO 16610), using the same experiment data in the actual measurement experiment. The results of RMS deviations of the entire surface are shown in Figure 20. It can be seen that the GPEF method has the best performance since it has the optimal solution

to address the lack of data in the edge area. The larger the cut-off length, the greater advantage of GPEF is found.

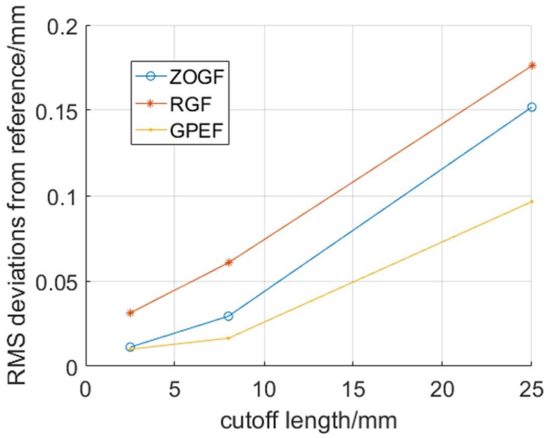

Figure 20 Influence of cut-off length of filters

4.4. Discussions

The proposed GPEF method demonstrates a significant improvement to the edge effect in surface filtering. In terms of RMS of deviation on the whole surface, the proposed GPEF method has improvement of approximately from 3-10% and 40-70% compared to the ZOGF and RGF methods, respectively, according to surfaces with different peak-to-valley values. The method utilizes the surface extrapolation method to assist the filtering process. The method can also be potentially applied in other research areas in precision engineering such as toolpath generation [39] for diamond-turned surface, where sometimes the design surface is not a round one and the diamond tool then needs to go outside the designed area and thus the data outside have to be calculated with an extrapolation method.

Currently, the calculation of Gaussian process machine learning requires long computation time especially when the dataset is large. Although the data size in the present study was limited to several hundreds × several hundreds and there was no need for special handling to reduce the data size, it would be difficult to deal with an extremely large dataset such as those in the high

dynamic range measurements [35, 40]. One solution to this is to first down-sample the original dataset to a reasonable dimension and then conduct the Gaussian process calculation for surface extrapolation. After conducting the extrapolation process, interpolation for the dataset is implemented to obtain a resolution as high as the original data. This can save significant computational time and yet have only a small influence on the result since the influence of the slight difference of the extrapolated data is negligible.

5. Conclusion

This paper presented a novel Gaussian process machine learning-based extrapolation and filtering (GPEF) method which attempts to improve the edge effect during surface filtering. The novelty of the GPEF method lies in the accurate extrapolation of the original measured surface before ordinary Gaussian filtering is applied. As a result, edge distortion is significantly reduced while the entire measured surface area is retained. Simulation and actual measurement experiments involving three different freeform surfaces have shown that the edge effect was improved by approximately 3-10% and 40-70%, respectively, over the commonly used zero-order Gaussian regression filter and robust Gaussian filter. Successful implementation of this method not only helps to improve the accuracy of surface characterization, but also provides a new method for other research fields dealing with signal processing. The limitation of the proposed method is that down sampling may be required for large datasets (e.g. more than one million points) in order to reduce computation time.

Acknowledgements

The work described in this paper was mainly supported by the National Natural Science Foundation of China (No. 51675456), China National Program. The authors sincerely thank the EMPIR organization. The EMPIR is jointly funded by the EMPIR participating countries within

EURAMET and the European Union (15SIB01: FreeFORM). The authors also thanks Dr Rong Su for valuable comments for the paper.

References

[1] B. Muralikrishnan, J. Raja, A Brief History of Filtering, Computational Surface and Roundness Metrology, (2009) 7-11.

[2] D.J. Whitehouse, Handbook of surface and nanometrology, CRC press, 2010.

[3] G. Taubin, T. Zhang, G. Golub, Optimal surface smoothing as filter design, in: B. Buxton, R. Cipolla (Eds.) Computer Vision — ECCV '96: 4th European Conference on Computer Vision Cambridge, UK, April 15–18, 1996 Proceedings, Volume I, Springer Berlin Heidelberg, Berlin, Heidelberg, 1996, pp. 283-292.

[4] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (1990) 629-639.

[5] R.B. Herrmann, Some aspects of band-pass filtering of surface waves, Bulletin of the Seismological Society of America, 63 (1973) 663-671.

[6] P. Horowitz, W. Hill, The art of electronics, Cambridge Univ. Press, 1989.

[7] B. Muralikrishnan, J. Raja, Computational surface and roundness metrology, Springer, 2009.

[8] I.T. Young, L.J. van Vliet, Recursive implementation of the Gaussian filter, Signal Processing, 44 (1995) 139-151.

[9] H. Zhang, Y. Yuan, W. Piao, A universal spline filter for surface metrology, Measurement, 43 (2010) 1575-1582.

[10] S. Lou, X. Jiang, P.J. Scott, Correlating motif analysis and morphological filters for surface texture analysis, Meas J Int Meas Confed, 46 (2013) 993-1001.

[11] K. Lingadurai, M.S. Shunmugam, Metrological characteristics of wavelet filter used for engineering surfaces, Measurement, 39 (2006) 575-584.

[12] S. Brinkmann, H. Bodschwinna, H.-W. Lemke, Accessing roughness in three-dimensions using Gaussian regression filtering, International Journal of Machine Tools and Manufacture, 41 (2001) 2153-2161.

[13] ISO 25178-2, Geometrical product specifications (GPS) -- Surface texture: Areal -- Part 2: Terms, definitions and surface texture parameters, 2012.

[14] ASME, B46.1-2009 Surface Texture (Surface Roughness, Waviness, and Lay), American Society of Mechanical Engineers, New York, NY, 2009.

[15] ISO 16610-31, Geometrical Product Specifications (GPS)–Filtration Part 31: Robust Profile Filters. Gaussian Regression Filters, International Standards Organization. British Standards Institute London, 2002.

[16] ISO 16610-71, Geometrical Product Specifications (GPS)–Filtration Part 71: Robust areal filters: Gaussian regression filters, International Standards Organization. British Standards Institute London, 2014.

[17] D.J. Whitehouse, Handbook of surface metrology, CRC Press, 1994.

[18] X.J. Jiang, D.J. Whitehouse, Technological shifts in surface metrology, CIRP Annals-Manufacturing Technology, 61 (2012) 815-836.

[19] D.J. Whitehouse, Surfaces and their Measurement, Elsevier, 2004.

[20] H. Li, C.F. Cheung, X.Q. Jiang, W.B. Lee, S. To, A novel robust Gaussian filtering method for the characterization of surface generation in ultra-precision machining, Precision Engineering, 30 (2006) 421-430.

[21] S. Brinkmann, Development of a robust Gaussian regression filter for three-dimentional surface analysis, 10. International Colloquim on Surfaces 31.1./1.2. 2000, Chemenitz, Germany, (2000).

[22] J. Seewig, Linear and robust Gaussian regression filters, J. Phys. Conf. Ser., IOP Publishing, 2005, pp. 254.

[23] W. Zeng, X. Jiang, P.J. Scott, Fast algorithm of the robust Gaussian regression filter for areal surface analysis, Measurement Science and Technology, 21 (2010).

[24] J. Raja, B. Muralikrishnan, S. Fu, Recent advances in separation of roughness, waviness and form, Precision Engineering, 26 (2002) 222-235.

[25] D. Janecki, Edge effect elimination in the recursive implementation of Gaussian filters, Precision Engineering, 36 (2012) 128-136.

[26] H. Li, X. Jiang, Z. Li, Robust estimation in Gaussian filtering for engineering surface characterization, Precision Engineering, 28 (2004) 186-193.

[27] J. Steinbring, U.D. Hanebeck, GPU-accelerated progressive Gaussian filtering with applications to extended object tracking, 2015 18th International Conference on Information Fusion (Fusion), 2015, pp. 1038-1045.

[28] Y. Su, Z. Xu, X. Jiang, GPGPU-based Gaussian filtering for surface metrological data processing, 12th International Conference Information Visualisation, IV08, London, 2008, pp. 94-99.

[29] M. Dai, Y. Wang, Fringe extrapolation technique based on Fourier transform for interferogram analysis, Optics Letters, 34 (2009) 956-958.

[30] L. Lundström, P. Unsbo, Unwrapping Hartmann-Shack Images from Highly Aberrated Eyes Using an Iterative B-spline Based Extrapolation Method, Optometry and Vision Science, 81 (2004) 383-388.

[31] M. Foracchia, E. Grisan, A. Ruggeri, Detection of optic disc in retinal images by means of a geometrical model of vessel structure, IEEE Transactions on Medical Imaging, 23 (2004) 1189-1195.

[32] H.K. Cigizoglu, Estimation, forecasting and extrapolation of river flows by artificial neural networks, Hydrological Sciences Journal, 48 (2003) 349-361.

[33] D. Janecki, Gaussian filters with profile extrapolation, Precision Engineering, 35 (2011) 602-606.

[34] M.Y. Liu, C.F. Cheung, C.H. Cheng, W.B. Lee, A Gaussian Process Data Modelling and Maximum Likelihood Data Fusion Method for Multi-Sensor CMM Measurement of Freeform Surfaces, Applied Sciences, 6 (2016).

[35] M.Y. Liu, C.F. Cheung, C.H. Cheng, R. Su, R.K. Leach, A Gaussian process and image registration based stitching method for high dynamic range measurement of precision surfaces, Precision Engineering 50 (2017) 99-106.

[36] C.K. Williams, C.E. Rasmussen, Gaussian processes for machine learning, the MIT Press, 2006.

[37] C.E. Rasmussen, H. Nickisch, Gaussian processes for machine learning (GPML) toolbox, The Journal of Machine Learning Research, 11 (2010) 3011-3015.

[38] P.J. Besl, N.D. McKay, A method for registration of 3-D shapes, IEEE Transactions on pattern analysis and machine intelligence, 14 (1992) 239-256.

[39] AMETEK Precitech Inc., DIFFSYS, 2016.

[40] M.Y. Liu, C.F. Cheung, D.J. Whitehouse, C.H. Cheng, An autonomous multisensor in situ metrology system for enabling high dynamic range measurement of 3D surfaces on precision machine tools, Measurement Science and Technology, 27 (2016) 115015.