# Application of Clustering Filter for Noise and Outlier Suppression in Optical Measurement of Structured Surfaces

Shan Lou, Dawei Tang, Wenhan Zeng, Tao Zhang, Feng Gao, Hussam Muhamedsalih, Xiangqian Jiang, and Paul J. Scott

***Abstract*— In comparison to tactile sensors, optical techniques can provide a fast, nondestructive profile/areal surface measurement solution. Nonetheless, high measurement noise, unmeasured points, and outliers are often observed in optical measurement, particularly for structured surfaces. To alleviate their detrimental impacts on the characterization of surface topography as well as the examination of micro/nanoscale geometries, a post processing filtering technique, i.e., the clustering filter, which is essentially an iterative process to find the aggregation center of a cluster of points, is implemented. The clustering filter is particularly useful for noises and outlier suppression for optical measurement of structured surfaces due to its edge-preserving capability. Five surface samples with structured features are measured by an in-house developed dispersive interferometer and a commercial white light interferometer, thereafter the measured surface data are filtered by the clustering filter. Both noise and outliers are suppressed, which not only facilitates the visualization and characterization of surface topography, but also enables the accurate evaluation of local functional geometries.**

***Index Terms*— Clustering filter, de-noising algorithm, optical metrology, structured surface, surface measurement.**

## I. Introduction

OPTICAL measurement techniques, e.g., optical interferometry, are nowadays popular in the field of surface texture measurement due to their fast measurement speed and high measurement resolution [1]–[3]. In comparison to the raster scanning style of tactile measurements, optical methods instantly acquiring raw data for producing profile/areal surfaces are more effective to determine the geometrical nature of surfaces without halting the manufacturing process. Another benefit of optical measurements is that they are noncontact, and thus have no risk of damaging the probes or the measurand surface as is the case for tactile measurements. Nonetheless, optical measurement techniques are not perfect. High measurement noise, unmeasured points, and artificial data, e.g., outliers, are usually noticed in optical measurement, especially when surface topography presents local geometries with a sharp gradient. For instance, in practical measurements of structured surfaces produced by laser texturing and additively manufactured surfaces, it is often noticed that large spikes and unmeasured points appear in the raw measurement data due to light diffraction at sharp edges.

Measurement values captured at high signal-to-noise ratio conditions are reported in optical measurement practice [4]. ISO/DIS 25178 Part 6 [5] defined measurement noise as noise added to the output signal occurring during the normal use of the instrument. Measurement noise can be caused by different sources [6]: instrument noises which are the internal noise of the instrument, e.g., electronic noise, optical noise; environmental noise generated by ground vibrations, air turbulence, and thermal issues [7]; and scanning noise generated by moving probe or samples during the measurement. Measurement noise usually features high spatial frequency and can be reduced by applying low-pass filtering techniques.

In addition to measurement noise, optical measurement techniques can often produce unmeasured or faulty points and outliers. Outliers are presented as sudden sharp jumps compared to the surrounding data and occur at certain localities on the surface [8]. In ISO 16610-1 [9], an outlier is defined as a local portion in a data set that is not representative, or not typical, for the partitioned integral feature. Outliers existing in surface measurement induce significant distortion to the assessment of surface texture and microgeometrical features on the surface, and thus should be removed or suppressed to obtain a reliable assessment result. They also affect the visualization of the surface in 3-D representation by flattening or contracting the scale in a false color display [10].

More specifically, the outliers commonly presented in optical surface measurement are spikes and batwings. Spikes are usually due to half-wavelengths and $2\pi$ errors caused by misclassification of fringe order and slope dependent dispersion [11]. Batwings are often observed at the locations close

Manuscript received June 24, 2019; revised November 12, 2019; accepted January 3, 2020. Date of publication February 26, 2020; date of current version August 11, 2020. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC), U.K., funding of the Future Advanced Metrology Hub under Grant EP/P006930/1, in part by the EPSRC under Grant EP/S000453/1, and in part by the Royal Academy of Engineering under Grant RCSRF1516/2/7. The Associate Editor coordinating the review process was Branislav Djokic. *(Corresponding author: Xiangqian Jiang.)*

Shan Lou, Dawei Tang, Wenhan Zeng, Feng Gao, Hussam Muhamedsalih, Xiangqian Jiang, and Paul J. Scott are with the Engineering and Physical Sciences Research Council (EPSRC) Future Metrology Hub, School of Computing and Engineering, University of Huddersfield, Huddersfield HD1 3DH, U.K. (e-mail: s.lou@hud.ac.uk; d.tang@hud.ac.uk; z.wenhan@hud.ac.uk; f.gao@hud.ac.uk; h.muhamedsalih@hud.ac.uk; x.jiang@hud.ac.uk; p.j.scott@hud.ac.uk).

Tao Zhang is with the Quality Engineering and Smart Technology (QUEST) Research Centre, Brunel University London, Uxbridge UB8 3PH, U.K. (e-mail: tao.zhang@brunel.ac.uk).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2020.2967571

to the slope discontinuity at the edges of microgeometrical features on the surface. For instance, in step height measurement, batwings are often generated at the sharp corner of the steps, where the step height is less than the coherence length of the source light. The batwing effect is explained as the interference between reflections of waves normally, incident on the top and bottom surfaces following diffraction from the edge [11], [12]. Other reasons that can lead to the generation of outliers include the presence of dust, marks or dirt on the measured sample surface [13].

The last two decades evidenced the technological shift that the surface texture of modern engineering surfaces is changing from a stochastic pattern to a structured pattern [14], [15]. The structured surfaces have dominant microfeatures on surface texture which are intentionally designed to meet specific functional requirements, such as optical properties, tribological properties, hydrodynamic properties, and paintability [16], [17]. Having recognized that the optical measurement of structured surfaces usually presents high measurement noise coupled with the large spikes and batwing outliers, it is proposed to use a numerical filtration technique, i.e., the clustering filter, to reduce the high-frequency measurement noise as well as suppress the outliers, while preserving the edges of microfeatures on surface.

The remainder of this article is structured in the following format: Section II provides a literature review of the techniques employed to reduce surface measurement anomalies. The basic principle of the clustering filter and its associated algorithm are detailed in Section III. In Section IV, the setups of two interferometry systems are illustrated and the two systems are used to measure structured surfaces; the clustering filter is then applied to de-noise the raw measured structured profiles/surfaces; Finally, Section V gives the conclusion and the future work.

## II. Literature Review

The aforementioned surface measurement anomalies are originated from specific optical measurement principles and their associated signal processing algorithms. This literature first summarizes the approaches aimed to minimize the production of these anomalies at the level of optical setup and signal processing. It is, however, noticed that even with optimized optics, measurement noises and outliers are almost unavoidable. Thus, the postprocessing approaches are adopted to further reduce their impacts. This is briefed in the second part of the literature review.

### A. Optical Approaches

Since the optical measurement techniques produce specific measurement anomalies, the techniques adopted for identifying and removing (or attenuating) the outliers are of necessity due to their influence on the following parameterization which generates the quantitative evaluation of surface texture. Many researchers have shown their particular interests in the outliers problem such as batwings in the last two decades. Generally, the strategies developed so far for batwing reduction, work either through combining two signal analysis techniques and comparing the two measurement results, or adding another wavelength to the optical system to assist the measurement. For instance, Harasaki *et al.* [12] combined phase shifting and coherence envelop peak sensing (least-squares fitting of the modulation contrast) techniques to permit high-resolution measurement as well as remove the batwing effect. Specifically, the two profiles obtained from the envelope and phase evaluations at the best focus position are compared at each point and then the batwing effect is removed in the phase unwrapping procedure. Vo *et al.* [18] proposed another algorithm which combines white light phase-shifting algorithm and fast Fourier Transform (FFT) coherence peak sensing techniques to effectively solve the problem of positioning error in the maximum modulation and remove the batwing effect at the step edges. Niehues *et al.* [19] proposed a dual-wavelength phase unwrapping technique to correct the batwings and ghost peaks by adding a second LED (with a different mean wavelength) in the vertical scanning interferometry and comparing the measurement results from two different illumination schemes. As such, the batwings at the sharp edges can effectively be removed. Ghim and Davies [20] applied the same concept to remove the ghost steps due to the inconsistencies between phase and coherence. Instead of adding another light source, their method was based on the Fourier analysis of the extracted signal at two different wavelengths and by using the difference information of the two height profiles to reveal and correct the height errors. However, the comparison techniques are not applicable to all optical measurement systems, measured data sets, and applications.

Additionally, enhancement of the optical system for minimization of spikes, generated by diffraction effects, has been investigated as well. Roy *et al.* [21] proposed a white light polarization interferometer and reduced the spikes using a phase shifter operating on the geometric (Pancharatnam) phase. However, this phase-shifting technique was time-variant and slowing the measurement further as three measurements were made of the irradiance at each scanning point. Niehues and Lehmann [22] proposed another physical way to reduce the diffraction effect during measurement by applying a confocal aperture in the illumination path of the interferometer as well as a signal processing strategy similar to that proposed by Harasaki *et al.* [12]. However, this technique suffered from longer measurement and computation time as well.

### B. Postprocess Approaches

The linear Gaussian filter is a commonly used technique to reduce high-frequency measurement noises. The measurement data are convolved by a Gaussian kernel function with a cutoff wavelength controlling the shape of the Gaussian function. The weights are determined by the Gaussian filter and, along with the neighborhood points, used to generate a weighted average. The Gaussian function has the bell shape, and thus the central kernel takes larger weights whereas two ends take less weight. When outliers are presented in measurement data, the linear Gaussian filter does not discriminate these outliers, and thus the filtration is distorted around the outlier areas due to the high impact of these outliers to the weighted averaging.

The robust Gaussian filter was then introduced to overcome the sensitivity of the linear Gaussian filter to outliers. Robust statistical estimates, e.g., the Turkey estimate, are incorporated into the filtering process, acting as a vertical weighted averaging as a supplementation to the lateral Gaussian weighted averaging [23]–[25].

The linear Gaussian filter works well for stochastic surfaces and its robust version can deal with outliers. However, its application is limited for structured surfaces, which usually contain functional geometrical features with steep edges. When the Gaussian kernel is acted on the edge portion of geometrical features, it tends to blur the sharp edges due to its averaging characteristic [26] and thus exerts a negative impact on the accurate evaluation of these features. A remedy solution can be adaptively changing the kernel function at the feature edges [27].

The spline filter is another commonly used surface filtration technique, which is specified by the filtration equation rather than the weighting function [28]. The spline filter originates from the form of a flexible natural cubic spline under the load of the measured profile [29]. The resulted profile is a spline which can be solved by a constrained optimization problem [30]. In comparison to the Gaussian filter, the spline filter could handle most of the surfaces in form measurement. To suppress outliers, the spline filter can integrate the robust estimates, the same as the robust Gaussian filter. However, in terms of structured surfaces, the spline filter is of limited use, because it is extremely difficult for the spline filter, by its nature, to approximate sharp geometries of structured features, and therefore, these geometries will be rounded by the filtration.

The median filter is a nonlinear filtering technique, which has a good performance in removing salt-and-pepper noises/speckle outliers from an image or a signal. However, it will fail when applied to data with clustered outliers since these outliers can result in biased medians. Ismail *et al.* [8] developed an optimized median filter. The outlier detection is based on the median relative height of the surrounding data within a defined detection window. The size of the detection window $n$ was initially determined based on the size of the largest outlier cluster $w$ observed on the measured data: $n > \sqrt{2}\ w$ and then iteratively incremented by 2 if needed. The outliers were identified by comparing the median of the data within the detection window to a threshold which is calculated from a cumulative probability curve of the medians of all data points. The median filter can reduce impulsive noise and preserve high contrast-edges, especially when applied to small speckle patterns [31]. However, it cannot achieve a good smoothing of nonimpulse noise and was reported not suitable for surfaces containing adjacent outlier clusters [32].

Anisotropic diffusion is an image processing technique used to reduce image noise without removing significant parts of the image content, e.g., the edges of an object [33]. It is a generalization of the linear diffusion process given by the equation $(\partial I/\partial t) - \mathrm{div}(c \cdot \nabla I) = 0$, where $I$ is the image data to be diffused, $\nabla$ denotes the gradient, $\mathrm{div}(\cdots)$ is the divergence operator, and $c$ is the diffusion coefficient. In linear diffusion, $c$ is a constant, thus the diffusion will not only reduce noise, but also smooth feature edges. The anisotropic diffusion filter takes an adaptive $c$, which is no longer a constant but a function that adaptively changes its value according to the local geometry. For instance, $c$ can have small values at the feature edges whereas have large values within the interior region of the features [34]. Thus the anisotropic diffusion filter can reduce the measurement noise, while preserving feature geometries.

Morphological filters are other types of nonlinear filters that can be used to suppress the outliers in measurement data. By moving spherical or flat structuring elements across measurement data, morphological closing operation can remove the pierces, and in opposite morphological opening operation can suppress the spikes [35]. Podulka *et al.* [13] presented an example of using a morphological closing filter with two sizes of flat structuring elements to detect the spikes on a cylinder liner profile. Since the result of morphological filters depends on the local geometry of the surface and the geometry of the structuring element, morphological filters are superior to the linear Gaussian filter in preserving significant surface features in some cases. However, when applying morphological filters to structured surfaces, the size of structure element, e.g., the radius of the sphere, is restricted by the geometry of local features on the surface, and thus might not be able to suppress the outliers and preserve the edge of surface features simultaneously. Another problem with morphological filters is that they are more computationally expensive in comparison to the Gaussian filter, and thus can be slow for large areal measurement data set [36].

Multiscale decomposition can be viewed as multiple bandpass filters, which decomposes the input signal into a sequence of wavelength bandwidth, while keeping their space/time information [15]. Wavelet decomposition, as a widely used multiscale approach, has successfully found its application in surface texture characterization. Machined surface topography usually contains multiscale topography signature from multiple process resources. Wavelet techniques were used to decompose the surface texture into different scales to evaluate tool marks, machining vibrations, and machine tool errors [37], [38]. The wavelet techniques can be of great use if the outliers exist in limited bandwidths. However, if the outliers spread in broad bandwidths, removing those bandwidths will cause the loss of other valuable information on other features of surface topography [8]. Le Goic *et al.* [10] proposed an alternative multiscale method based on the discrete modal decomposition, which can generate a discrete spectrum of the surface. The outliers are identified by the Grubbs' test with the merit of associating a risk level to the threshold for outlier detection. The outlier identification is performed at different scales of the measured surface.

Other practical outlier correction methods in optical measurement of engineering surfaces include material ratio curve and statistical approaches. Podulka *et al.* [13] developed an empirical method to detect spikes on the cylinder liner surfaces. The material ratio curve of the surface was truncated at the empirical thresholding ratio of 0.13% in order to remove the spikes from the underlying surface. Wang *et al.* [32] implemented ten statistical methods for outlier detection

which can be used for surface texture measurement. They generated a database of experimental data on the basis of two engineering surfaces, but coupled with varying levels of surface roughness, percentage of outliers, the amplitude of outliers and uncertainty, and compared the performance of these statistical methods. These aforementioned methods, however, are limited to stochastic surfaces. Yang *et al.* [39] proposed a filtering technique for structured surface using combined sparse regularizes. The use of two regularization terms with the first and second-order derivative, respectively, can divide filtered data into a piecewise constant and a piecewise smooth part, which enables the filter to smooth the noise while preserving edges and corners of structured geometries.

## III. Clustering Filter

### A. Basic Principle of Clustering Filter

The clustering filter used to reduce noise and suppress outliers of optical measured surfaces is originally developed by Wong [40], [41]. Clustering is a technique to divide a given data set into a set of compact groups. The clusters are scale-dependent. At the coarsest scale, all the data can be viewed as a single cluster; at the finest scale, each data in the data set are a cluster. The clustering filtering from the coarse scales to the fine scales is a scale-space decomposition and consistent with the physical melting process.

The center of a cluster of data is derived by maximizing the data entropy subject to a linear cost constraint. Given a 1-D data set $\{(x_i, z_i) : i = 1, \ldots, n\}$, where $z_i$ represents the height of the profile at the sampling position $x_i$, and suppose $\{x, z\}$ is its cluster center. The cost functions of each $x_i$ and $z_i$ are $e_x(x_i) = \|x - x_i\|^2$ and $e_x(z_i) = \|z - z_i\|^2$, where is $\|\ldots\|^2$ the squared Euclidean distance. Let $P$ denotes the contribution of $(x_i, z_i)$ to the cluster center $\{x, z\}$. Maximizing the entropy $S = -\sum_i P_i \log P_i$ subject to the two linear constraints $\sum_i P_i e_x(x_i) = C_x$ and $\sum_i P_i e_z(z_i) = C_z$ leads to the following equation [37]:

$$z' = \frac{\sum_i z_i \omega_i e^{-\beta(z_i - z)^2}}{\sum_i \omega_i e^{-\beta(z_i - z)^2}} \tag{1}$$

where $\omega_i = e^{-\alpha\|x_i - x\|^2}$, $\alpha$ and $\beta$ are two parameters which will be described in Section III-B. $z'$ can be interpreted as a cluster center in height and $\omega_i$ as its weight to give $z_i$. The filter essentially performs clustering for each sampled point. This process can be cast as a robust estimation. A cluster center is naturally a robust estimate of the data associated with it. Thus, the output of the clustering filter should be a better representation of the original signal/surface.

### B. Algorithm for the Clustering Filter

Fig. 1 illustrates the flowchart of the algorithm used to compute the clustering filter. Initially, the algorithm starts by inputting the $\alpha$ parameter so that the region of points which will contribute to the determination of the clustering center can be targeted. The $\beta$ parameter used in the equation is initially set to $\beta = (1/2\sigma_z^2)$, where $\sigma_z$ is a measure of local variance given by $\sigma_z^2 = (\sum_i (z_i - \bar{z})^2 \omega_i / \sum_i \omega_i)$, and $\bar{z}$ is a measure of local average and given by $\bar{z} = (\sum_i z_i \omega_i / \sum_i \omega_i)$. Then, iteratively applying (1) yields (2), which always converges

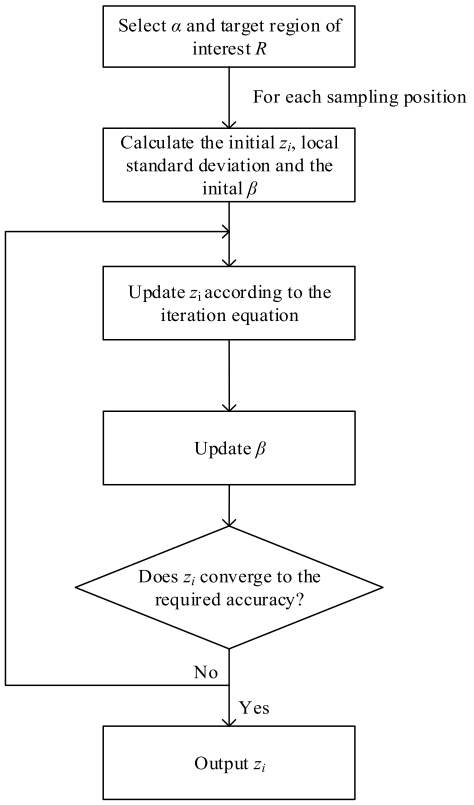



Fig. 1. Flowchart of the algorithm of the clustering filter to determine the clustering center.

$$z^{(t+1)} = \frac{\sum_i z_i \omega_i e^{-\beta^{(t)}(z^{(t)} - z_i)^2}}{\sum_i \omega_i e^{-\beta^{(t)}(z^{(t)} - z_i)^2}}. \tag{2}$$

The initial guess of the clustering center $z'$ can take $\bar{z}$ as the starting point. In each iteration, the clustering center $z^{(t)}$ is updated, and so does $\beta$. In practice, a small tolerance with enough precision is set for the termination of the iteration process. The same procedure is applied to each point of the measured data set. The generated cluster centers $(x, z)$ are taken as the filtered data.

### C. Control Parameters

Equation (2) has two input parameters $\alpha$ and $\beta$. $\alpha$ is a measure of scale in the input space. A large $\alpha$ implies that only points close to $x$ have significant $w_i$. For example, if $\alpha = \infty$, then $w_i = 1$ when $x = x_i$ and 0 otherwise, which means every point is preserved perfectly. Conversely, a small $\alpha$ implies that more neighbors of $x$ can contribute. When $\alpha = 0$, each datum is weighted equally and the estimate is the same for every point. $\beta$ is a measure of scale in the output space and is determined by the local variation of the neighborhood data. $\beta$ controls the degree that an edge of the features should be preserved. Increasing $\beta$ gives strong edge preservation. In the algorithm presented in Fig. 1, $\beta$ is adaptively updated in response to the local data variation, thus it possesses the ability to perverse the feature edges.

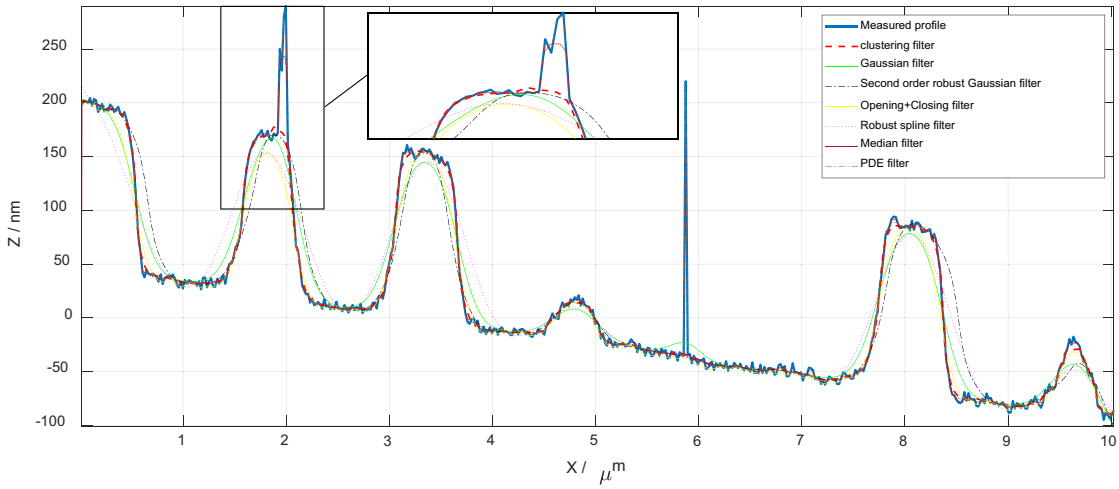



Fig. 2. Comparison of the cluster filtering and other filtration techniques on a measured profile from a CD surface.

Fig. 2 illustrates an example where a measured profile from a CD surface is filtered by the clustering filter ($\alpha = 13$) as well as other filtration techniques, including the linear Gaussian filter ($\lambda_c = 0.8$ $\mu$m), the second-order robust Gaussian filter ($\lambda_c = 0.8$ $\mu$m), the morphological filter (opening filter followed by closing filter with disk radius 0.5 $\mu$m), the robust spline filter ($\lambda_c = 0.8$ $\mu$m), the median filter (window width 50 pixels) and the partial differential equation (PDE) filter, i.e. the anisotropic diffusion filter. The profile is showing a couple of significant pits intentionally made for the purpose of data storage. All these pit features have sharp edges. Two simulated spike outliers are added to the profile data, i.e., a wide spike on the left edge of the second pit, and a single sharp outlier on the bottom area of the middle right pit. The simulated profile and all the applied filters are implemented by MATLAB.

The filters are compared in terms of their capability of outlier suppression and edge preservation. The Gaussian filter, although it can generate a very smooth filtered data, is unable to keep the corner geometries of pits. It is also severely biased by the single sharp outlier. The second-order robust Gaussian filter can suppress two spike outliers. However, it still suffers from rounding pit geometries. The morphological filter takes the combination of the opening operation followed by the closing operation, which is designed to suppress spike outliers. Similar to the robust Gaussian filter, the corner geometries are significantly smoothed. Even worse is the robust spline filter in preserving the corners. The median filter can preserve the corners and cope with the single outlier (salt-and-pepper outlier), but fails to smooth the wide spike outlier on the second pit. The PDE filter is able to retain the corner geometries while suppressing the noises. However, it is unable to deal with the spike outliers. On the contrary, the clustering filter can substantially keep the corner geometries, while perfectly suppress the outliers.

## IV. Interferometric Measurement and Cluster Filtering of Measured Data Sets

The proposed clustering filter can be applied to both profile and areal surface data. In this section, it is applied to the measurement data obtained from an in-house developed interferometric sensor known as single-shot dispersive profile interferometer (SDPI) and a commercial white light interferometer (WLI)–Taylor Hobson Coherence Correlation Interferometry (CCI) 3000.

### A. Optical Setup of SDPI and WLI

The basic configuration of the developed SDPI system is shown in Fig. 3 [42], which includes four modules, namely the light source, the optical probe, the spectrometer, and the console. A halogen bulb with broadband spectrum provides the white light illumination for the system. A 4× Michelson interferometric objective serves as an optical probe and observes the tested surface without contact. A number of fringes will occur when the optical path difference (OPD) between the reference arm and measurement arm of the interferometer is within the scope of the coherence length. After passing through the spectrometer, the interference beam is then spectrally decomposed along the rows of the charge-coupled device (CCD) pixels. The generated interferogram is a 2-D image with one axis (horizontal) providing the phase information encoded as a function of wavenumber, and the other (vertical) representing the length of the measured surface profile [43]. Therefore, SDPI system can achieve surface profile measurement with a single spectral interferogram. During the signal processing, fringe order determination can be carried out to retrieve the phase information with higher accuracy, hence, measurement resolution can be enhanced. However, this method suffers from discrete spikes errors equaling to $\lambda/2$ or an integer multiple of this value presenting in the height profiles.

Fig. 4 shows the schematic of a WLI system (for full-field measurement). This technique uses a broadband illumination as well and generates interference fringes localized in space due to the low temporal coherence [44]. By combining with the vertical scanning techniques, a series of sequential interferograms are captured during the mechanical scanning. The maximum visibility of the fringes occurs when the OPD between the reference arm and measurement arm equals zero.

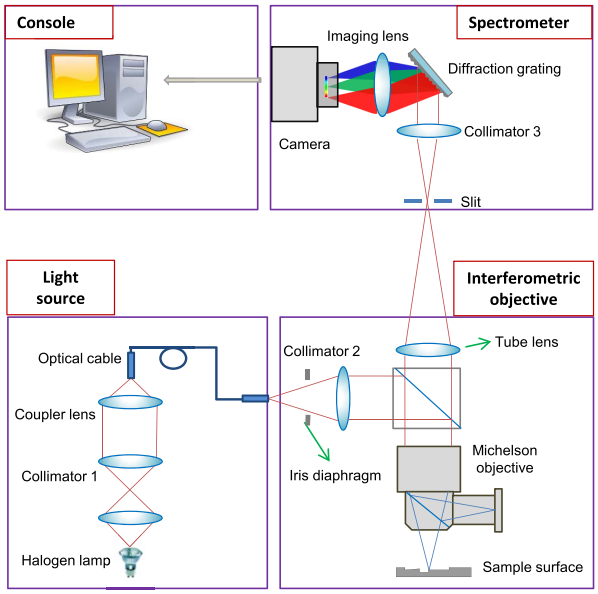



Fig. 3. Schematic of a SDPI system.

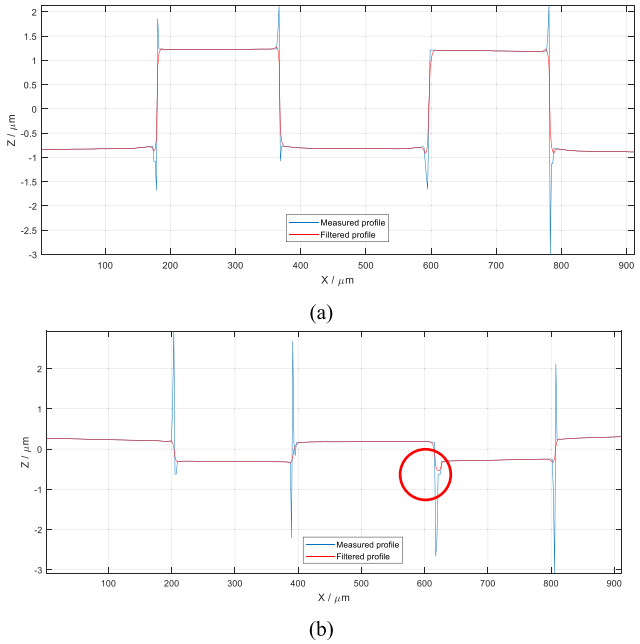



Fig. 5. Surface profiles measured by SDPI and filtered by the clustering filter. (a) 2.1-$\mu$m NPL step height sample. (b) 500-nm NPL step height sample.

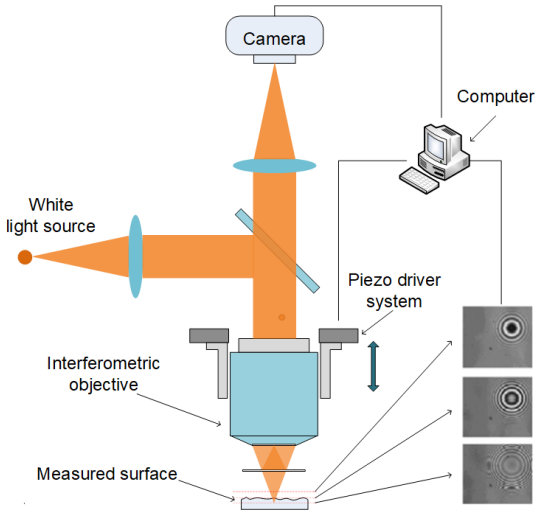



Fig. 4. Schematic of a WLI system.

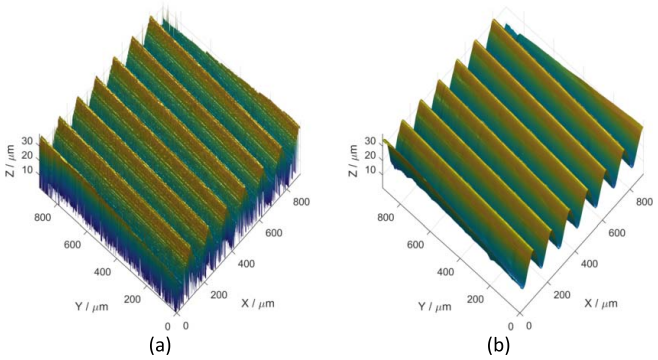



Fig. 6. Clustering filtering of a V-grooves surface with $\alpha$ value 16. (a) Raw measured surface. (b) Filtered surface.

The visibility falls quickly when the OPD increases. Surface topography can be acquired after tracking all coherence peaks or phase retrieval within the field of view of the objective [12], [18].

### *B. Cluster Filtering of SDPI Measured Profiles*

Two Step Heights (2.1 $\mu$m and 500 nm) from the National Physical Laboratory (NPL) were measured by SDPI, see Fig. 5. Large spikes appear on two corners of each step height. Particularly in the case of 500-nm NPL sample, the amplitudes of spikes are four to five times higher than the step height itself, see Fig. 5(b). The clustering filter with $\alpha$ value 12 is applied to filter the measured profiles. In Fig. 5, large spikes are effectively suppressed, while the geometry of step heights are successfully preserved. This facilitates the subsequent surface characterization, such as form/roughness measurement and step height evaluation. It is, however, noticed that in the highlighted area of Fig. 5(b) that the filtered profile sank down about 100–150 nm, which is inconsistent with the original geometry. This unexpected sinking is caused by the relatively broad spatial distribution of the highlighted spike which is wider than the rest of spikes. Wide spikes/outliers can impede the correct determination of the clustering center and thus generate undesirable filtering results. To enable the successful application of the clustering filter, the bulk of measured data shall be "good" data such that the anomalies can be corrected using these good data. If not, it is unlikely that the clustering filter will generate reasonable results, as this is almost the same for other post processing techniques mentioned in Section II-B.

### *C. Cluster Filtering of WLI Measured Surfaces*

Three structured surfaces, including a V-grooves surface, a gauge block surface, and a nanometer step surface, are measured by WLI, see Figs. 6–8. Both high-frequency noise and

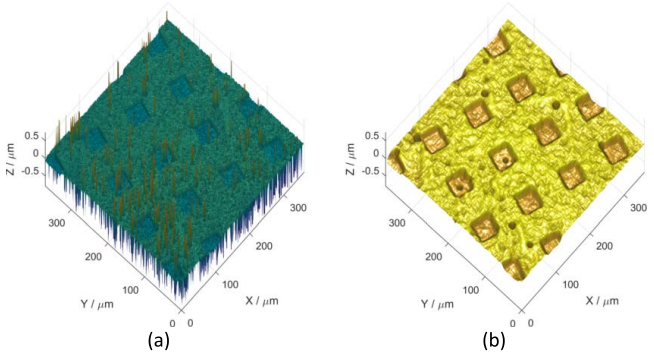



Fig. 7. Clustering filter of a gauge block surface with $\alpha$ value 12. (a) Raw measured surface. (b) Filtered surface.

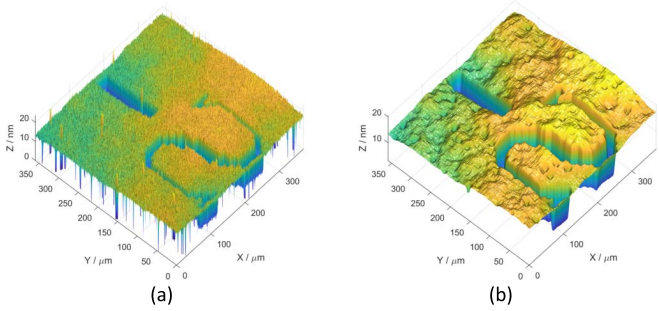



Fig. 8. Clustering filter of a nanometer step surface with $\alpha$ value 12. (a) Raw measured surface. (b) Filtered surface.

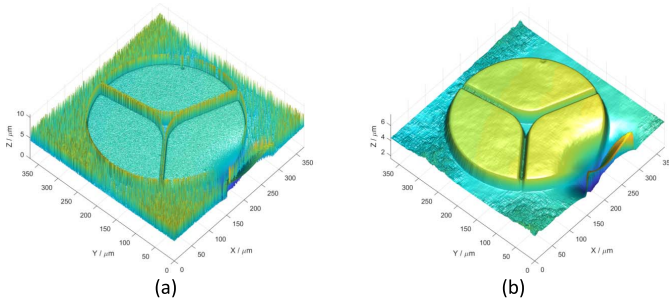



Fig. 9. Clustering filter of a microstructure surface with $\alpha$ value 12. (a) Raw measured surface. (b) Filtered surface.

large spikes appear on the measured surfaces. Their presence not only contracts the scale of original surface topography and thus makes the visualization of surface topographical features very difficult, but also distorts the examination of surface geometries.

Fig. 6(b) presents the smoothed surface filtered by the clustering filter with $\alpha$ value 16. This allows the accurate evaluation of the V shapes on the surface, including the height, the spatial distribution, and the inclination angle. Both surfaces in Figs. 7 and 8 contain nanoscale features. While large spikes impede the assessment of these nanoscale geometries, the surface topography is almost submerged by widespread noises. After applying the clustering filter, both nanoscale geometries and surface topography are clearly presented. Particularly, in Fig. 7(b), a bunch of small circular dents, originally veiled by large spikes, now clearly present on the surface. Fig. 9(a) presents a microstructure surface. The microstructures are buried in spike noises and thus it is hard to recognize their true geometries. In comparison, these geometries are well presented on the filtered surface.

## V. Conclusion

While optical measurement techniques are becoming increasingly popular in surface metrology due to their superior abilities over the tactile counterpart, there are a couple of shortcomings that limit the uptake of these technologies into a wider range of applications. A fatal problem of using optical sensors to measure structured surfaces is that they tend to produce noises and outliers, particularly at the edge of sharp geometry transitions. To alleviate the negative impacts of these undesirable noises and outliers on the characterization of surface topography and micro/nanoscale geometries, the clustering filter is implemented, which is an iterative process to find the aggregation center of a cluster of points. The clustering filter can suppress the noises and outliers in measurement data sets while preserving the edge of geometrical features. This capability is evidenced by the examples of filtering five structured profiles/surfaces, including two profiles measured by SDPI and three surfaces measured by WLI. It, however, shall be noted that if the outliers have wide spatial amplitudes, which severely bias the measurement data, the clustering filter might not generate reasonable results, as this is almost true for all post processing techniques.

Potential applications of the clustering filter are optical measurement of laser texturing produced surfaces and additively manufactured surfaces. The surface topographies generated by these technologies are usually rough and often consist of local sharp microgeometries. These surface topographies also present inconsistent optical properties. As a result, a large number of noises, unmeasured points, and large outliers are often observed in these measurements. It is, therefore, interesting to examine the capability of the clustering filter on filtering these kinds of challenging surfaces.

## References

[1] X. Jiang, "Precision surface measurement," *Philos. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 370, no. 1973, pp. 4089–4114, 2012.

[2] A. Magnani, A. Pesatori, and M. Norgia, "Real-time self-mixing interferometer for long distances," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 7, pp. 1804–1809, Jul. 2014.

[3] J. H. Galeti, R. T. Higuti, E. C. N. Silva, and C. Kitano, "Nanodisplacement measurements of piezoelectric flextensional actuators using a new interferometry homodyne method," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1256–1265, May 2015.

[4] F. Gao, J. Coupland, and J. Petzing, "V-groove measurements using white light interferometry," *Photon*, vol. 6, pp. 4–7, Sep. 2006.

[5] *Geometrical Product Specifications (GPS)—Surface texture: Areal—Part 6: Classification of Methods for Measuring Surface Texture*, ISO Standard ISO 25178-6:2010, 2007.

[6] C. L. Giusca, R. K. Leach, F. Helary, T. Gutauskas, and L. Nimishakavi, "Calibration of the scales of areal surface topography-measuring instruments: Part 1. Measurement noise and residual flatness," *Meas. Sci. Technol.*, vol. 23, no. 3, Mar. 2012, Art. no. 035008.

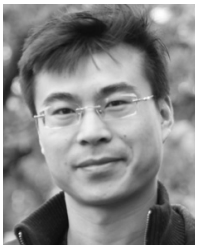
[7] H. Fujimoto, M. Tanaka, and K. Nakayama, "Noise reduction in an optical interferometer for picometer measurements," *IEEE Trans. Instrum. Meas.*, vol. 44, no. 2, pp. 471–474, Apr. 1995.

[8] M. F. Ismail, K. Yanagi, and A. Fujii, "An outlier correction procedure and its application to areal surface data measured by optical instruments," *Meas. Sci. Technol.*, vol. 21, no. 10, Oct. 2010, Art. no. 105105.

[9] *Geometrical Product Specifications (GPS)—Filtration—Part 1: Overview and Basic Concepts*, ISO Standard 16610-1:2015, 2015.
[10] G. Le Goic, C. A. Brown, H. Favreliere, S. Samper, and F. Formosa, "Outlier filtering: A new method for improving the quality of surface measurements," *Meas. Sci. Technol.*, vol. 24, no. 1, Jan. 2013, Art. no. 015001.
[11] F. Gao, R. K. Leach, J. Petzing, and J. M. Coupland, "Surface measurement errors using commercial scanning white light interferometers," *Meas. Sci. Technol.*, vol. 19, no. 1, Jan. 2008, Art. no. 015303.
[12] A. Harasaki, J. Schmit, and J. C. Wyant, "Improved vertical-scanning interferometry," *Appl. Opt.*, vol. 39, no. 13, pp. 2107–2115, May 2000.
[13] P. Podulka, P. Pawlus, P. Dobrzański, and A. Lenart, "Spikes removal in surface measurement," *J. Phys., Conf. Ser.*, vol. 483, no. 1, Mar. 2014, Art. no. 012025.
[14] X. Jiang, P. J. Scott, D. J. Whitehouse, and L. Blunt, "Paradigm shifts in surface metrology. Part II. The current shift," *Philos. Trans. Roy. Soc. A Math., Phys. Eng. Sci.*, vol. 463, no. 2085, pp. 2071–2099, 2007.
[15] X. J. Jiang and D. J. Whitehouse, "Technological shifts in surface metrology," *CIRP Ann.*, vol. 61, no. 2, pp. 815–836, 2012.
[16] L. De Chiffre, H. Kunzmann, G. Peggs, and D. Lucca, "Surfaces in precision engineering, microengineering and nanotechnology," *CIRP Ann.*, vol. 52, no. 2, pp. 561–577, 2003.
[17] A. Bruzzone, H. Costa, P. Lonardo, and D. Lucca, "Advances in engineered surfaces for functional performance," *CIRP Ann.*, vol. 57, no. 2, pp. 750–769, 2008.
[18] Q. Vo, F. Fang, X. Zhang, and H. Gao, "Surface recovery algorithm in white light interferometry based on combined white light phase shifting and fast Fourier transform algorithms," *Appl. Opt.*, vol. 56, no. 29, pp. 8174–8185, Oct. 2017.
[19] J. Niehues, P. Lehmann, and K. Bobey, "Dual-wavelength vertical scanning low-coherence interferometric microscope," *Appl. Opt.*, vol. 46, no. 29, pp. 7141–7148, Oct. 2007.
[20] Y.-S. Ghim and A. Davies, "Complete fringe order determination in scanning white-light interferometry using a Fourier-based technique," *Appl. Opt.*, vol. 51, no. 12, pp. 1922–1928, Apr. 2012.
[21] M. Roy, J. Schmit, and P. Hariharan, "White-light interference microscopy: Minimization of spurious diffraction effects by geometric phase-shifting," *Opt. Express*, vol. 17, no. 6, pp. 4495–4499, Mar. 2009.
[22] J. Niehues and P. Lehmann, "Improvement of lateral resolution and reduction of batwings in vertical scanning white-light interferometry," *Proc. SPIE*, vol. 8082, May 2011, Art. no. 80820W.
[23] J. Seewig, "Linear and robust Gaussian regression filters," *J. Phys., Conf. Ser.*, vol. 13, no. 1, pp. 254–257, Jan. 2005.
[24] W. Zeng, X. Jiang, and P. J. Scott, "Fast algorithm of the robust Gaussian regression filter for areal surface analysis," *Meas. Sci. Technol.*, vol. 21, no. 5, May 2010, Art. no. 055108.
[25] *Geometrical Product Specifications (GPS)—Filtration—Part 30: Robust Profile Filters: Basic Concepts*, ISO Standard 16610-30:2015, 2015.
[26] B. Zhong and K.-K. Ma, "On the convergence of planar curves under smoothing," *IEEE Trans. Image Process.*, vol. 19, no. 8, pp. 2171–2189, Aug. 2010.
[27] D. J. Whitehouse, "Theoretical enhancement of the Gaussian filtering of engineering surfaces," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 469, no. 2158, Oct. 2013, Art. no. 20130184.
[28] W. Zeng, X. Jiang, and P. Scott, "A generalised linear and nonlinear spline filter," *Wear*, vol. 271, nos. 3–4, pp. 544–547, Jun. 2011.
[29] M. Krystek, "Form filtering by splines," *Measurement*, vol. 18, no. 1, pp. 9–15, May 1996.
[30] T. Goto, J. Miyakura, K. Umeda, S. Kadowaki, and K. Yanagi, "A robust spline filter on the basis of $L_2$-norm," *Precis. Eng.*, vol. 29, no. 2, pp. 157–161, Apr. 2005.
[31] P. D. Ruiz, "Evaluation of a scale-space filter for speckle noise reduction in electronic speckle pattern interferometry," *Opt. Eng.*, vol. 37, no. 12, p. 3287, Dec. 1998.
[32] C. Wang, J. Caja, and E. Gómez, "Comparison of methods for outlier identification in surface characterization," *Measurement*, vol. 117, pp. 312–325, Mar. 2018.
[33] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[34] W.-H. Zeng, X. Jiang, P. J. Scott, and L. Blunt, "Diffusion filtration for the evaluation of MEMs surface," *Int. J. Precis. Eng. Manuf.*, to be published.
[35] S. Lou, X. Jiang, and P. J. Scott, "Correlating motif analysis and morphological filters for surface texture analysis," *Measurement*, vol. 46, no. 2, pp. 993–1001, Feb. 2013.
[36] S. Lou, X. Jiang, and P. J. Scott, "Geometric computation theory for morphological filtering on freeform surfaces," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 469, no. 2159, Nov. 2013, Art. no. 20130150.
[37] X. Q. Jiang, L. Blunt, and K. J. Stout, "Development of a lifting wavelet representation for surface characterization," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 456, no. 2001, pp. 2283–2313, Sep. 2000.
[38] X. Chen, J. Raja, and S. Simanapalli, "Multi-scale analysis of engineering surfaces," *Int. J. Mach. Tools Manuf.*, vol. 35, no. 2, pp. 231–238, Feb. 1995.
[39] H. Yang, X. Zhang, H. Zhang, X. He, H. Xiao, and M. Xu, "Feature-preserving filtering for micro-structured surfaces using combined sparse regularizers," *Measurement*, vol. 104, pp. 278–286, Jul. 2017.
[40] Y.-F. Wong, "Clustering data by melting," *Neural Comput.*, vol. 5, no. 1, pp. 89–104, Jan. 1993.
[41] Y.-F. Wong, "Nonlinear scale-space filtering and multiresolution system," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 774–787, Jun. 1995.
[42] D. Tang, P. Kumar, F. Gao, and X. Jiang, "Phase retrieval algorithm for line-scan dispersive interferometry," *Proc. SPIE*, vol. 10827, Jul. 2018, Art. no. 108270F.
[43] D. Tang, F. Gao, and X. Jiang, "On-line surface inspection using cylindrical lens-based spectral domain low-coherence interferometry," *Appl. Opt.*, vol. 53, no. 24, pp. 5510–5516, 2014.
[44] D. Malacara, *Optical shop testing*, 3rd ed. Hoboken, NJ, USA: Wiley, 2007, pp. 711–712.

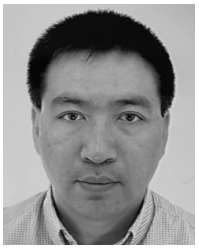
**Shan Lou** received the B.Sc. and M.Sc. degrees from Central South University, Changsha, China, in 2002 and 2006, respectively, and the Ph.D. degree in precision metrology from the University of Huddersfield, Huddersfield, U.K., in 2013.

He worked as a Coordinate Measurement Machine Software Developer with Hexagon Metrology Ltd., Shanghai, China. He is currently a Senior Research Fellow with the Engineering and Physical Sciences Research Council (EPSRC) Future Metrology Hub, Centre for Precision Technologies (CPT), University of Huddersfield. His research interests are precision geometrical metrology, metrology for additive manufacturing, and X-ray computed tomography metrology.



**Dawei Tang** received the B.Sc. degree in mechanical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2008, the M.Sc. degree in optical engineering from the Graduate University of Chinese Academy of Sciences (UCAS), Beijing, China, in 2011, and the Ph.D. degree in optical engineering from the University of Huddersfield, Huddersfield, U.K., in 2016.

He is currently a Research Fellow with the Engineering and Physical Sciences Research Council (EPSRC) Future Metrology Hub, Centre for Precision Technologies (CPT), University of Huddersfield. His research interests include in-process surface metrology, precision engineering, and optical instrumentation.



**Wenhan Zeng** received the B.Sc. degree in mechanical engineering, the M.Sc. degree in precision instrumentation, and the Ph.D. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1997, 2000, and 2005, respectively.

He is currently a Principal Research Fellow with the Engineering and Physical Sciences Research Council (EPSRC) Future Metrology Hub, Centre for Precision Technologies, University of Huddersfield, Huddersfield, U.K. His current research interests include precision metrology, surface metrology, image and signal processing as well as machining learning and deep learning.

**Tao Zhang** received the B.E. degree in electronic science and technology from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2008, the M.E. degree in optical engineering from the Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences (CAS), Changchun, China, in 2011, and the Ph.D. degree in mechanical engineering from the University of Huddersfield, Huddersfield, U.K., in 2018.

He is currently a Research Fellow with the Quality Engineering and Smart Technology (QUEST) Research Centre, Brunel University London, Uxbridge, U.K. He is working on a 3-D printing robot and the Internet of Things (IoT) infrastructure for smart buildings. His current research interests include robotics, the IoT, surface metrology, computer vision, and data analytics.

**Feng Gao** received the B.Sc. and M.Sc. degrees in precision measurement and instrumentation from Tianjin University, Tianjin, China, in 1985 and 1988, respectively, and the Ph.D. degree in precision measurement and instrumentation from Coventry University, Coventry, U.K., in 2001.

He worked as an Assistant Engineer and an Engineer with the National Institute of Metrology of China, Beijing, China, a Visiting Scholar with the Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany, and a Research Associate with Loughborough University, Loughborough, U.K. He joined as a Senior Research Fellow with the Centre for Precision Technologies (CPT), University of Huddersfield, Huddersfield, U.K., where he is currently a Reader of metrology and instrumentation. His main research interests are in the field of 3-D form measurement using fringe projection and deflectometry technologies and interferometry for on-line surface inspection.

**Hussam Muhamedsalih** received the B.Sc. degree in laser and optoelectronics engineering from Nahrain University, Baghdad, Iraq, in 2001, and the M.Sc. degree in control systems engineering and instrumentation and the Ph.D. degree in interferometry instrumentation from the University of Huddersfield, Huddersfield, U.K., in 2008 and 2013, respectively.

Since 2013, he has been with the Research Group, Centre for Precision Technologies (CPT), University of Huddersfield. He is currently a Senior Research Fellow with CPT, focusing on the development of on-line/in-process optical instruments for surface measurement, including the optics design and signal analysis.

**Xiangqian Jiang** received the Ph.D. degree in engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1995, and the D.Sc. degree from the University of Huddersfield, Huddersfield, U.K., in 2007.

She is currently a Professor of precision metrology with the Centre for Precision Technologies, University of Huddersfield. She is also the Director of the Engineering and Physical Sciences Research Council (EPSRC) Future Metrology Hub and the Renishaw/Royal Academy of Engineering Chair in precision metrology. Her research mainly lies in surface metrology, precision engineering and manufacturing, and optic instrumentation.

Dr. Jiang is a fellow of the Royal Academy of Engineering (FREng), Royal Society of Arts (FRSA), and Institute of Engineering Technologies (FIET).

**Paul J. Scott** received the B.Sc. degree in mathematics and the M.Sc. and Ph.D. degrees in statistics from Imperial College London, London, U.K., in 1979, 1980, and 1983, respectively.

He is currently a Professor with the School of Computing and Engineering, University of Huddersfield, Huddersfield, U.K. He is a Visiting Industrial Professor with Taylor Hobson Limited, Leicester, U.K. His research interests are foundations of specifying and characterizing geometrical products, stable, robust and fast algorithms for computational geometry, and philosophy of the measurement of geometry.

Dr. Scott is a fellow of Royal Statistical Society (FRSS). He is the Taylor Hobson Chair for Computational Geometry.