A REVIEW OF DIGITAL FILTERING IN EVALUATION OF SURFACE ROUGHNESS

Baofeng He, Haibo Zheng, Siyuan Ding, Ruizhao Yang, Zhaoyao Shi

Abstract

Surface roughness is an important indicator in the evaluation of machining and product quality, as well as a direct factor affecting the performance of components. A rapidly developing filtering technology has become the main means of extracting surface roughness. The International Organization for Standardization (ISO) is constantly updating and improving the standard system for filtering technology in order to meet the requirements of technological development. Based on the filters already accepted by the international standard ISO 16610, this study briefly introduces the filtering principle of each filter, reviews the development of each filter in the application of surface roughness, and compares the advantages and limitations of their individual performances. The application range of each filter is summarized and, finally, the future direction of the digital filtering used in surface roughness is extrapolated.

Keywords: surface roughness, profile filtering, areal filtering, surface engineering.

1. Introduction

Surface roughness is an important indicator of the microgeometry error of mechanical parts. Surface roughness is generally caused by the machining method and other factors such as the friction between the cutting tool and the surface of the processed part, plastic deformation during metal cutting and high frequency vibration in the process system. It affects the fitting of parts, corrosion resistance, contact stiffness, fatigue resistance, sealing, and part appearance. Therefore, the extraction and evaluation of surface roughness play an important role in ensuring the reliability and stability of parts, reducing the friction coefficient, improving the working accuracy and sensitivity of machinery and instruments, and reducing the surface wear of parts. The surface topography signal consists of three components: surface roughness, surface waviness and macroscopic form error [1]. The low frequency, intermediate frequency and high frequency components in the signal correspond to the macroscopic form error, surface waviness, and surface roughness, respectively. The premise of evaluating surface roughness is to effectively separate...
the surface roughness from the surface topography signal. At present, filtering technology is the main method used to separate surface roughness and other low frequency components. It can decompose the measured surface topography signal in the frequency domain, yielding different components. The wavelength components are separated into well-defined bandwidths without having to represent the surface topography as a specific function [2].

In the 1950s, the passive filtering technology became more mature. In the 1960s, with the continuous development of computer technology, microelectronic technology, active filters developed rapidly and were widely used. The “2RC” filter specified in the 1975 first edition of international standard ISO 3274 was applied to surface topography separation, but its own non-linearity seriously distorted the profile shape and could not truly reflect the surface texture [3]. Whitehouse also noticed the drawbacks of the 2RC filter, and proposed the concept and method of the phase correction filter [4]. In the 1970s, the rapid development of digital computers enabled digital analysis techniques that could be applied to surface topography measurements. For example, Raja et al. used a regression algorithm to digitally process the 2RC filter, which improved the speed of profiles filtering [5]. In the 1980s, numerous entities in industrial countries focused on investigating various new types of filter performance, and gradually expanded the scope of application. In the 1990s, classical Gaussian filters were widely used for surface texture extraction, replacing the 2RC filter as the standardized filter [6]. However, there are some disadvantages of the Gaussian filter, such as the end effect. The introduction of Gaussian regression and spline filter can effectively suppress the end effect [7]. At the beginning of the 21st century, many experts and researchers introduced the theory of robust estimation into the filters, which has enhanced the ability of filters to process outlier signals [8, 9].

As a significant operation in the Geometrical Product Specification and Verification (GPS) system, filtering technology plays an important role in the measurement and evaluation of surface microtopography. The 1996 Technical Committee ISO/TC 213 revised the filter technology standards ISO 11562 [6] and ISO 13565 [10]. ISO 11562 specified phase correction filters for surface profile measurement and delineated the process for separating the roughness and waviness of surface profiles. This standard also replaced the traditional 2RC filter standard. ISO 13565 proposed two Gaussian filters with the “Rk” filter in order to suppress the effects of deep valleys, but did not overcome the effects of larger and deeper valleys, nor did it inhibit peaks. The continuous development of microelectronic technology, manufacturing technology, bioengineering and other technologies has imposed higher requirements for the extraction and characterization of surface texture. ISO/TC 213 continuously updates and refines the filtering technology system to meet the needs of technological development, and replaces the traditional standard ISO 11562 by revising the new standard ISO 16610. The current published filter standards are shown in Fig. 1, however, it does not cover all the existing filtering techniques. The nomenclature and taxonomy are given in Table 1 and Table 2. Although international standards have formed a relatively complete filtering technology system, their application scope has not been discussed and summarized in detail, nor can the conversion of existing theoretical systems to applied technology be guaranteed.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Type</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>F = Filter</td>
<td>P = Profile (2D)</td>
<td>L = Linear</td>
</tr>
<tr>
<td></td>
<td>A = Areal (3D)</td>
<td>M = Morphological</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R = Robust</td>
</tr>
</tbody>
</table>

Table 1. Classification of filters.

218
This study first introduces the principle of standard profile filters and areal filters, and next it reviews the development and application of each filter in the evaluation of surface roughness, and then compares the advantages and limitations of their individual performances. For the filtering characteristics of each filter, the application range for surface roughness is summarized. Finally, the development direction of digital filtering in surface roughness is discussed.

2. Linear profile filters

2.1. Gaussian profile filters

In 2011, the international standard ISO 16610 Part 21 included Gaussian profile filters [11]. The Gaussian filter, utilizing Gaussian functions as conversion weight functions, is a zero-phase-shift filtering method that completely corrects the phase distortion defect of traditional 2RC
filters, and was the first type of phase correction filter certified with international standards. The Gaussian function is an ideal filter which can satisfy the minimum product of time width and frequency width defined by energy moment density. The information of roughness and waviness can be obtained with a Fast Fourier transform (FFT), and Gaussian profile filters are the most scale used profile filters.

There are two different types of surface profiles: open profiles and closed profiles. An open profile refers to a surface profile having two endpoints and a limited length, such as a core-like part for molding the inner surface of a plastic part. The weight function of the Gaussian profile filter for filtering an open profile is

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

where \( x \) is the distance to the center of the Gaussian weight function, \( \lambda_c \) is the cut-off wavelength. To ensure that the Gaussian profile filter has a transmission rate of 50% at the cut-off wavelength, \( \alpha \) is taken as \( \sqrt{\ln 2/\pi} \approx 0.4697 \). The weight function of the Gaussian profile filter for closed profile filtering is

\[ s(x) = \frac{f_c}{\alpha L} \exp\left(-\pi \left(\frac{x f_c}{\alpha L}\right)^2\right), \]  

where \( x \) is the distance of the center of the weight function along the closed profile, \( L \) is the closed profile length, and \( f_c \) is the cut-off frequency in undulations per revolution (\( f_c = L/\lambda_c \)). The transmission characteristic of a Gaussian filter is determined from the weighting function by means of the Fourier transformation. The transmission characteristic of the long wave component (mean line) for an open profile is

\[ \frac{a_1}{a_0} = \exp\left(-\pi \left(\frac{\alpha \lambda_c}{\lambda}\right)^2\right), \]  

where \( a_0 \) is the amplitude of a sinusoidal wave profile before filtering, \( a_1 \) is the amplitude of this sinusoidal profile in the mean line, \( \lambda \) is the wavelength of this sinusoidal profile. The transmission characteristics of the short-wave component are obtained by equation (1) \(-a_1/a_0\). The transmission characteristic of the long wave component (mean line) for a closed profile is

\[ \frac{a_1}{a_0} = \exp\left(-\pi \left(\frac{\alpha f}{f_c}\right)^2\right), \]  

where \( f \) is the frequency of the sinusoidal profile.

When Gaussian filters are used to filter the open profiles, the so-called end effect will be produced. It should be noted that “edge effect” has been used in many publications. In this paper, “end effect” was used instead of “edge effect”. In fact, they refer to the same phenomenon. The effective part of the weight function at both ends will exceed the edge of the profile. Therefore, the convolution of the weight function and the profile end data will change unexpectedly, that is, the data at both ends of the profile after the filtering are distorted. Extending the ends of the profile is one way to eliminate the end effect. Three common methods for profile expansion are proposed in ISO 16610 Part 28, namely the zero-padding method, linear extrapolation method and the symmetric extension method [12]. The zero-padding method is a simple way to preserve the traversal length after profile filtering. This method, however, is only suitable for profiles
without slopes. The linear extrapolation method uses the least squares method to fit the end effect region profile and obtain the fitted line segment as an extended region. This method is suitable for profiles without slopes as well as those with a certain degree of inclination. The symmetric extension method consists of symmetric expansion of the two endpoints of the end effect region, and includes axis symmetry and point symmetry. The extended profile of this method has a tilting direction that differs significantly from the original tilting direction, and this approach cannot cope with complicated filtering situations. These three methods also have the disadvantage of poor adaptability to a given profile, and it is cumbersome to select the appropriate method for different profiles in actual projects. ISO 16610-28 also proposes a moment retainment criterion. It is a method suitable for all profiles. This technique changes the Gaussian weight function according to the position in the end effect region, and ensures that the total weight of the weight function is 1. Therefore, the Gaussian regression filter eliminates the end effect of the Gaussian filter by modifying the weight function. Fig. 2a shows the weighting function of the zero-order Gaussian regression filter at ten different positions of the profile. In order to show the change of shape near the boundary clearly, only the first 2 mm of the weighting function are drawn in Fig. 2a. When the first moment is 0, a reliable filtering effect can be achieved, and the distortion problem caused by the profile expansion is avoided. To further verify the boundary performance of the Gaussian regression filter, MATLAB is used to generate a normally distributed random noise as shown in Fig. 2b. It is found that the Gaussian filter has an obvious end effect at the boundary, while zero-order Gaussian regression filter can better embrace the profile. Therefore, the Gaussian regression filter is useful when the end effect cannot be tolerated such as a short trajectory measurement.

![Fig. 2. Modified Gaussian weighting function (λc = 0.8 mm) (a) and profile and the waviness profiles using a Gaussian and zero-order Gaussian regression filter (λc = 0.8 mm) (b).](image)

Many experts and scientists carried out in-depth research on the application theory and method of Gaussian profile filters which has laid the foundation for the application of Gaussian profile filters in surface roughness measurement. In 1979, Raja et al. introduced the fast Fourier transform into the traditional Gaussian filter algorithm in order to achieve digital technology for extraction of roughness [13]. In 1996, Krystek proposed a fast and reliable convolution algorithm based on the recursive relationship of weighting functions that speeds up the extraction of surface roughness.
by Gaussian filters [14]. In 1998, Hara et al. proposed an online Gaussian filtering algorithm that not only improves the computational efficiency, but also saves storage space and is very suitable for real-time surface roughness measurement system [15]. In 2000, Yuan et al. used the sampling function \((\sin u/u)^n (n \to \infty)\) to approximate the Gaussian weight function. The Gaussian profile filter constructed by this simplified algorithm has the advantages of fast calculation, high precision, and no phase distortion [16]. In addition, Yuan et al. proposed a fast recursive algorithm for Gaussian profile filters using the central limit theorem and approximation method. Based on the recursive algorithm, the undistorted filtering mean line can be determined [17]. In 2016, Xu et al. applied the error theory to Gaussian filtering which effectively suppressed abnormal signals of the groove surface. Firstly, the abnormal signals are judged and corrected by the 5\(\sigma\) criterion. Then the Gaussian filter is realized with the moving average approximation method to obtain the average of surface profile. Finally, the method achieves robust filtering based on the original profile and the mean line to separate the surface roughness [18]. Compared with other robust methods of improved Gaussian filter, this method abandons the weight function. The algorithm is simple and easy to be applied in engineering. In addition, researchers are also working on the methods for end effect elimination, as listed in Table 3.

Gaussian filters must have three prerequisites for surface roughness measurements: (1) Removal of the form error in order to solve the problem of Gaussian filters not tracking the profile signals with large forms well. (2) The assumption that the signal of the original profile consists of a series of harmonics of different wavelengths. (3) The surface roughness parameter obeys normal distribution. Gaussian profile filters can be applied to the extraction of most types of surface roughness, including grinding, stamping, turning, milling, reaming (the reamer cuts off the micro metal layer from the hole wall of the workpiece), granite (brittle material consisting of a combination of different crystals) and after abrasive blasting, etc. However, there are still some specific engineering surfaces that do not meet these three prerequisites. For example, surfaces with deeper pores, structural feature surfaces (non-normal distribution), surfaces with larger form components (aspheric lenses, F-theta lenses, free-form surfaces), stratified surfaces (honied surfaces), and surfaces for which the measurement area is limited so that the surface of the boundary cannot be reduced. Gaussian profile filters are not robust to outliers such as deep valleys and peaks in surface topography signals. Therefore, the filter mean line is distorted during filtering, affecting the authenticity of the measured data. In the evaluation of surface topography, the default length of international standard is \(7\lambda_c\). Since the algorithm of the Gaussian filters usually has end effects, the middle \(5\lambda_c\) are taken as the evaluation length.
2.2. Spline profile filters

To solve the end effect problem of Gaussian profile filters, ISO 16610 Part 22 recommends spline profile filters for the extraction of two-dimensional surface roughness [26]. The spline profile filter is defined using the matrix decomposition algorithm [27]. At the same time, the standard specifies two conditions, aperiodic and periodic, for open-profile and closed-profile filtering, respectively, with different end conditions corresponding to different filter equations. The spline filter eliminates end effects by adjusting boundary conditions $P$ and $Q$. $P$ is a linear spline term, and $Q$ is a cubic spline term. The filter equation for the aperiodic spline profile filter is

$$P = \begin{bmatrix}
1 & -1 \\
-1 & 2 & -1 \\
& \ddots & \ddots & \ddots \\
& -1 & 2 & -1 \\
& -1 & 1
\end{bmatrix}, \quad Q = \begin{bmatrix}
1 & -2 & 1 \\
-2 & 5 & -4 & 1 \\
1 & -4 & 6 & -4 & 1 \\
& & \ddots & \ddots & \ddots \\
1 & -2 & 1
\end{bmatrix},$$

(5)

$$\left[ 1 + \beta \alpha^2 P + (1 - \beta) \alpha^4 Q \right] w = z,$$

(6)

$$\alpha = \frac{1}{2 \sin(\pi \Delta x / \lambda_c)},$$

(7)

where $z$ is the vector of the profile data before filtering, $w$ is the vector of profile values in the filtered profile, $\lambda_c$ is the cut-off wavelength of the profile filter, and $\Delta x$ is the sampling interval, $\alpha$ is the constant value determined by the cut-off wavelength. $\beta$ is the tension parameter ($0 \leq \beta \leq 1$) which controls the tightness of the data points through the spline curve and also objectively controls the similarity of the transmission characteristics of the spline profile filter and the Gaussian profile filter. The filter mean line value $w$ can be obtained from (6). The vector $w$ represents a long-wave component, while the short-wave component $r$ (the surface roughness value) is obtained by the difference vector $r = z - w$. When filtering the closed profile, matrices $P$ and $Q$ become the corresponding augmented matrices. Part 22 introduces the transmission characteristics based on cubic base splines. The transmission characteristics of the filter for long-waves are approximated by the following expression:

$$\frac{a_1}{a_0} = \left[ 1 + \beta \alpha^2 \sin^2 \frac{\pi \Delta x}{\lambda} + 16(1 - \beta) \alpha^4 \sin^4 \frac{\pi \Delta x}{\lambda} \right]^{-1},$$

(8)

where $a_0$ is the amplitude of the sinusoidal profile before filtering, $a_1$ is the amplitude of the sinusoidal profile in the long-wave component, and $\lambda$ is the wavelength of the sinusoidal profile ($\lambda \geq 2 \Delta x$). The short-wave transmission characteristic approximation expression is

$$\frac{a_2}{a_0} = 1 - \frac{a_1}{a_0},$$

(9)

where $a_2$ is the amplitude of the sinusoidal profile’s short-wave component.

In 1996, Krystek compared the computational efficiency and transmission characteristics of spline profile filters and Gaussian profile filters, finding that the spline profile filter is faster than the Gaussian profile filter in extracting surface roughness. The spline profile filter can eliminate
the impacts of end effects and thus became a popular type of profile filter [28]. Krystek not only constructed cubic spline filter, but also introduced the numerical solution of the spline function matrix. This work laid a solid foundation for the application of the spline filter in the field of surface topography measurement. In 2007, Numada et al. proposed a high-order profile spline filter. With the increase of the order of the spline, the transmission characteristics become smoother. The experimental results showed that the end effect of high-order spline filter with seventh order or above was smaller than that of traditional low-pass filter. Through the frequency domain calculation method, the calculation efficiency was improved [29]. Conventionally, the end effect occurred in the ideal low-pass filter due to ringing artifacts. Because of the order limit of the high-order spline filter, the ideal low-pass filter with a step-edge-type cut-off characteristic was realized to provide a method to eliminate the end effect of an ideal filter. In 2011, Zeng et al. constructed a generalized spline filter model that can derive linear spline filters and nonlinear spline filters based on M-estimation theory. Experiments revealed that the spline profile filter derived from this model is suitable for extracting surface roughness and has high computational efficiency [30]. In 2012, Zhang et al. reduced the relative deviation of the spline filter and Gaussian filter transmission characteristics by using the cascade method for approximating the spline filter [31]. As shown in Fig. 3a, the transmission characteristics of the first-order approximating spline filter can approach the Gaussian filter well. The transmission characteristic deviation between the first-order approximating spline and Gaussian filter reaches the minimum value of 4.26% as shown in Fig. 3b. This is in line with the deviation range (~5.0% ~5.0%) of the simplified Gaussian filtering algorithm specified in the international standard ISO 11562. With the increase of the cascade order, the transmission characteristic deviation is further reduced. More importantly, the cascade approximation also provides a method to realize the isotropic areal spline filter. The ISO standard provides the tension parameter $\beta = 0.62524$ to make the transmission characteristic of the spline filter approximately similar to the Gaussian filter. However, when the tension parameter $\beta$ is not zero, end-effects appear. In 2015, Tong et al. proposed a new boundary condition that can provide satisfactory end portions of the output form without end-effects for the spline filter while maintaining $\beta = 0.62524$ [32]. In 2019, He et al. used a smooth B-spline filter to accurately extract the surface roughness of the S-shaped test part, and obtained the uncertainty model of surface roughness by using quadratic curve subsection fitting [33].

![Fig. 3. Transmission characteristics of approximating spline and Gaussian filter (a) and transmission characteristic deviation ($\lambda_c = 0.8$ mm) (b).](image-url)
The transmission characteristic of the spline profile filter is steeper than that of the Gaussian profile filter as shown in Fig. 4a, which makes the separation effect of the spline profile filter better. The spline filter can also effectively suppress the end effect when a modified weighting function is constructed. At the same time, they can smooth the data. As shown in Fig. 4b, the waviness curve of gear tooth profile filtered by spline filter and Gaussian filter. It can be found that the Gaussian filter has an end effect at the inclined boundary, while the spline filter follows the profile well. At the same time, the output line of spline filter is smoother. In the bearing industry, there are some narrow surfaces that have technical requirements for surface roughness. For example, while the reference end face of the tapered roller and the roughness \( Ra \) of the large rib of the tapered roller are required to be between 0.1–0.32 \( \mu m \), the effective length of the surface is 1–2 mm, which is not suitable for filtering with Gaussian profile filters with end effects due to the effective evaluation length. The sampling length of the spline profile filter can be reasonably selected, and is not limited by the length of the evaluated surface (the measured length is the evaluation length), so it is widely used in the bearing industry. The spline profile filter is also suitable for profile filtering of large forms, especially for roughness-measuring instruments without lead. A wide range of surface-measuring instrument analysis and evaluation software incorporates spline filters. However, the spline profile filter is less robust to outliers. This is primarily due to the fact that the spline filter is based on the least squares estimation and the variational method. When the error is non-normally distributed, the least squares estimation is no longer the most effective approach since the least squares estimate must follow the data of deep valleys and peaks in order to achieve the goal of minimizing the sum of squared residuals. These data (outliers) exert a much larger effect on the sum of the squares of the residuals than other data, greatly affecting the output signal of the filter, and thus causing the filter mean line to deviate from the true surface. For example, the surface of the crankshaft of an automobile engine and the surface of a laser hole both have obvious outliers.

\[ a \]

\[ b \]

Fig. 4. Transmission characteristics of the Gaussian and spline filters (a) and waviness curve generated by filtering gear tooth profile (\( \lambda_c = 0.8 \) mm) (b).

2.3. Wavelets

Wavelets are the small waves that are oscillatory and limited in the range. Wavelets are special type of sets of basis functions that are useful in the description of function spaces [34]. Wavelet transform is a new transformation method developed on the basis of short-time Fourier
transform [35]. It is a local transformation based on the time-frequency domain, which can effectively extract the mutation information from the changing signal, and achieve the multi-resolution analysis of the signal through the expansion and translation operation of the basis function [36, 37]. The profile $s$ is obtained by sampling at a fixed spacing ($x_i = i\Delta x$) (among them: $\Delta x$ is the sampling spacing, $i = 1, 2, \ldots, n$ and $n$ is the number of sampling points). The discrete wavelet transform with $g(x)$ as the parent wavelet is implemented,

$$S(i\Delta x, a) = \Delta x \sum_j s [(i - j)\Delta x] g_{a, j}(j\Delta x), \quad (10)$$

where $a$ is the dilation parameter for the wavelet of frequency band. $j = -m, \ldots, -2, -1, 0, 1, 2, \ldots, m$, and $m$ is the number of coefficients of the filter on one side from the centre.

Engineering surfaces are typically machined using a variety of processes, each producing a different frequency component. In actual engineering analysis, researchers are more prone to extract the frequency components of different scales in the signal, both in order to analyze the quality of the parts and to realize the monitoring process. Wavelet technology can decompose the profile signal into different frequency components, after which the corresponding frequency components can be examined using different scale resolutions. The basic process of wavelet analysis consists of decomposing the original signal into scale space, converting the time domain signal into a scale signal, and using the narrow and wide time-frequency windows to perform time-frequency analysis on the high and low frequencies, respectively. In 1998, S. Mall proposed the multiresolution analysis (MRA), also known as the multiscale analysis. It is a theory based on the concept of function space and its core idea is to use the multi-scale property of orthogonal wavelet basis to expand the signal so as to obtain the information of the signal in different scale space. As shown in Fig. 5, the multi-resolution structure divides the signal into low (a1) and high frequency components (b1). Then the low frequency components are further decomposed, and so on until the required layer number. The ultimate goal of decomposition is to construct an orthogonal wavelet basis that approximates the space in frequency. These orthogonal wavelet bases with different frequency resolutions correspond to bandpass filters with different bandwidths. MRA is one of quite often used methods of filtration and also being implemented in ISO 16610 series and it is the link between wavelet transform and filter theory.

![Fig. 5. Principle of multi-resolution analysis [38].](image)

In 1996, Sweldens proposed a lifting scheme and used it to construct interpolated spline wavelets [39]. The lifting scheme algorithm consists of three phases: splitting, predicting, and updating. In the splitting phase, the lifting scheme first divides the data to be analyzed into even and odd sequences, with each sequence containing half of the sample data. In the forecasting phase, the wavelet algorithm predicts odd sequences from even sequences and then removes the predicted values from the odd sequences in order to ensure that the average of the original signals remains unchanged. When predicting cubic interpolation spline wavelets, the following 5 cases need to be considered.
(1) There are 2 adjacent points on each side of the interval;
(2) There is 1 sampling point on the left boundary of 1 interval and 3 sampling points on the
right boundary;
(3) There is 1 sampling point on the right boundary of 1 interval and 3 sampling points on the
left boundary;
(4) All 4 sampling points are on the left;
(5) All 4 sampling points are on the right.

The main consideration for these situations is to better eliminate the end effect. In the updating
phase, the odd sequences are updated to even sequences in order to preserve as many profile
moments as possible. In 2015, the interpolated spline wavelet was officially incorporated into
the international standards. Among them, ISO 16610 Part 29 specifically recommends the cubic
interpolation spline wavelet as the wavelet filter for surface topography analysis [40].

At present, wavelets have first-, second-, and third-generation wavelet models; Table 4 com-
pares the properties of these three types of wavelet models. The basic biorthogonal wavelet
satisfies the three requirements for filtering the surface profile signal: linear phase, finite im-
pulse response, and complete reconstruction. Therefore, researchers often use basic biorthogonal
wavelets for wavelet analysis. For example, the bior6.8 biorthogonal wavelet not only has smooth
cut-off characteristics but also displays low-pass amplitude–frequency characteristics that are al-
most horizontal, with no overshoot. The basic biorthogonal wavelet belongs to the first generation
of wavelets. In 1996, the promotion of the lifting scheme led to the emergence of the second
generation of wavelets. The biorthogonal wavelets constructed with the lifting scheme (known as
second-generation wavelets) overcame the shortcomings of the first-generation wavelet computa-
tion while retaining the characteristics of the basic biorthogonal wavelet. The cubic interpolation
spline wavelet introduced by Part 29 also adopted the lifting scheme. However, the cubic in-
terpolation spline wavelet has a serious amplitude–frequency characteristic distortion problem,
and the obtained filter mean line seriously deviates from the surface profile. Therefore, the cubic
interpolation spline wavelet is suitable for surface topography analysis, but is not recommended
for extracting the surface profile mean line. The first- and second-generation wavelet models
were established by actual discrete wavelet transforms. This transformation caused the first- and
second-generation wavelets to effectively extract linear and curved scratches on the surface of
the microstructure, such as honing the deep valley features of the surface; characteristics of arc
scratches, regular scratches, random depth scratches, and so on caused by surface wear of arti-
ficial joint heads in bioengineering; and features such as micro-steps, grooves and channels on
the surfaces of micro-boards. The third-generation wavelet adopts complex wavelet transform
theory, which has translation invariance and directional selectivity, and can extract features, such

<table>
<thead>
<tr>
<th>Properties</th>
<th>Basic biorthogonal wavelet</th>
<th>Cubic interpolation spline wavelet</th>
<th>Complex wavelet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>Zero phase shift</td>
<td>Zero phase shift</td>
<td>Zero phase shift</td>
</tr>
<tr>
<td></td>
<td>Linear phase shift</td>
<td>Suppression of end effects</td>
<td>Fully refactored</td>
</tr>
<tr>
<td></td>
<td>Perfect reconstruction</td>
<td></td>
<td>Shift-invariant</td>
</tr>
<tr>
<td></td>
<td>Finite impulse response with tracking and positioning characteristics</td>
<td></td>
<td>Perfect reconstruction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Directional selectivity</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Low computational efficiency; end effects</td>
<td>Abnormal distortion of amplitude-frequency characteristics</td>
<td>At the nanoscale, robustness is poor</td>
</tr>
</tbody>
</table>
assuring linearity and curve scratches. This type of wavelet also exhibits good surface metrology properties and is suitable for surface roughness extraction.

In 1995, Chen et al. studied potential applications of wavelet transform in surface characterization, proving that wavelet analysis can link the multi-scale features of engineering surfaces with manufacturing and functional aspects [41]. In 1996, Liu et al. studied the difference between Daubechies wavelets and spline wavelets applied to engineering surface analysis. Compared with the spline wavelet filter, the Daubechies wavelet filter exhibits an improved amplitude frequency response and reconstruction accuracy [42]. In 2000, Jiang et al. proposed a lifting wavelet model and applied it to bioengineering, initially verifying the feasibility and applicability of a lifting wavelet model for surface topography analysis [43]. In 2003, Wang et al. proposed a surface roughness evaluation method based on wavelet analysis and fractal dimension that could also detect the self-similarity of a surface [44]. In 2008, Jiang et al. introduced basic biorthogonal wavelets, lifting wavelets and complex wavelets. Lifting wavelets are suitable for surface frequency analysis, surface recognition, and nanoscale precision reconstruction. Complex wavelets can better extract surface morphological features [45]. In 2009, Bakucz et al. proposed a wavelet filter based on fractional spline wavelet transform. Experiments revealed that the fractional spline wavelet filter can analyze the surface fractal scale characteristics well and has a good smoothing effect when filtering [46]. In 2015, Miller et al. investigated the influence of different filtering methods on surface roughness, and found that spline wavelet had caused little change to values of the roughness parameter compared with big values of morphological filters for the turning and milling surfaces [47]. Compared with other existing filtering techniques, the main advantages of wavelet transform are its space-frequency location and multi-scale view of signal components. In addition, it is a powerful tool for characterisation of local morphology by pinpointing the local motif using the modulus and the phase of the wavelet transform to quantify the roughness and waviness components.

Wavelet filters help investigators analyze the engineering surfaces formed by a variety of processes, as well as monitor these processes. In addition, wavelets can analyze non-static profiles. The spline wavelet filter can be used to divide a profile into different bands, and is an ideal multi-scale analysis and characterization filter. As shown in Fig. 6, the profile is decomposed into ten layers by coif4 wavelet, and the differential profile and approximate profile are drawn for multi-resolution analysis. Sintered, honed and similar surfaces are formed by multi-step processes and
have special processing requirements. The honing surface texture consists of rough honing, which results in deeper scratches, and fine honing, which removes sharp edges and leaves finer machine marks. This surface texture is formed by a combination of different wavelengths produced by the two processes. In fact, in order to suppress the influence of outliers on the roughness evaluation criterion, a double Gaussian filter or a robust filter is usually used to filter the honed surface of the platform. The spline wavelet filter can also perform surface topography analysis and surface roughness extraction on engineering surfaces such as turning, milling, grinding, rolling, and reaming of cylindrical surfaces and carbon surfaces. Compared to Gaussian filters and spline filters, spline wavelet filters are more conducive to analyzing the surface topography formed by multiple processes.

3. Robust profile filters

Under the combined effects of different processing techniques and material properties, a surface will produce microscopic geometric shapes of various shapes and sizes, such as deep valleys and peaks. These outliers do not follow a strict normal distribution. Outliers interference can cause distortion of the filtered reference line, which affects the authenticity of the filtered data. In order to suppress the influence of outliers on the filter reference, the international standard ISO 16610 specifies two robust filters: robust Gaussian regression filters and robust spline filters.

3.1. Robust Gaussian regression profile filters

Brinkmann and Seewing proposed a robust Gaussian regression filtering method [21, 48]. In 2010, the international standard added a robust Gaussian regression profile filter [49]. The robust Gaussian regression filter combines the robust weight function with the Gaussian regression filter to calculate the weight of each sampling point in an iterative way. The robust Gaussian regression filter is a regression filter based on a Gaussian weight function and a robust weight function. The filter equation is

\[ w = (1 \ 0 \ 0)(X_k^T \ S_k \ X_k)^{-1}X_k^T S_k z, \]

\[ X_k = \begin{bmatrix} 1 & x_{1,k} & x_{2,1,k} \\ \vdots & \vdots & \vdots \\ 1 & x_{n,k} & x_{2,n,k} \end{bmatrix}, \quad S_k = \begin{bmatrix} s_{1,k}\delta_1 & 0 & \cdots & 0 \\ 0 & s_{2,k}\delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{n,k}\delta_n \end{bmatrix}, \]

\[ x_{l,k} = (l - k)\Delta x, \quad l = 1, \ldots, n. \]

\[ s_{l,k} = \frac{1}{\gamma \lambda_c} \exp \left(-\pi \left(\frac{x_{l,k}}{\gamma \lambda_c}\right)^2\right), \quad k, l = 1, \ldots, n. \]

\[ \delta_l = \begin{cases} \left(-\pi \left(\frac{z - w}{c}\right)^2\right), & |z - w| \leq c \\ 0, & |z - w| > c \end{cases}, \]

\[ c \approx 4.4478 \times \text{median} |z - w|, \]
where \( w \) is the vector of dimension \( n \) of the profile values in the filtered profile, \( z \) is the vector of dimension \( n \) of the profile values before filtering, \( X_k \) is the matrix form of the regression function, \( S_k \) is the space variant weighting function, \( s_{l,k} \) is the Gaussian function, \( k \) is the index of the profile ordinates, \( \lambda_c \) is the cut-off wavelength, \( \Delta x \) is the sampling interval, \( \gamma \) is a constant (\( \gamma \approx 0.7309 \)), \( \delta \) is the additional weight, \( c \) is the scale parameter. The filter equation needs to be solved iteratively and by weighted least squares. Since at least \( n^2 \) operations are required for each iteration, the calculation efficiency is slow.

The robust weight function generally uses a multiple of the median or standard deviation as the scale parameter. The selection of the robust weight function is a very important step. Different weight functions may produce completely different filtering results when processing the same measured data. Therefore, the user needs to reasonably select the robust weight function based on the actual situation. The international standard uses the Tukey function in the M estimation method. In order to facilitate the user’s targeted selection of weight functions, Table 5 summarizes

<table>
<thead>
<tr>
<th>Estimation function</th>
<th>Additional scale</th>
<th>Characteristic</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLAV</td>
<td>MAD</td>
<td>Strongest resistance</td>
<td>The weighting factor gradually increases to infinity with decreasing error, which is contrary to the definition of correct observation.</td>
</tr>
<tr>
<td>Huber</td>
<td>( \sigma/MAD )</td>
<td>Balanced least squares efficiency and robustness of WLAV</td>
<td>The effect of outliers on estimates cannot be completely eliminated.</td>
</tr>
<tr>
<td>Hampel</td>
<td>MAD</td>
<td>The essence is a further fine-tuning of the Huber estimate</td>
<td>The slope of the function changes abruptly and has an adverse effect on the estimation of the data.</td>
</tr>
<tr>
<td>Tukey</td>
<td>MAD</td>
<td>Smooth truncation estimation with elimination points at both ends</td>
<td>Insensitive to small data estimates distributed in the middle and low computational efficiency.</td>
</tr>
<tr>
<td>Andrews</td>
<td>MAD</td>
<td>Smooth truncation estimation with elimination points at both ends; completes the operation with triangular waves</td>
<td>Insensitive to small data estimates distributed in the middle and low computational efficiency.</td>
</tr>
<tr>
<td>IGGI</td>
<td>( \sigma )</td>
<td>Synthesis of least squares estimation and WLAV estimation; high computational efficiency</td>
<td>Robustness is lower than the robust weight function using MAD as a scale parameter.</td>
</tr>
<tr>
<td>QC</td>
<td>( \sigma )</td>
<td>Good convergence; high computational efficiency</td>
<td>If there are more bad data or the value deviates from the normal value, it will lead to convergence problems or lower accuracy.</td>
</tr>
<tr>
<td>ADRF</td>
<td>MAD</td>
<td>Combines the computational efficiency of QC with the smoothness of the Tukey, achieving both robustness and convergence speed</td>
<td>Distinctive distortion on the surface of large-width deep valleys.</td>
</tr>
</tbody>
</table>
the characteristics of several typical robust weight functions. MAD = \text{med} |r|$, med represents the median of roughness data, and $\sigma$ represents standard deviation, which are both used as scale parameters.

In 2004, Li et al. proposed a new robust weight function, ADRF. Experimental results revealed that robust filtering based on the ADRF weight function not only improves the robustness of the filter, but also preserves the accuracy of the Gaussian filter [50]. The robust Gaussian filter composed of the ADRF function and the Tukey function is used to process the honing surface, and it is found that the corrugation curves of the two surfaces are basically the same, as shown in Fig. 7. However, there are significant differences in the number of iterations, 4 and 8 respectively. In general, the ADRF function has good filtering performance. Dobrzanski et al. compared the influence of the traditional Gaussian filter, Gaussian regression filter and spline filter on the variations of surface roughness profile parameters. The experimental results showed that the roughness parameter variations caused by the three kinds of filters for the profile filtering with longer wavelength were quite different. However, the calculation error of the roughness parameters after the spline filtering was large, and the calculation time was longer than that for the Gaussian regression filter [51]. In addition, Dobrzanski et al. used the Andrews, ADRF, Huber, Tukey, and Hampel weight functions to establish robust Gaussian profile filtering techniques, and filtered the random profiles generated by computer simulation, with triangular scratches, and with stratified surfaces. The degree of profile distortion and calculation time after filtering were then compared. Ultimately, the Tukey function of scale parameter 4.4, Hampel (1:1:5:3), and Andrews (1:5) were recommended [52, 53]. In 2014, Gurau et al. studied the convergence of robust Gaussian regression filters and applied them to wood surfaces of different grain sizes in order to determine tolerance values, thereby minimizing the number of iterations. Gurau et al. examined the convergence of the RGRF by monitoring the effect of some predefined tolerances (0.1 and 0.01 \( \mu \text{m} \)) on the variation of roughness parameters of wood profiles from various sanded surfaces, discovering that the roughness parameters of wood surfaces converge to a tolerance of 0.01 \( \mu \text{m} \) [54].

In surface roughness measurement, the Gaussian profile filter is suitable for filtered surfaces, and the robust Gaussian regression profile filter is also generally applicable. Meanwhile some engineering surfaces contain functionally related structural elements, such as laser holes and hard
particles left by laser processing of metal-based composite materials, deep valleys left by honing processing, pits left by EDM, and so on. For these surfaces, it is recommended not to use Gaussian profile filters, instead utilizing robust Gaussian regression profile filters, which are more suitable and yield more authentic profile data. The robust Gaussian regression profile filter is also suitable for filtering highly layered surfaces and wood surfaces. However, robust Gaussian regression filters exhibit neutrality on surfaces whose surface roughness follows a normal distribution. In addition, the surface roughness error calculated by the Gaussian regression filter is less than that of the spline filter, and the calculation efficiency is higher. Finally, the number of iterations of the robust filter should be noted. The more iterations, the stronger the suppression of outliers. The number of iterations depends on the filter threshold. As shown in Fig. 8, the robust Gaussian regression filter is used to filter the honing surface for 2 and 5 iterations respectively. It can be found that 2 iterations cannot well restrain the influence of large deep valleys. On the contrary, a satisfactory mean line could be obtained by 5 iterations. Therefore, the selection of threshold value needs to be determined by repeated experiments according to the surface morphology.

![Graph showing honing profile and waviness profiles](image)

**Fig. 8.** Honing profile and waviness profiles are generated by using the robust zero-order Gaussian regression filter with \( \lambda_c = 0.8 \) mm (iterations: 2, 5).

### 3.2. Robust spline profile filters

The spline profile filter solves the end effect problem. However, its robustness to outliers needs to be improved. Therefore, ISO 16610 Part 32 recommends a robust spline profile filter [55]. Based on the concept of the variational method, the robust spline filter constructs a compromise function between the sum of the absolute residual values and the approximate bending energy of the data. This function is solved by a loop iterative operation. The filter equation can be composed of any order. Part 32 introduces a robust spline profile filter based on cubic splines. Equation (12) is the filter equation for the open profile, however, the matrices \( P \) and \( Q \) become the corresponding augmented matrices when the closed profile is filtered.

\[
\begin{bmatrix}
\beta a^2 P + (1 - \beta) a^4 Q
\end{bmatrix} \cdot w = \frac{\text{sgn}(z - w) \sum |(z - w)|}{\sum |(z - w)|},
\]

\[
\text{sgn}(t) = \begin{cases} 
+1, & t \geq 0 \\
-1, & t < 0
\end{cases}
\]
In 2005, Goto et al. proposed a robust algorithm based on the L-2 norm. This new robust filter inherits all the advantages of ordinary spline filters and exhibits good robustness to outliers [56]. In 2005, Krystek discussed the general characteristics of spline profile filters and robust spline profile filters, and experimentally proved that the health sample strip filter is more advantageous for extracting surface roughness [57]. In 2013, Zhang et al. proposed a new robust spline profile filter. Compared with the robust spline filter specified in the international standards, the new filter is more robust and more computationally efficient. In addition, experiments revealed that this new robust spline filter has the same filtering characteristics as the Gaussian regression filter. Therefore, it can be used for the comparison of international standards and surface assessment results [58]. In 2018, Nakamura et al. used robust spline filter to extract the roughness of coated wood surface and further verified the feasibility of evaluating the sensory roughness of coated wood surface through image analysis [59].

The robust spline profile filter is an improvement on the spline profile filter, and its filtering effect is better than that of the spline profile filter. Due to its robustness, the robust spline profile filter can filter porous surfaces well, eliminating the effects of outliers on the filtering mean line. This filter is also available used in the automotive industry, for example, the mean line of the measured profile on a metal plate (which usually has a large-form deviation) and the extraction of the profile roughness of an engine cylinder liner. Similar to a robust Gaussian regression filter, the robust spline filter also has strong robustness. However, the construction principle and implementation algorithm of the robust spline filter are quite different from that of the robust Gaussian regression filter, which leads to large differences in surface filtering. In order to more accurately assess the surface roughness, it is necessary to study new application algorithms to narrow the gap in filter characteristic differences between the two filters.

The theoretical and technological developments of the above five types of profile filters are relatively mature. In 2015, Markov et al. briefly discussed the robust filtering algorithm for surface roughness profiles and gave an example of the algorithm applied to robust Gaussian regression and robust spline [60]. However, the relevant documentation does not provide a detailed summary of their differences in surface roughness applications. Table 6 summarizes the advantages and disadvantages of each filter, along with their typical applications.

Table 6. Advantages, disadvantages, and applications of profile filters.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Typical applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian filter</td>
<td>Phase linearity Easy to calculate Can be used for interval data</td>
<td>End effects Not robust Sensitive to outliers Need to remove the shape before filtering</td>
<td>Grinding, stamping, turning, milling, reaming, granite, and others</td>
</tr>
<tr>
<td>Spline filter</td>
<td>Phase linearity Can eliminate end effects Can filter profiles large-form Not subject to assessed length limit Can be used for random spacing data Acquired filtered mean line smoothing</td>
<td>Poor robustness Roughness parameter calculation error is large [61]</td>
<td>Ball bearing inner ring surface, tapered roller reference end face and inner sleeve large rib</td>
</tr>
<tr>
<td>Spline wavelet</td>
<td>Multi-scale analysis Extracts a single feature Locates and identifies outliers Closed profile: no end effects Suitable for unstable surfaces Can be applied to short profiles Acquired filtered mean line smoothing No need to remove the form before filtering</td>
<td>Multiple types of mother wavelets</td>
<td>Turning, milling, grinding, rolling, reaming cylindrical surfaces and carbon surfaces [42,43]</td>
</tr>
</tbody>
</table>
Table 6 (cont.)

<table>
<thead>
<tr>
<th>Filter</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Typical applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust Gaussian</td>
<td>Robust against outliers</td>
<td>Nonlinear filter; and cannot be described by filter transfer function</td>
<td>Laser holes or hard particles, plateau honing surfaces, and spark pits in metal matrix composites [48]</td>
</tr>
<tr>
<td>regression</td>
<td>Can eliminate end effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust spline filter</td>
<td>Robust against outliers</td>
<td>Nonlinear filter; and cannot be described by filter transfer function</td>
<td>Porous surfaces, plateau honing [62], such as engine cylinder liner surface</td>
</tr>
<tr>
<td></td>
<td>Can eliminate end effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Can be used for random spacing data</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Morphological profile filters

Digital morphological filters are a new class of nonlinear filters developed from mathematical morphology. These range from the early binary morphological filters to the later multi-value morphological filters. In terms of signal processing, the morphological filter achieves signal filtering by selecting a smaller structural element to interact with the digital signal. Introducing morphological filters overcomes some of the limitations of linear filters in processing signals, such as end effects and lack of robustness.

International standards define two structural elements: the disk and the horizontal line-segment [63]. The discrete morphological filter takes profile $z$ and structural element $s$ as inputs and outputs profile $w$ with the same array length as profile $z$. A basic method for calculating discrete morphological filters is provided in the standard. The structural elements are positioned at each point of the input profile, and the extreme points at each sampling point are collected to form an envelope of the output. Dilation and erosion are two basic morphological operations, and opening and closing operations are two secondary morphological operations. The morphological filter includes a morphological closed filter (opening operation) and a morphological open filter (closing operation). The former process is dilation followed by erosion, and the latter is erosion followed by dilation. In general, the closing operation forms an output envelope above the input profile and the opening operation forms an output envelope below it. As shown in Fig. 9, MATLAB

![Fig. 9. Dilation, erosion, opening and closing mean lines with a 0.5 mm ball radius.](image-url)
is used to generate a profile as the sum of the following three components defined in the interval $0 \leq x < 8 \text{ mm}$, with a spacing $\Delta x = 1 \mu\text{m}$: (a) sinusoidal profile of amplitude $0.2 \mu\text{m}$ and wavelength $16 \text{ mm}$, (b) sinusoidal profile of amplitude $0.05 \mu\text{m}$ and wavelength $0.1 \text{ mm}$ and (c) superimpose normally distributed random noise of zero mean and $0.03 \mu\text{m}$ standard deviation. Then, the structure element of size $0.5 \text{ mm}$ was used to output the dilation mean line, and the erosion mean line was a simple dilation mean line calculated on the inverse profile. After dilation and erosion, erosion and dilation output the closing and opening mean line respectively.

Morphological filters can be used for the analysis of outliers such as deep valleys and peaks. For these outliers, morphological closed filters can eliminate deep valleys with widths smaller than their own scale, and morphological open filters can eliminate peaks with widths smaller than their own scale. The principle is illustrated in Fig. 10. SE1 and SE2 are horizontal line segment structural elements of different lengths (the length of SE1 is slightly shorter than that of SE2); SE1 and SE2 are used for closed and open filters, respectively. The left sides of (a) and (b) show the use of the morphological closed filter, and the right sides use the morphological open filter. The peak on the left side of the figure is not eliminated, while the deep valley smaller than the structural element SE1 is eliminated. In contrast, the peak on the right side of the figure that is smaller than the construction element SE2 is eliminated. It can be seen that the morphological filter acts as a “sieve”. In the closed filter, all peaks and deep valleys with structural widths greater than the length of the structural elements are retained in the sieve. In the open filter, all deep valleys and peaks with structural widths greater than the length of the structural elements are retained in the sieve. An alternating-sequence filter is a combination of closed and open filters with different-sized structural elements. Deep valleys and peaks can be removed simultaneously when using alternating-sequence filters. If the structural elements of the closed and open filters are the same size, and the structural elements used are symmetric about their origin, the alternating-sequence filter is called an alternating symmetric filter. The structure of the alternating-symmetric filter is similar to multi-resolution analysis, allowing continuous observation of profiles at different scales. The profile signal is divided into low-frequency and high-frequency components at each layer, and the low-frequency components then continue to be decomposed on a larger scale [64]. Table 7 summarizes the advantages and disadvantages of the morphological open (closed) filters as well as the alternating-symmetric filter.

![Filtering diagram of simulated peaks (a) and deep valleys (b) by morphologically open/closed filters with structural elements of different sizes.](image)

Fig. 10. Filtering diagram of simulated peaks (a) and deep valleys (b) by morphologically open/closed filters with structural elements of different sizes.

In 1974, Shunmugam et al. discussed the process of calculating two-dimensional and three-dimensional envelopes, comparing the difference between the roughness value measured by a sphere and that measured by a circle. In addition, they found that the radii of the sphere and the rolling circle played a major role in surface roughness measurement [65]. In 1998, Srinivasan provided a MATLAB algorithm for implementing discrete morphological filters, proving that the morphological filter is computable [66]. In 2006, Kumar et al. used two circles of different
Table 7. Advantages and disadvantages of morphological open (closed) filters and alternating symmetric filters.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Morphological open (closed) filter</th>
<th>Alternating symmetric filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>Closed profile: no end effects</td>
<td>Robust against outliers</td>
</tr>
<tr>
<td></td>
<td>Suitable for random spacing data</td>
<td>Can eliminate end effects</td>
</tr>
<tr>
<td></td>
<td>Shift-invariant and monotonic decreasing</td>
<td>Can perform multi-resolution-type analysis</td>
</tr>
<tr>
<td></td>
<td>No need to remove the shape before filtering</td>
<td>No need to remove the shape before filtering</td>
</tr>
<tr>
<td>Disadvantage</td>
<td>Slow calculation</td>
<td>Slow calculation</td>
</tr>
<tr>
<td></td>
<td>Certain degree of end effects</td>
<td>Uniform sampling profiles only</td>
</tr>
<tr>
<td></td>
<td>Uniform sampling profiles only</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open (closed) filter only has partial robustness</td>
<td></td>
</tr>
</tbody>
</table>

diameters as structural elements, with small circles and large circles used successively to operate on profiles. Experiments revealed that this morphological filtering method is robust to deep valleys and has good shape-tracking ability [67]. In 2011, Lou et al. proposed an alpha-shape algorithm. The shape profile filter and the area filter formed by this algorithm have better filtering effects, and experiments demonstrated that the alpha-shape algorithm is more efficient than the naive algorithm [68]. In addition, Lou et al. proposed two algorithms for implementing morphological profile filters, an improved Graham scanning algorithm and a recursive algorithm [69]. In 2013, Lou et al. applied the alpha-shape algorithm to the inner surface of automotive engine cylinder liners, wear metal surfaces, new ceramic surfaces, and a range of bioengineered surfaces. The experimental results demonstrated that the alpha-shape algorithm can successfully extract surface topography features [70].

In the surface roughness measurement, morphological closed filters and open filters can eliminate topographical deep valleys and peaks, respectively. However, the individual use of the open and closed operations produces a reference profile with a slanted discontinuity, and as a result, the contour after filtering is distorted, which affects the authenticity of surface roughness measurement data. Combining the form-closing and opening operations produces a smooth reference profile. The smooth reference profile more accurately reflects the waviness of the bottom layer, which is more conducive to the evaluation of surface roughness. Alternating symmetric morphological filters can be applied to engineering surfaces such as milling, ceramics, and honing, in which surface texture defects and topographical features can be well identified on the basis of scale space technology. However, traditional morphological filtering algorithms have not been widely used, primarily because: (1) For larger structural elements, all surface data can be involved in the calculation of each envelope ordinate, resulting in a calculation time that is too long. Due to the large amount of calculation, the maximum size of the structural elements is also limited. (2) In the implementation process of the traditional algorithm, there is a modified end effect after infinite filling which will cause extraction distortion of surface topography features.

5. Areal filters

The acquisition of the two-dimensional surface roughness parameters depends on the direction and length of the sampling profile segment, and only reflects the geometrical features of the single profile of the surface of the material being tested. The surface topography of the workpiece obtained in actual production is often anisotropic, so the two-dimensional surface roughness parameters have great limitations and can be misleading when characterizing the surface topography of the workpiece to be tested. The three-dimensional surface roughness parameters can provide and reflect the spatial topography of the surface of the material and can comprehensively reflect the surface topography.
5.1. Gaussian areal filters

ISO 16610 Part 61 divides Gaussian areal filters into Gaussian planes and Gaussian cylinders which are applied to nominal plane and cylindrical plane filtering respectively [71]. The weight function of the Gaussian plane filter is

\[
s(x, y) = \begin{cases} \frac{1}{\alpha^2 \lambda_c^2} \exp \left[ -\frac{\pi}{\alpha^2} \left( \frac{x^2 + y^2}{\lambda_c^2} \right) \right], & -L_c \lambda_c \leq \sqrt{x^2 + y^2} \leq L_c \lambda_c \\ 0, & \text{otherwise} \end{cases}
\]

(14)

where \(x, y\) is the distance from the center (maximum) of the weighting function in the \(x\) and \(y\) directions, respectively, and \(L_c\) is the truncation index of the Gaussian filter. \(\alpha\) is the constant, to provide 50% transmission characteristic at the cut-off \(\lambda_c\). The Gaussian weight function has non-compact support, meaning that the weight function has no boundaries in the \(x\) and \(y\) directions. In order to determine the moving average, the weight function other than the interval \((-L_c \lambda_c, L_c \lambda_c)\) must be set to 0. \(L_c \lambda_c\) can control the surface area that can be evaluated. In order to prevent excessive reduction of the measured area, \(L_c \geq 0.5\) is generally selected without significantly distorting the filter transfer function. The weight function of the Gaussian cylindrical filter is determined by the cut-off wavelength in the axial and circumferential directions, and is expressed as

\[
s(t|f_c) = \begin{cases} \frac{f_c}{\alpha L} \exp \left[ -\pi \left( \frac{tf_c}{\alpha L} \right)^2 \right], & -\frac{L_{ct}}{f_c} \leq t \leq \frac{L_{ct} L}{f_c} \\ 0, & \text{otherwise} \end{cases}
\]

(15)

where \(t\) is the distance of the weight function from the circumferential direction to the center (maximum value), \(f_c\) is the cut-off frequency in undulations per revolution, \(L\) is the length of the closed profile, and \(L_{ct}\) is the truncation index of the Gaussian filter. The Gaussian weight function is separable, so the Gaussian areal filter can be regarded as a Gaussian profile filter superposed in two directions.

The areal Gaussian filter also exhibits end effects, as shown in Fig. 11. In 2013, Seewig summarized two methods for correcting the end effects of Gaussian areal filters [38]. One is to reflect the measurement data in the boundary area in order to increase the measurement area; the

![Fig. 11. Sinusoidal surface (a) and waviness surface are generated by using the Gaussian filter (\(\lambda_c = 0.8\) mm) (b).](image-url)
measurements can then be filtered all the way to the boundary area. The other technique uses the asymmetric weight function of the truncated boundary area, with the volume being corrected by rescaling it to the value of 1.

Although the Gaussian filter is the preferred standardized filter, it is theoretically a non-causal system that is physically unrealizable. However, the theoretical amplitude-frequency and phase-frequency characteristics of the Gaussian filter can be approximated by either an approximating polynomial or an approximation function. Table 8 summarizes three approximation methods for implementing digital Gaussian filtering.

Table 8. Approximation methods for digital Gaussian filtering.

<table>
<thead>
<tr>
<th>Approximation method</th>
<th>Characteristic</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIR type digital Gaussian approximation filter</td>
<td>Structural feedback</td>
<td>There is an end effect, the length of the end effect interval is fixed, and the data will converge stably after a certain amount of iterative calculations</td>
</tr>
<tr>
<td></td>
<td>Small amount of calculations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zero-phase-shifting characteristics</td>
<td></td>
</tr>
<tr>
<td>FIR type digital Gaussian approximation filter</td>
<td>Small amount of calculations</td>
<td>As the number of approximations increases, the degree of end effects increases</td>
</tr>
<tr>
<td></td>
<td>Unit impulse response is limited</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Essence is a rectangular window</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Structural feedback does not exist</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zero-phase-shifting characteristics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phase linearity is stricter than IIR type</td>
<td></td>
</tr>
<tr>
<td>Triangular window approximation</td>
<td>Simple algorithm and high computational efficiency</td>
<td>End effect</td>
</tr>
</tbody>
</table>

In 1993, Luo et al. investigated the application of the Gaussian filter in three-dimensional surface topography in detail, and pointed out its applicability in surface measurement. At the same time, they also noted that Gaussian filters usually have poor frequency selectivity [72]. In 2001, Yuan et al. proposed a Gaussian filter recursive algorithm for areal surface topography measurement, which uses zero-phase filtering technology to render the Gaussian filtered surface free of phase distortion [73]. In 2005, Xu et al. proposed a Gaussian filter with a cascaded moving average filter. The larger the number of stages, the higher the filtering accuracy. The amplitude deviation of the 8-stage cascaded filtering algorithm and the 2D Gaussian filter is less than 1%, which proves the feasibility of the filtering algorithm [74]. In 2019, Gathimba et al. used a Gaussian filter to investigate the influence of corrosion degree and thickness change of steel pile on surface roughness characteristics in marine environment [75]. In 2019, Solhjoo et al. studied the idea of using Gaussian filters to eliminate the effect of surface roughness on the contact stress field of atoms [76]. In 2020, Todhunter et al. introduced a novel method for the validation of surface texture filtration. The method utilizes a combination of mathematically defined surfaces and an accurate evaluation of the linear Gaussian filter transmission characteristic function to produce a mathematically traceable reference pair suitable for the performance assessment for surface texture analysis software [77].

Unlike linear profile filters, Gaussian areal filters must establish a reference plane (i.e., the mean plane) when extracting surface roughness. Compared to the profile mean line, the reference surface contains more roughness information, which is more helpful when assessing the surface topography. Gaussian areal filters can be used on most engineered surfaces such as grinding, milling, stone, and carbon fiber. It does not, however, apply to surfaces that contain pits, valleys, scratches, and so on, or to surfaces that have almost no frequency-period characteristic signals.
such as wood, ceramics, and honed surfaces. The turning surface and the plastic surface are filtered by Gaussian filter as shown in Fig. 12 and Fig. 13 respectively. The morphology of a turning surface is close to normal distribution, whereas the morphology of a plastic surface can be regarded as normal random distribution. There is no distortion for the reference plane obtained by the Gaussian filter and the accurate roughness information is retained.

![Fig. 12. Turning surface (a) and waviness surface are generated by using the Gaussian filter ($\lambda_c = 0.25$ mm) (b).](image)

![Fig. 13. Plastic surface (a) and waviness surface are generated by using the Gaussian filter ($\lambda_c = 0.25$ mm) (b).](image)

5.2. Robust Gaussian regression areal filters

The introduction of a robust Gaussian regression filter allows for better filtering of surfaces containing functional properties such as honing, laser texturing, or metal matrix composites. With the use of the robust filter, the filter plane mainly follows the plateau, providing a clear basis for evaluating functionally related surface structures. The robust Gaussian regression areal filter is a regression filter based on the Gaussian weight function and the influence function [78]. The filter equation for a robust Gaussian plane filter is

$$w_{ij} = (1 0 0 0 0) \left(X_{ij}^T S_{ij} X_{ij}\right)^{-1} X_{ij}^T S_{ij} z_{ij},$$  \hspace{1cm} (16)
where \( w_{ij} \) is the filtered surface value, \( z_{ij} \) is the surface values before filtering, \( X_{ij} \) is a regression function, and \( S_{ij} \) is a spatial variable weight function, both of which are represented in the form of a matrix. When filtering a radial surface such as a cylindrical surface, \( X_{ij} \) and \( S_{ij} \) are replaced by corresponding augmented matrices, and \( z \) is replaced by radial surface value \( \rho \). The above matrix equation cannot be solved directly, requiring the use of the vertical robust estimation weight function \( \delta \) and an iterative algorithm.

In 2001, Brinkmann et al. studied the difference between the surface filtering of different-shape components with a zero-order Gaussian regression filter and a second-order Gaussian regression filter. They discovered through experiments that while the zero-order Gaussian regression filter can better deal with end effects, the second-order Gaussian regression filter has a better approximation ability. Brinkmann also proposed a robust Gaussian regression algorithm for hierarchical surfaces with asymmetric distribution functions [21]. In 2004, Fujiwara et al. studied the characterization of the surface roughness of wood. They found through experimentation that the robust Gaussian regression filter greatly reduced the influence of deep valleys on the surface roughness of wood [79]. In 2010, Zeng et al. studied the general model of the Gaussian regression region filter and proposed an algorithm for analyzing the second-order Gaussian regression filter of three-dimensional surfaces. The practicality and feasibility of the second-order Gaussian regression filter and the robust Gaussian regression filter from the model were experimentally proven [80]. In 2019, Podulka proposed a valley excluding method (VDEM) to remove the areal form and found that the dimples distortion on the surface of plateau-honed cylinder liner could be minimized with the application of VDEM besides using Gaussian regression, robust Gaussian regression and other filters [81].

For any engineering areal surface roughness measurement, the robust Gaussian regression areal filter is an effective tool for determining the reference surface. A stone surface representation obtained by the Talysurf CCI measurement system is shown in Fig. 14a. The existing curved shape has been removed from the surface, but the surface details are retained. In order to test the robust performance of the Gaussian regression filter, five square deep valley areas are intentionally added to the figure to increase the surface deformity characteristics. Fig. 14b shows waviness surfaces which were obtained using a first-order robust Gaussian regression filter with a nested exponent of 0.8 mm. According to the output results, there is no obvious distortion in and around the square area. The low frequency waviness is extracted effectively by robust treatment and the influence of

![Fig. 14. Stone surface (a) and waviness surface using a first-order robust Gaussian regression filter (\( \lambda_c = 0.8 \text{ mm} \)) (b).](image)
abnormal characteristics is eliminated. A robust Gaussian regression filter can effectively track a curved surface. Its algorithm was specifically designed to analyze high-strain contact surfaces in mechanical engineering, such as honed cylinder liners and sintered bearings. Robustness allows the filter-to-filter surfaces with obvious peaks and deep valleys without distorting the filter reference plane. The robust filter does not work without an offset, so the effect of calculating amplitude parameters can be ignored. There is no difference between the robust filter and the non-robust filter when correcting the zero offset of the reference face value. Therefore, a robust Gaussian regression areal filter can be applied to surfaces that do and do not require robust processing. Brinkmann et al. applied robust Gaussian regression region filters and non-stable filters to turning, milling, ground, and honing surfaces. They discovered that there is almost no difference between the two filters in turning, milling, and ground surfaces, but for filtering honing surfaces with deep valleys, the robust Gaussian regression areal filter provides more realistic “flat” data [8]. In order to facilitate the practical application of robust Gaussian regression areal filters, ISO 16610 Part 71 only specifies definitions in discrete cases.

5.3. Morphological areal filters

Areal segmentation divides surface topography into different regions. In particular, morphological segmentation has received extensive attention in surface metrology. In 2012, the international standard ISO 25178 Part 2 first defined morphological segmentation, which was partitioned into Maxwellian hills and dales [82]. In 2013, the international standard ISO 16610 Part 85, as revised by the Technical Committee, replaced the segmentation in ISO 25178 Part 2 and defined segmentation as a filtering technique [83]. Watershed segmentation (Maxwellian dales) is the main morphological segmentation method. The basic idea is to immerse a topography model vertically in water, then prick holes at the lowest point of each dale, so that the water slowly and evenly rises in each hole. When the water levels of the different dales rise to the point where they merge, a dam is constructed at the junction of the two dales. As the water level gradually rises, all the dales are completely flooded. Each valley is completely surrounded by dykes, so that each dam (watershed) and dale (target object) are separated by dams, finally achieving the goal of terrain segmentation. Watershed segmentation itself is very robust and can cope with complex topographical surfaces. It is also fully deterministic, that is, the same result is guaranteed from the same topographical data. However, segmentation also creates insignificant dales that mask more important features. Also, Wolf pruning can be applied which is a method based on measuring the height difference between the peak/pit and the saddle point in order to apply a threshold, thereby eliminating unimportant hills/dales and thus simplifying redundant areal features. The use of Wolf pruning also faces a problem, that is, appropriate thresholds need to be selected in order to determine which hills/dales to merge. The threshold is typically determined by the percentage of the given areal parameter $S_z$ for which the typical threshold for Wolf pruning is 5% $S_z$ [82], although this value is not optimal for all situations.

Surface topography features play an important role in specific functions. For example, larger peaks and hills function as contact bands in engineering applications. The segmentation technique based on Wolf pruning can effectively identify and extract the important features of a surface, and then calculate the feature parameters in order to quantify the characteristics of the selected features. This also indicates that morphological segmentation techniques have significant potential for capturing important surface features. In 2011, Wang et al. applied segmentation and Wolf pruning pattern recognition technology to extract the characteristic parameters of a titanium surface. This new feature recognition technology is highly efficient and stable, and has played a leading role in establishing new surface characterization techniques [84]. Senin et al. examined
the differences between Maxwellian hills and dales segmentation and Maxwellian hills and dales segmentation with edge detection in practical application, finding that Maxwellian hills and dales segmentation can effectively extract the salient traits of a surface, but is less suitable for extracting individual features. Dale segmentation combined with edge detection is particularly useful for areal features that are high-slope feature boundaries, or any other type of feature boundary that allows for successful edge detection [85].

Micro-electromechanical system surfaces typically contain regular stepped features and planar portions, defined as structured surfaces. The segmentation technique based on Wolf pruning can reduce excessive segmentation caused by local roughness, thereby effectively extracting the salient features of the surface. The combination of edge detection and the watershed segmentation technique can accurately identify feature boundaries, thus solving the feature boundary blur problem caused by areal data preprocessing. For example, deeper and slender scratches usually appear on the worn hip-head surface and Wolf pruning can merge the unimportant dales at the bottom of the scratches. Wolf pruning, however, may also merge adjacent scratches, resulting in inaccurate recognition of scratch boundaries. The combination of edge detection and watershed segmentation techniques can cope with this problem. The segmentation technique based on Wolf pruning can correctly calculate the number of grinding wheel particles, and can distinguish the active peaks in the segmentation area, thus helping to characterize the grinding wheel. In addition, in the automotive industry, it effectively identifies individual moats in the car body panel. Wrinkle surface of Hills defined from Wolf pruning using 10% of $S_z$ is shown in Fig. 15. The “smaller” segments are pruned out to leave a suitable segmentation of the scale-limited surface. It should be noted that, unlike filtering techniques such as Gaussian, spline or wavelet, morphological areal filtering is supposed to isolate relevant topographic traits encoded by the main hills and dales, whilst discarding smaller, less relevant, topographic formations. Once the decomposition is complete, geometric attributes of the salient topographic formations (height, projected area, perimeter and other properties of each hill and/or dale) are aggregated back into synthetic descriptors of the overall properties of the surface texture, i.e., feature parameters (a class of areal roughness parameters).

![Fig. 15. Wrinkle surface (a) and hills defined from Wolf pruning using 10% of $S_z$ (b).](image)

In areal surface analysis, morphological filters can be used not only for segmentation, but also for opening, closing, erosion and dilation filters. At present, the international standard plans to
include sphere, horizontal planar segment filter and scale space technology into the morphological areal filter (ISO 16610-81, ISO 16610-89). For instance, abrasive surface illustrated in Fig. 16a, a sphere with a diameter of 0.2 mm is used as the construction element to perform the closing and opening operations. Fig. 16b shows the filtered surface and Fig. 16c shows the residual surface. Although the international standard has not officially published the corresponding areal filtering file, the feasibility of the filtering of opening, closing, erosion and dilation filters has been verified and used in the software of surface topography analysis, such as MountainsLab Premium.

![Fig. 16. Abrasive surface by Talysurf CCI (a) and filtered surface (b) and residual surface (c).](image)

**5.4. International standards are planning to include areal filters**

WG 15 of ISO TC 213 is currently working on proposals for areal filters such as spline filters, robust spline filters, and spline wavelets. Although these areal filters have not been officially included in international standards, researchers and scientists have never stopped studying their related theories and engineering applications.

When the spline filter is used as the areal filter, the spline profile filter is generally used to superimpose the spline areal filter in the $x$ and $y$ directions. Unlike the Gaussian filter, the spline areal filter constructed by this method has anisotropic amplitude characteristics, which limits its application in 3D surface topography. In 2009, Goto et al. proposed a discrete optimal operator design method based on the least squares method. The discrete spline region filter resulting from this method displayed the best isotropic amplitude characteristics when $\beta \approx 0.056$ [86]. In 2013, Janecki et al. constructed a two-dimensional isotropic spline filter by eliminating the discretization of the part of the function representing the bending energy and treating the filtered surface as a continuous function of spatial variables [87]. In 2015, Zhang et al. constructed a spline areal filter with isotropic amplitude characteristics and no end effects based on the high-order spline areal filter [88]. The introduction of weighting factors and new boundary conditions in the isotropic spline areal filter makes it possible to evaluate surfaces containing blank areas and highly inclined shapes, such as strike needle impression surfaces [89]. The study of robust spline filters mainly focuses on profile surface filtering, and less on the areal surface. Goto et al. designed a robust spline filter that can be used for both planar and spatial curves using the filter equation vector expression of the spline filter [56].

Tuning fork surface (6.12 mm × 2.44 mm) measured by Talysurf CCI is shown in Fig. 17a. Fig. 17b shows the waviness surface obtained by Gaussian filter with $\lambda_c = 0.8$ mm. The waviness surface shown in Fig. 17c is obtained by decomposing a spline wavelet with scaling level 5. It can be found that the datum obtained by the two filters are similar, which indicates the feasibility of roughness filtering with a spline wavelet. International standards are also planning to incorporate sample wavelets into linear areal filter standards (ISO 16610-69). Wavelet transform provides
a new filtering technique for surface topography evaluation and analysis, and has developed into a “3-generation” wavelet model. The earliest application of the first generation of biorthogonal wavelets is also the most extensive. The cubic interpolation spline wavelet recommended by the international standard is suitable for surface topography analysis. The third generation of wavelets inherited the advantages of the first two generations and can be successfully used for surface topography analysis and surface roughness extraction. With the increasing demand for 3D surface filtering, many researchers have begun to study the application of wavelet transform on areal surfaces. In 2001, Josso et al. used wavelet transform to identify eight different engineering surface roughness (casting, grinding, grit blasting, hand filing, horizontal milling, linishing, shot blasting, and milling) [90]. Also in 2001, Xiao et al. used the biorthogonal B-spline wavelet and cubic spline wavelet to extract the surface roughness and texture characteristics of the lapped ceramic femoral head surface. Experiments have shown that the two spline filters maintain similar properties in terms of linear phase shift, smoothness, and second-order moment, although the boundary problem of the biorthogonal B-spline wavelet still needs to be completely solved [91]. In 2001, Jiang et al. used a lifting wavelet to extract the morphological features of the scratches and grooves of the new metal femoral head bearing surface. Experimentation proved the feasibility and applicability of the lifting wavelet in the extraction of surface topography features [92]. In 2005, Zeng et al. proposed a dual-tree complex wavelet transform for surface analysis, which was used for 3D milling surfaces. Their results revealed that the dual-tree complex wavelet can accurately separate surface roughness and waviness [93]. In 2010, Zawada-Tomkiewicz used a neural network to estimate the surface roughness parameter $R_a$. The increment of processing surface image parameters was used as the input of a neural network estimator. Six statistical parameters were calculated by the small decomposition method. This method has great practicability for monitoring the quality of turning surface roughness [94]. In 2010, Huang et al. proposed a novel roughness separation method based on B spline wavelet to analyze engineering surface. This method improved the accuracy of roughness signal separation. The experimental results showed that B spline wavelet could accurately extract the roughness component [95]. In 2014, Ren et al. added bilateral filtering to the dual-tree complex wavelet algorithm. Their improved dual-tree complex wavelet not only improves robustness but also accurately separates the roughness and waviness of the three-dimensional nanoscale surface [96]. In 2015, Jana et al. used wavelet analysis to extract wavelength-based surface features from surface profiles. The surface roughness was separated from the actual profile, and the roughness value $R_a$ was obtained [97]. In 2017, Wang et al. used wavelet packet transform technology and a profiler to extract the areal roughness parameters of milling, turning, and grinding surfaces, respectively. The experimental results showed that the roughness parameters extracted by the wavelet packet transform technology and profiler were basically consistent, further proving the feasibility of wavelet packet transform in surface roughness evaluation [98]. In 2017, Pawlus et al. studied the surface texture measurement error caused by the presence of non-measured points, and used a Talysurf CCI Lite interferometer to measure the surface texture under different illumination conditions. Fifteen surfaces (one- and two-process, isotropic and anisotropic, random and deterministic) were measured and analyzed. The experimental results show that the following parameters are susceptible to errors: skewness $S_{sk}$, areal material ratio $S_{mr}$, as well as the following feature parameters: $S_{pd}$, $S_{da}$, $S_{dv}$, $S_{ha}$ and $S_{hv}$. Compared with deviations in peak portions, inaccuracies of measurement in valley parts of two-process textures usually led to larger errors of parameter computations [99]. In 2018, Duan et al. proposed a surface roughness extraction method based on a dual-tree complex wavelet, which accurately extracted the surface roughness value of an Al/SiCp composite material and maintained high stability [100]. In 2018, García-Plaza et al. applied wavelet packet transform to the monitoring of workpiece surface roughness in CNC machine tools. The results show that mother
wavelet biorthogonal 4.4 in decomposition level L3 with the fusion of the orthogonal vibration
components \( (a_x + a_y + a_z) \) were the best option in the vibration signal and surface roughness
correlation [101]. Wavelet transform can effectively remove the shape deviation and waviness of
the areal surface, and use multiple thresholds to identify features such as peaks and grooves on
the surface. Wavelet transform cannot, however, extract the salient features of a surface based on
the morphological type [102].

6. Conclusions

The surface quality of a mechanical part is an important factor in product performance analysis
and prediction. The surface topography of a part not only can measure the quality of part processing
but also contains a wealth of information regarding features and functions. Its geometrical features
greatly influence the various performance indices of the part and the product. The measurement
and characterization of surface morphology have important theoretical and practical significance
for improving the assembly quality, production process, and product performance.

ISO 16610 provides designers and metrologists with a variety of filtering tools, covering
a wide range of applications. However, not all filters can be applied for a given application case.
In the analysis of surface topography, the user should choose the appropriate filtering method
according to the surface characteristics. Therefore, the future standards will provide users with
the recommendation and feedback to help them select filtering tools and configure the correct
parameters. Future digital filters will address the following challenges in the development of
surface roughness applications:

1. Reducing the difference between various filtering technologies when extracting reference
   benchmarks.
   Although the spline filter and the spline wavelet filter suppress the end effect to some
   extent, they are quite different from the filtering characteristics of the Gaussian filter.
   Different filter references are generated when different filtering techniques are used to filter
   the same surface, resulting in significant differences in the extracted surface roughness
   parameter values. Robust spline filters and robust Gaussian regression filters face the same
   problem. In order to more accurately assess surface roughness while avoiding the huge
   differences caused by different evaluation benchmarks, international standards need to
   achieve uniformity for different filtering technology standards.

2. Improving the areal filtering technology standard system.
   Continuous science and technology improvements place higher requirements on the surface
   roughness of mechanical parts. Two-dimensional surface roughness has significant limi-
tations and is misleading; three-dimensional surface roughness can make up for the lack of two-dimensional surface roughness information. At present, three-dimensional surface roughness parameters have been widely used in the fields of manufacturing, tribology and biomedical research. In the future, the characterization and application of three-dimensional surface roughness will become popular, even replacing the characterization and application of two-dimensional surface roughness. Therefore, international standards need to focus on improving areal filtering technology standards in order to meet development needs. At present, some new filtering techniques have been implemented in the software used for analysis. However, some of the methods described in ISO/TS 16610 have affected the newly developed standards in ISO/TC 213 which have been or will be published in the future.

3. Establishing relevant documentation to guide practical application of the filters.

Although the international standard has developed a relatively complete filtering technology system, many difficulties still remain when applying the filtering technology to practical engineering. It should guide the users to find the appropriate analysis technology for selected application cases. At present, the surface topography analysis software package provides enough filter technology, and the choice for the users is no longer limited. However, the applicability of the selected filter should be thoroughly and rigorously checked for the expected evaluation strategy. It is wise for the researchers to establish a link between the analysis process and the functions they track, and to disseminate good practices to peers.

References


Baofeng He received her B.Sc. degree and M.Sc. degree both from Harbin Institute of Technology in 2006 and 2009 respectively, and Ph.D. degree from Loughborough University in the UK in 2012. Currently, she is a lecturer at Beijing University of Technology. Her main research interests are precision measurement technology and instruments.

Siyuan Ding received his B.Eng. degree from Zhengzhou University in 2016, and M.Eng. degree from Beijing University of Technology in 2020. His main research interests are precision measurement technology and instruments.
Haibo Zheng received his B.Sc. degree in mechatronic engineering from Shandong Agricultural University in 2018. Currently, he is a Post-graduate of Beijing University of technology. His main research interests are precision measurement technology and instrument.

Ruizhao Yang received his B.Eng. degree from Henan University of Science and Technology in 2019. Currently, he is a student at Beijing University of Technology. His main research interests are precision measurement technology and instruments.

Zhaoyao Shi received his B.Sc. degree from Hefei University of Technology in 1984, and M.Sc. degree from Shanxi Institute of Mechanical Engineering in 1988, and Ph.D. degree from Hefei University of Technology in 2001. Currently, he is a professor and Ph.D. supervisor in Beijing University of Technology, and a “Yangtze River Scholar” special professor awarded by the Ministry of Education. His main research interests are precision measurement technology and instruments.