Renishaw plc

New Mills, Wotton-under-Edge,
Gloucestershire, GL12 8JR
United Kingdom

T +44 (0)1453 524524
F +44 (0)1453 524901
E uk@renishaw.com

www.renishaw.com

Introduction to filtering

Contents

1 Why do we filter?

When size and position are reported these are determined by fitting a plane, a circle, or some other geometry using the least-squares method. This means that any noise in the measurement data is averaged out.

This is NOT the case for form, however, as it is calculated as the difference between the maximum and minimum material deviations with respect to the fitted geometry, i.e. the difference between the highest and lowest points. Any noise in the measured scan data directly feeds into the reported value.

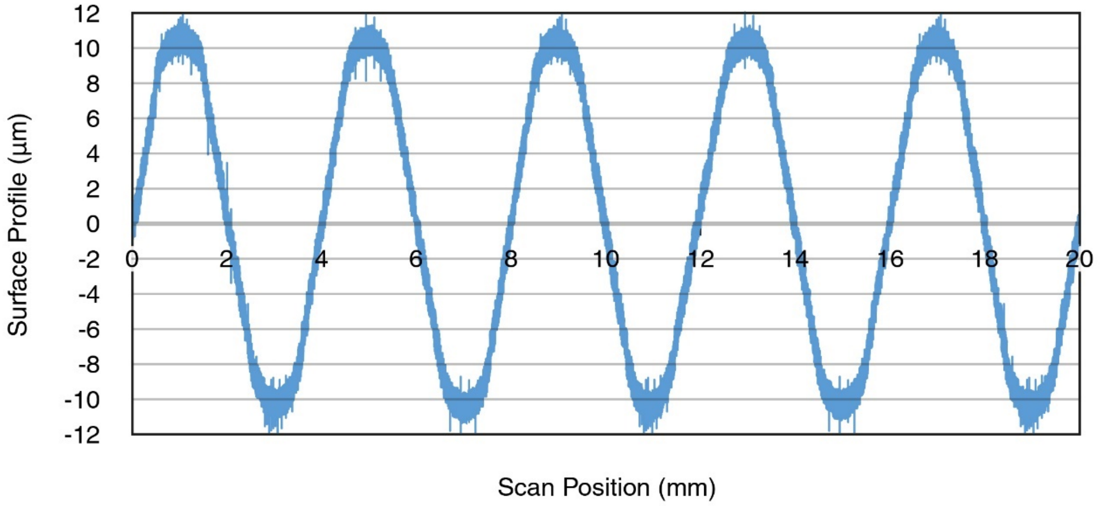
When parts are scanned using a stylus there are many sources of noise. These include:

- electrical noise in the scanning circuitry, particularly the analogue signals in the probe and module
- mechanical noise (affected by the speed and acceleration of the scan)
- environmental conditions (temperature, humidity etc)
- vibration of the stylus on the surface

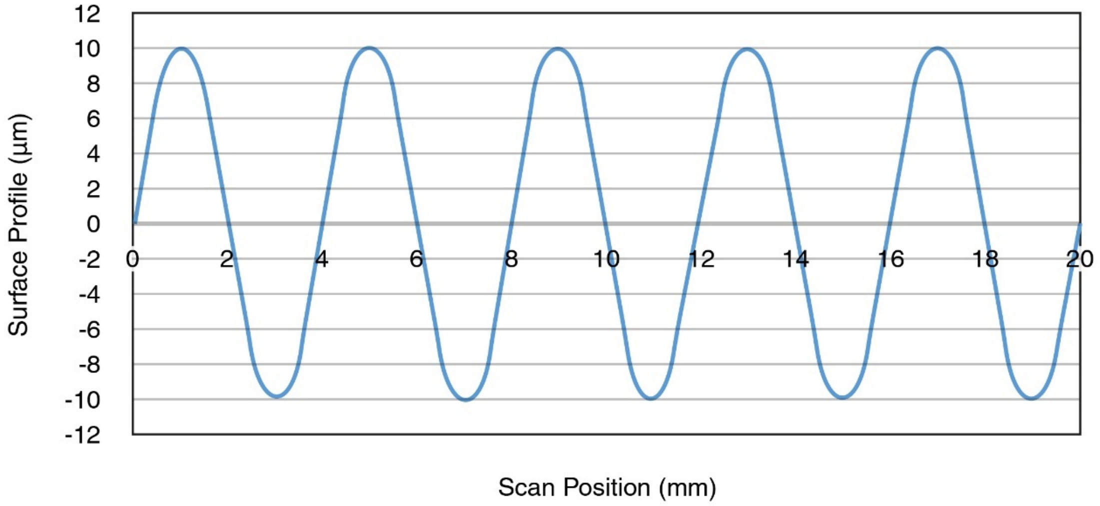
The combined effect of all noise sources appears as random values added to the profile of the surface being measured. Figure 1.1 shows a cross section of a prismatic component with a sinusoid milled into the surface. The top graph shows the raw surface profile with noise and the bottom the surface profile without noise.

Figure 1.1

The surface profile with measurement noise:



The true surface profile:



In this example the flatness of the raw measurement data would be reported as 24 µm, whereas the flatness of the true surface profile is 20 µm. The presence of noise will cause form results to be reported high. Noise also makes repeatability worse so a GR&R will not pass. Noise will also have an adverse impact on material fits / functional fits (e.g. maximum material fit) as the contact points wander from one run to the next. Filtering stabilises the contact areas. So filtering is performed to reduce the noise and provide a better estimate of the form of the part or improve fitting.

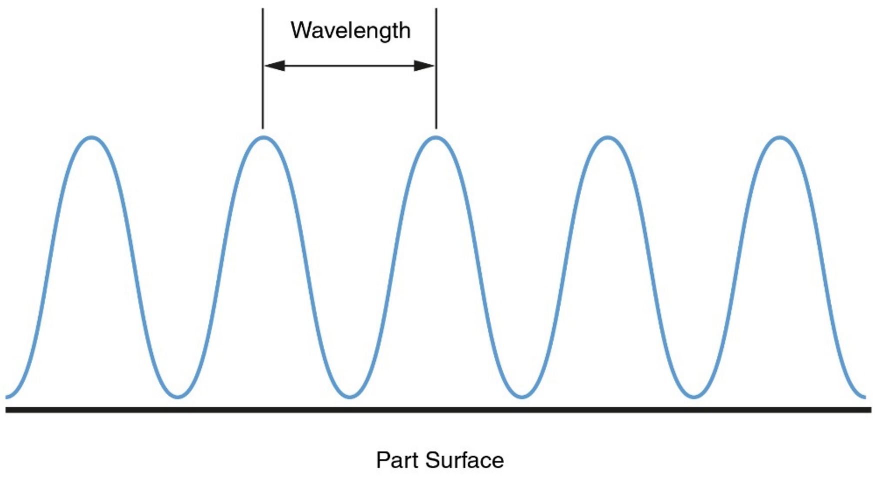
2 Filter concepts

Before we examine how we filter there are a few simple concepts that are key: wavelength, undulations per revolution (UPR), frequency and types of filter.

2.1 Wavelength

Wavelength, often represented as λ (the Greek letter lambda), is the main attribute that is used to specify a filter on flat features such as lines and planes. The wavelength is the distance between the peaks as shown in the cross section of the surface of a part in figure 2.1.

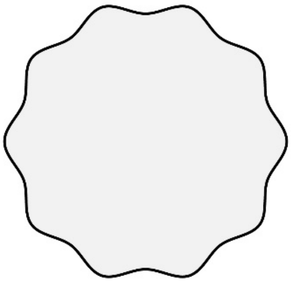
Figure 2.1 - The cross section of surface of a part showing the wavelength is the distance between peaks



2.2 Undulations per revolution (UPR)

Undulations Per Revolution (UPR) is the main attribute that is used to specify a filter on features with a circular cross-section such as circles, cylinders, cones, and spheres. This UPR is the number of peaks over a single revolution (360º). A round bar with a milled sinusoidal form of 10 UPR is shown in figure 2.2.

Figure 2.2 - The surface of a round bar with 10 Undulations Per Revolution (UPR)



2.3 Relationship between UPR and wavelength

As undulation is just another word for wave then the relationship between UPR and wavelength is:

Wavelength = Circumference / UPR (2.3.1)

and

UPR = Circumference / wavelength (2.3.2)

For example, if the round bar with a sinusoidal form of 10 UPR is shown in figure 2.2 has a diameter of 10 mm then its circumference is 10π = 31.4 mm (to nearest 0.1 mm). Using wavelength = circumference / UPR we can determine that the wavelength = 31.4 / 10 = 3.14 mm.

2.4 Frequency

Frequency is defined as the number of waves or undulations per second and is related to scan speed according to:

Scan speed = Frequency × wavelength (2.4.1)

Frequency is not really used when we apply filters in a metrology setting as wavelength and UPR are more convenient to deal with. It is, however, important in understanding the different types of filter described in the next section. The important concepts to take on board are that frequency is proportional to UPR and inversely proportional to wavelength. For example:

- Frequency is proportional to UPR
- Frequency is proportional to 1 / wavelength

2.5 Amplitude

Amplitude is the deviation from the fitted feature along the normal for each scan point. For a plane feature this will be the height above or below the fitted plane. Material-on has a positive amplitude, material-off has a negative amplitude. For a bore this will be the radial distance from the fitted circle. Material-on (smaller radial distance from fitted centre than the fitted radius) has a positive amplitude, material-off on (larger radial distance from fitted centre than the fitted radius) has a negative amplitude. The difference between the maximum amplitude and minimum amplitude yields the form of the measurement.

2.6 Filter types (low-pass, high-pass and band-pass)

Fundamentally, a filter reduces the amplitude of the signal. This amplitude reduction is dependent on the frequency (UPR or wavelength) of the signal.

The types of filter are described using terminology from the world of electronics - low-pass, high-pass and band-pass filters. They are defined in terms of frequency so that:

- Low-pass means that low frequencies are “passed” or allowed through and high frequencies are “blocked”. This in turn means that low UPR are passed and high UPR blocked and that high wavelengths are passed and low wavelengths are blocked.
- High-pass means that high frequencies are “passed” or allowed through and low frequencies are “blocked”. This in turn means that high UPR are passed and low UPR blocked and that low wavelengths are passed and high wavelengths are blocked.
- Band-pass means that a range of frequencies are “passed” or allowed through and frequencies above and below are “blocked”.

In a real filter, frequencies are not fully passed or fully blocked - there will be a known relationship between the frequency and the amplitude called the transmission characteristic or frequency response. Figure 2.6 shows graphs of the transmission characteristic for a Gaussian filter - one of the standard types of filter used in metrology.

Each graph shows the reported form after applying the 2.6(a) shows the low-pass transmission characteristic in terms of UPR and 2.6(b) shows the low-pass transmission characteristic in terms of wavelength.

Figure 2.6(a) - Transmission characteristic for a low-pass Gaussian filter in terms of UPR

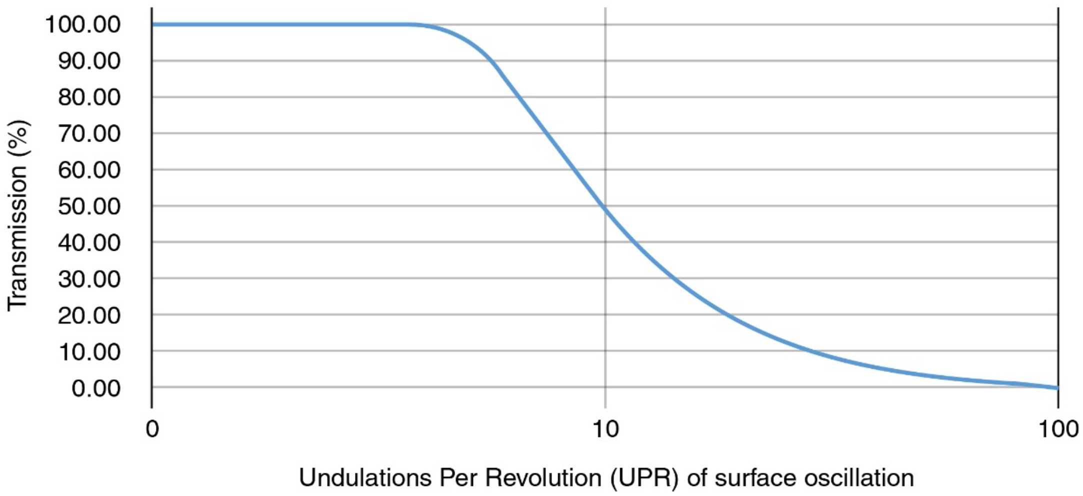
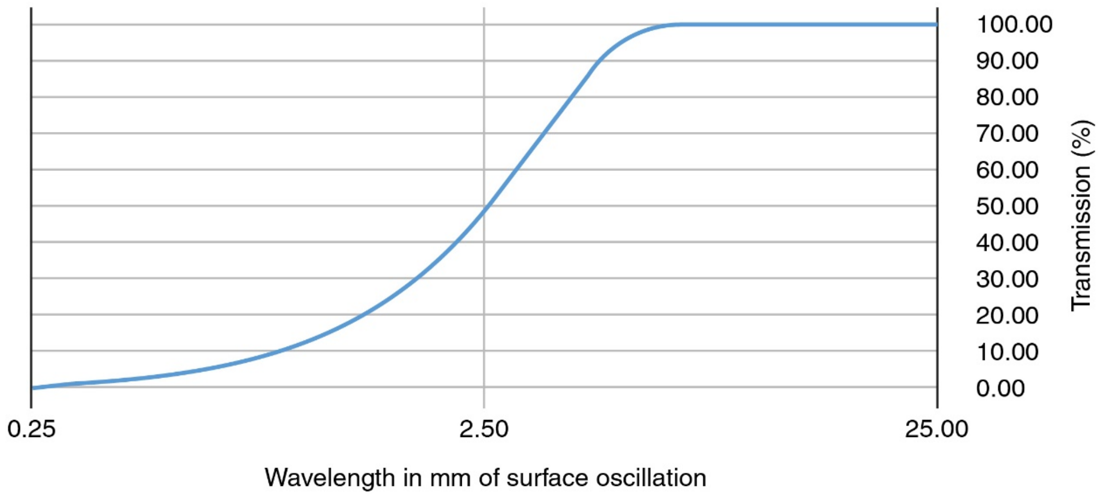


Figure 2.6(b) - Transmission characteristic for a low-pass Gaussian filter in terms of wavelength



2.7 Cut-off frequency and cut-off wavelength

A low-pass and high-pass filter has a single setting that determines the position of the transmission characteristic along the UPR / wavelength axis. This is called the cut-off frequency if we are talking about UPR or cut-off wavelength if we are talking about wavelength. The cut-off determines the UPR or wavelength at which the amplitude is attenuated (reduced) by 50%. Looking back at figure 2.6(a) you can see that this occurs at 10 UPR so 10 UPR is the cut-off frequency for this filter. Similarly, in figure 2.6(b) you can see that this occurs at a wavelength of 2.5 mm so 2.5 is the cut-off wavelength for this filter.

To make it easier to visualise the effect of these filters, imagine a series of parts like those shown in figures 2.1 and 2.2 with a sinusoidal oscillation of 10 microns peak to trough machined into the surface with different UPR / wavelengths. All parts would report 10 microns form prior to filtering. After filtering the parts would report:

Filtered form = Transmission percentage / 100 * unfiltered form

The round bar with a 10 UPR sinusoid would report 50 / 100 * 10 = 5 microns and the plane with the 2.5 mm wavelength sinusoid would report 5 microns.

Typically, the terms “filter UPR”, “filter frequency” or “filter wavelength” are used and these refer to the cut-off values.

Sometimes the term sampling frequency is also used, meaning the number of scan points per unit distance. This is clearly related to the scan pitch and is distinct from the filter cut-off frequency.

3 How do we filter?

3.1 High-pass filter

High-pass filtering is very rarely needed and currently MODUS does not support it.

3.2 Low-pass filter

Low-pass filtering is the main stay of filtering in a metrology setting. This is because the underlying form of a part will have lower frequencies (UPR) or longer Wavelengths than the noise.

Filtering is undertaken using the following steps:

- Fit the appropriate feature (plane, line, circle, cylinder, cone, sphere etc) using the least squares method.

- Obtain the distance error between each raw data point and the fitted feature (material-on with respect to fitted feature has a positive error, material-off with respect to fitted feature has a negative error). This is commonly called the raw error profile.

- Filter the raw error profile using the selected filter type configured for the required cut-off UPR or Wavelength. This gives the filtered error profile.

- Construct the filtered points by using each filtered error profile value as an offset from the fitted feature in the same location as its corresponding raw value.

An example for a cylindrical bore with a 10 UPR sinusoid of 20 microns peak to trough milled into the surface is shown in figure 3.2(a) and 3.2(b).

NOTE: That for clarity the peaks and troughs are exaggerated and only a small number of points together with their error profiles are shown. Real-world data will contain many more points.

Figure 3.2(a) - A bore with a 10 UPR sinusoid milled into the surface showing the filtered point data, the fitted circle, and the raw and filtered error profiles

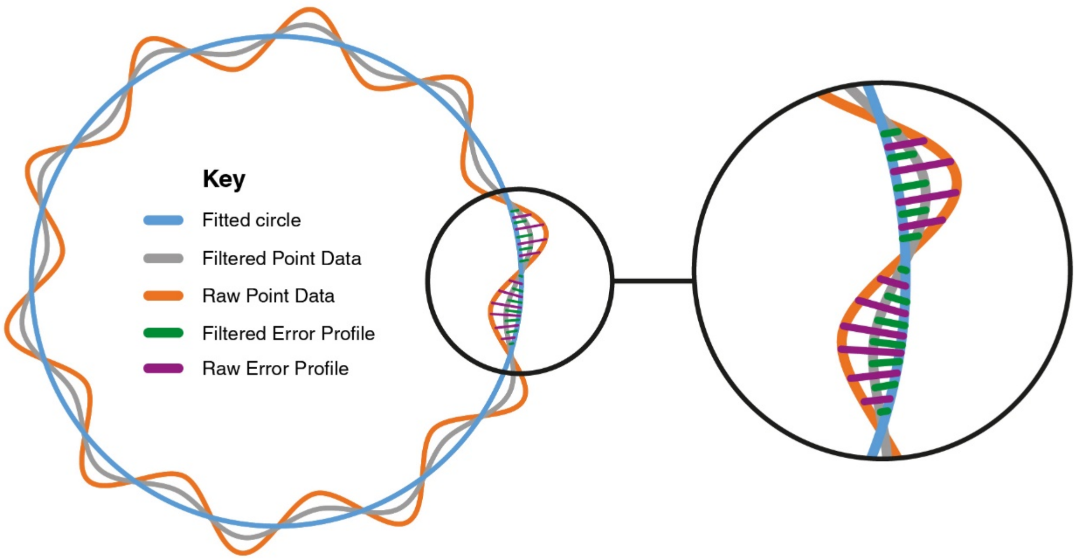
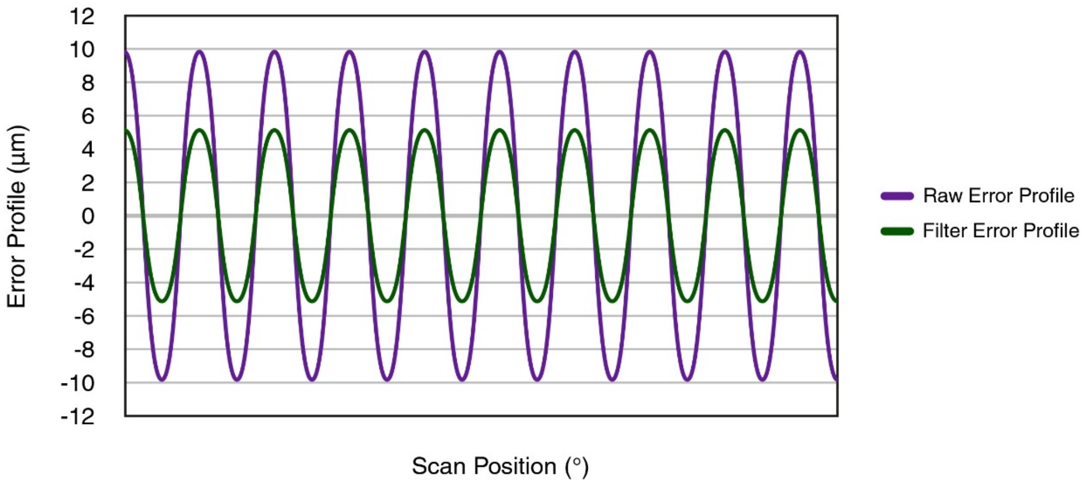


Figure 3.2(b) - The raw and filtered error profiles associated with the bore shown in figure 3.2(a)



4 Filter methods

The main filter methods that are used for metrology are:

- 2RC filter (historical, poor performance)
- Gaussian filter (standard, fast - ISO16610-21 and ISO16610-28)
- Spline filter (standard, fast - ISO16610-22)
- Robust Gaussian filter (iterative, slow- ISO16610-31)
- Robust Spline filter (iterative, slow- ISO16610-32)

These filters were developed in this order chronologically, with each method an improvement on the previous method. These filters are all available in MODUS 1.12 apart from the robust spline filter. The robust filters are both iterative and this can commonly impart an unacceptable time penalty in an application (where 50K points might take a few seconds using the standard filters the robust filters can take several minutes). It is also worth pointing out that the 2RC filter is not commonly used as the Gaussian filter superseded it due to its better performance. The 2RC filter would only ever be used if a customer wished to correlate to an application employing it and was unwilling to use a better filter. Consequently, with the 2RC filter having poor performance and the robust filters being slow the Gaussian and standard spline filters are the filters most commonly used in 99% of applications.

4.1 Gaussian low-pass filter

The Gaussian low-pass filter smooths a region around each point specified by the Gaussian (bell) curve as shown in figure 4.1(a). The Gaussian low-pass filter has a closed and open form. The closed form is used if the scan ends where it started. This is commonly found on bores when scanning 360 degrees or on a gasket type plane scan. The open form is used if the scan start and finish are not in the same place on the part. This filter method can be implemented as a Fourier Transform or Convolution (ISO16610-21) with both methods producing the same result. You can picture the Convolution as a moving window that calculates a weighted average over a fixed number of raw error profile values. Figure 4.1(b) shows the process: a is the “kernel” of filter weights, m is the raw error profile and n is the filtered error profile. Only three filter weights are shown for simplicity; in the real filter there will be tens or hundreds of weights in the kernel.

In the case of the closed filter, we have a continuous set of data and the filter can operate correctly at all positions. However, in the case of the open profile there is no data at the start and finish of the scan to use to calculate the filtered error. This would cause unintentional changes to the filter response and these changes are called end-effects. One solution is to just truncate the filtering at the start and finish to end up with a smaller data set than the original. This is not always desirable and there are a number of end-effect correction methods used to obtain the full data set (see Appendix A - ISO16610-28). Each of these methods has pros and cons but the best general method is called Moment Retainment, and this is the default in MODUS 1.12. Prior to MODUS 1.12 zero-padding was used for open profiles. The user could choose between accepting a non-ideal result or to trim the poor data from each end.

Figure 4.1(a) - Gaussian weighting function

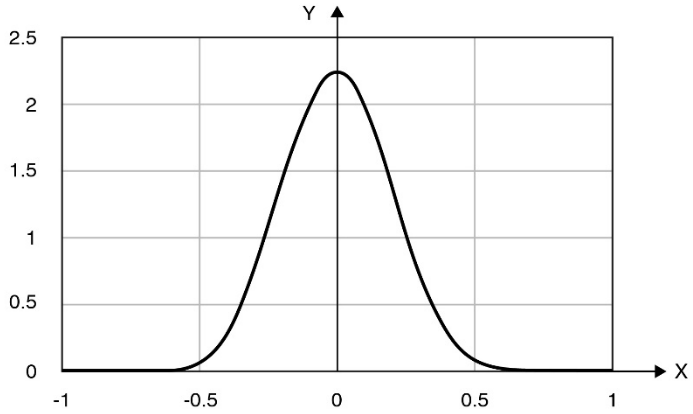
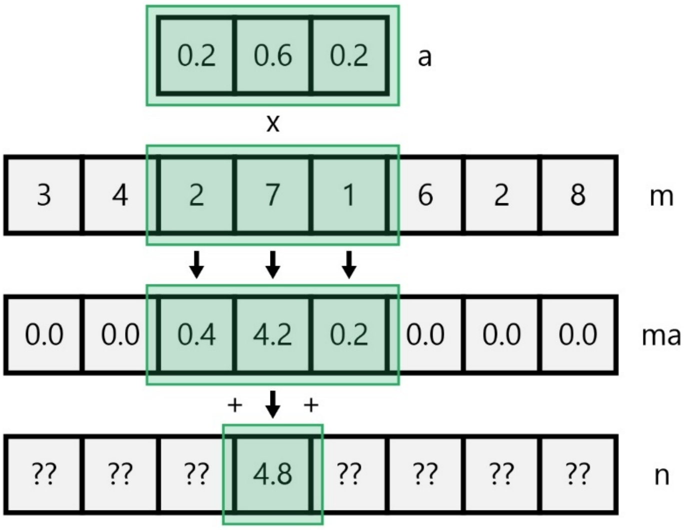


Figure 4.2(b) - Applying the Gaussian filter by Convolution



Even when the closed filter is used for circle scans a 360-degree scan may actually only be 359 degrees and have a small gap between beginning and end. In addition, the 360-degree scan will contain the acceleration and deceleration phases of the scan. These portions may be short but can be noisier mechanically. For this reason, to achieve ultimate performance in terms of best GR&R repeatability it is good practice to over-scan by half a wavelength at each end. For example, if the filter UPR is 50 then the wavelength in terms of degrees is 360/50 = 7.2º. We would scan a circle of minimum 367.2º. To calculate how many points to remove at each end of the scan we work out the percentage of points in the over-scan and multiply this by the total number of points, i.e.

No. points to remove = over-scan / total scan * total points

= 7.2 / 367.2 * 1000

= 20 points

Half this number would be removed from each end.

Some customers specify the amount of over-scan and may be quite a lot more than the minimum, 450º for example (this allows for 45º to be removed either end).

The convolution-based Gaussian filter has an associated variable called the truncation constant (TC) that determines the length of the “tails” of the Gaussian weighting function. Figure 4.1(a) shows tails to ±1 (TC=1) but it could be truncated at ±0.5 (TC=0.5) for example. A lower value increases the speed of the filter (less weights to multiply and add together) but at the expense of accuracy, however with modern fast microprocessors it is best to go for accuracy as the speed overhead is negligible. Therefore, a truncation constant of 1 should be used for all applications.

4.2 Spline low-pass filter (ISO16610-22)

The spline filter is a smoothing filter and is so called because it least-squares-fits a spline curve (cubic curve) around each point of the raw error profile to calculate each filtered value. The spline filter has a closed and open form just like the Gaussian filter. The advantage of a spline filter is that it behaves in a more orderly manner at the ends of open profiles by design. There is a trade-off between the performance at the ends of the scan and how closely the transmission characteristic mirrors the Gaussian filter. This trade-off is determined by the spline tension (B), a value between 0 and 1 that governs how tightly the spline curve fits through the data points. A value of 0.625 provides the response closest to a Gaussian filter. A value of 0 gives a better filter performance at the ends of the scan but with a less Gaussian transmission characteristic.

MODUS 1.12 supports the spline filter and allows the spline tension (B) to be set to 0 or 0.625.

4.3 Filter method comparison

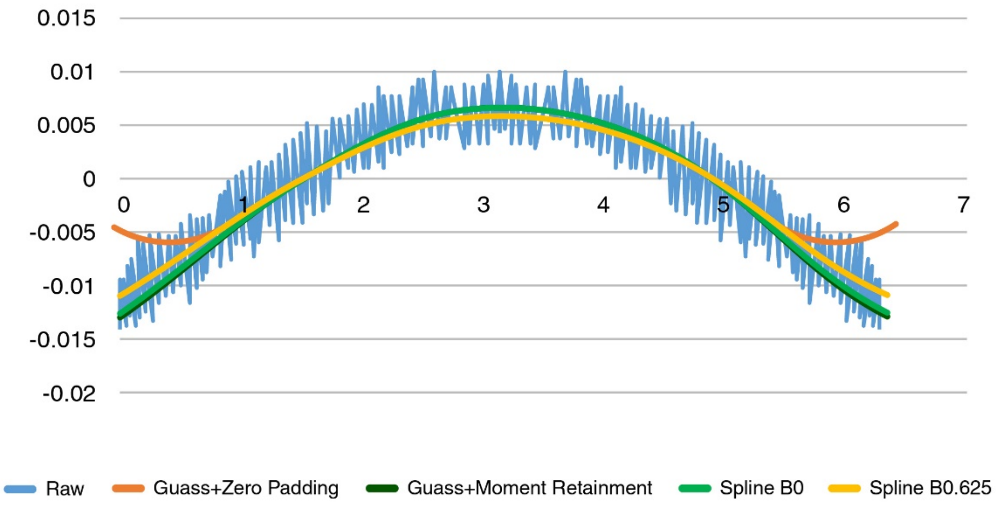
4.3.1 Closed profiles

For closed profiles both the spline filter with B=0.625 and Gaussian filters will give very similar results. Do not use the spline filter with B=0.0 as there is no benefit from improved end-effects but the response deviates more from the ideal Gaussian one.

4.3.2 Open profiles

For open profiles in many situations the filtered data can look quite different depending on the filter method employed. This is shown in figure 4.3. It can be clearly noted that the Gaussian filter with zero padding (the only option prior to MODUS 1.12) performs very poorly at the ends of the scan. The spline filter with Spline Tension (B) =0.625 is also not that good. In this example both the Gaussian filter with moment retainment and the spline filter with spline tension (B) = 0 give good results that track the raw data. Consequently, the spline filter with spline tension (B) = 0 is the preferred method for open profiles. For example, a series of short, line-on-plane scans used to construct a datum plane on a part.

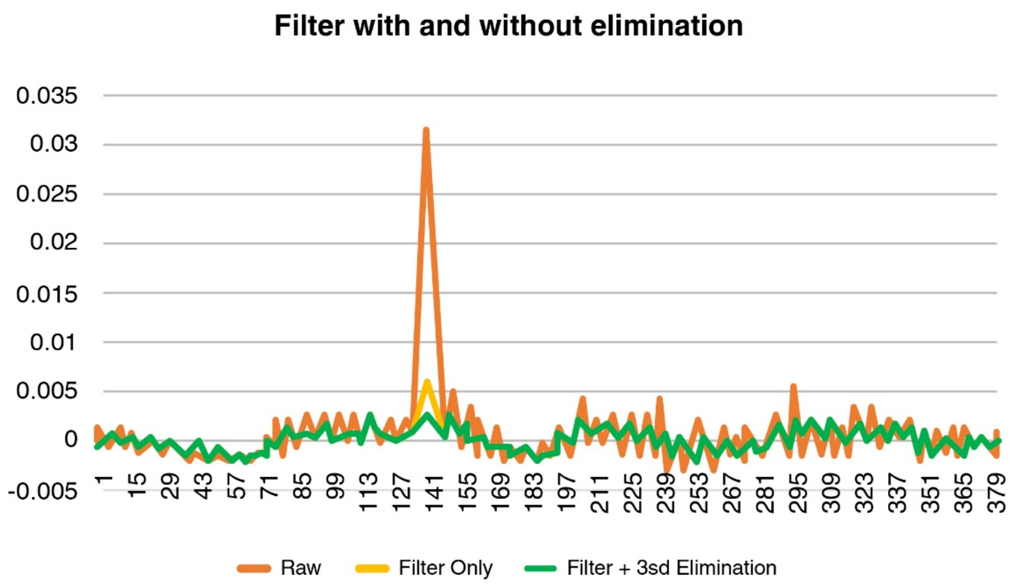
Figure 4.3 - Effect of different filter methods on a short plane in Z measured in the X direction. filter wavelength = 2.5 mm



4.4 Transient response / elimination

It is quite common for transient peaks / troughs of a small number of scan points to be present in the scan data. If a filter is applied to data of this type the transient is “smoothed out” and this can easily be the dominant material condition in the scan. It is common practice to apply statistical elimination prior to filtering to better approximate the true underlying form of the part. This is commonly set to 3 standard deviations from the average raw error profile value. Any data points that exceed plus or minus 3sd are removed from the data set prior to filtering. The effect of filtering with and without elimination are shown in figure 5.3. In this example the forms of the raw, filtered and filtered with elimination are 35, 7 and 5 microns respectively. In some instances, outliers (sometimes referred to as fliers) could have a greater effect and so elimination of 3sd is always recommended.

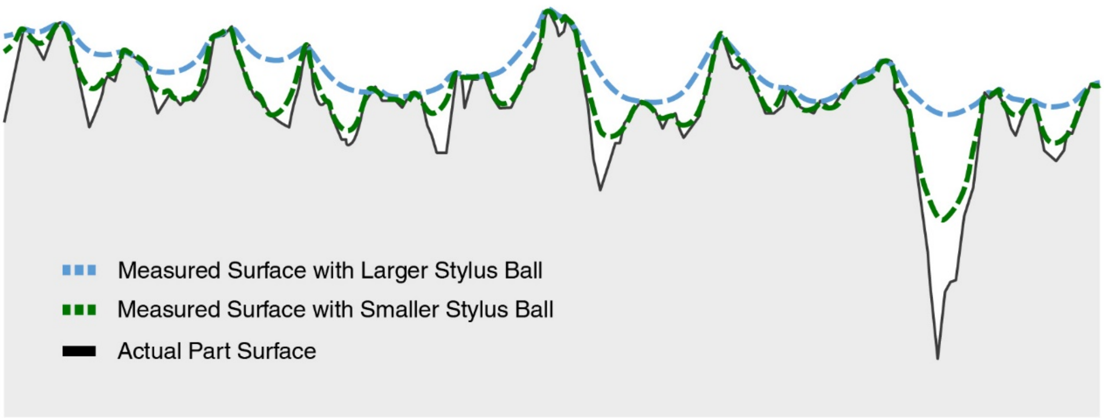
Figure 4.3 - Effect of filter with and without elimination of 3sd



5 Stylus filtering

During scanning the stylus acts as a mechanical low-pass filter (short wavelengths attenuated more than long wavelengths). Larger stylus diameters filter more than smaller diameters. Figure 5.1 shows the effect of different sized styli.

Figure 5.1 - Mechanical filtering with different sized styli



Form can be measured adequately using most standard sized styli in the 1 mm - 6 mm range. In fact, larger styli may be beneficial in that they help reject surface roughness even before filtering is applied in software. If surface finish needs to be measured, then the Renishaw surface finish probe is designed for this purpose and uses styli with dimensions in the microns (rather than in mm). Occasionally there is a requirement to detect wavelengths in the 0.1 mm - 0.5 mm range that are longer than surface finish and shorter than form, referred to as waviness. A smaller stylus to reduce the amount of mechanical filtering will be needed in these cases.

Appendix B quantifies these mechanical filtering effects on some common stylus sizes.

6 Practical considerations

6.1 Filter method

The following filter methods are strongly recommended when using MODUS 1.12:

- If the profile is closed (start and end join up, e.g. bore, gasket-type plane scan) or adequately over-scanned and trimmed use the Gaussian filter.

- If the profile is open, then use the spline filter with spline tension (B) = 0. This will give the best performance at the end regions. If a Gaussian filter is mandated by the customer that is OK as the moment retainment end-effect correction method is used by default.

A number of alternative Gaussian filter end-effect correction methods are available in MODUS 1.12 but these should not normally be selected - use the default moment retainment method. See Appendix A.

6.2 Minimum point spacing

Ideally, there should be a minimum of 10 points per undulation / wavelength:

- Maximum point spacing = Wavelength / 10

- Maximum point spacing = (Diameter * π) / (UPR * 10)

For example:

- A plane feature has a filter of wavelength 2.5 mm applied.
  Maximum point spacing = 2.5 / 10 = 0.25 mm

- A bore has a diameter of 50 mm and a filter of 50 UPR is applied.
  Maximum point spacing = (50 * π) / (50 * 10) = 0.314 mm

MODUS will filter with any number of points but will give a worse form approximation if there are less than 7 points per undulation/wavelength.

In most practical applications it is likely that there will be many times the number of points needed to achieve the maximum point spacing. This is more likely to need to be considered when scanning at high speed with Equator / Equinox / REVO where the number of points needed to filter satisfactorily will determine the upper limit on the scanning speed.

6.3 Recommended settings

6.3.1 Linear features (wavelength – plane / line features)

> 8 mm	Used on poor surface finish when only position is required
2.5 mm	Commonly used assuming part has good surface finish
0.8 mm	Commonly used assuming part has good surface finish
< 0.25 mm	Surface finish

6.3.2 Circular features (UPR – circle, cylinder, cone, sphere)

The UPRs / minimum numbers of points are shown in Table 6.3.2.

Table 6.3.2 – UPRs / minimum numbers of points for different sized features:

Diameter (mm)	Form		Waviness and form		Waviness	
	UPR	Min no. points	UPR	Min no. points	UPR	Min no. points
Up to 8	15	150	50	500	15 - 150	1500
Between 8 and 25						
Between 25 and 80	50	500	150	1500	50 - 500	5000
Between 80 and 250	150	1500	500	5000		
Greater than 250	500	5000	1500	15000	150 - 1500	15000

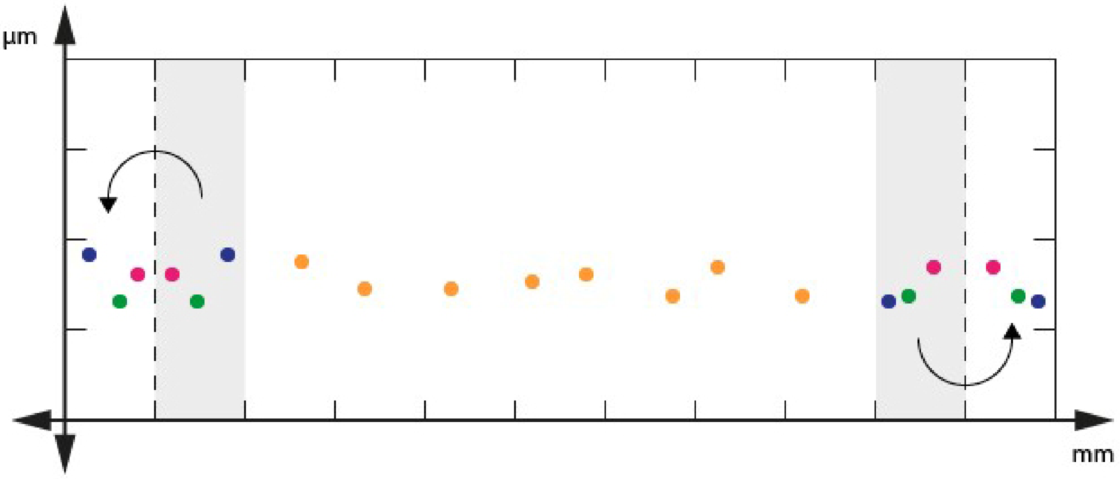
7 Appendix A - End effect correction methods for Gaussian filter

7.1 Linear extrapolation

Extrapolates the data at each end using a straight line fit. Values are placed into the end-effect regions of the profile buffer. This maintains the gradient of the profile.

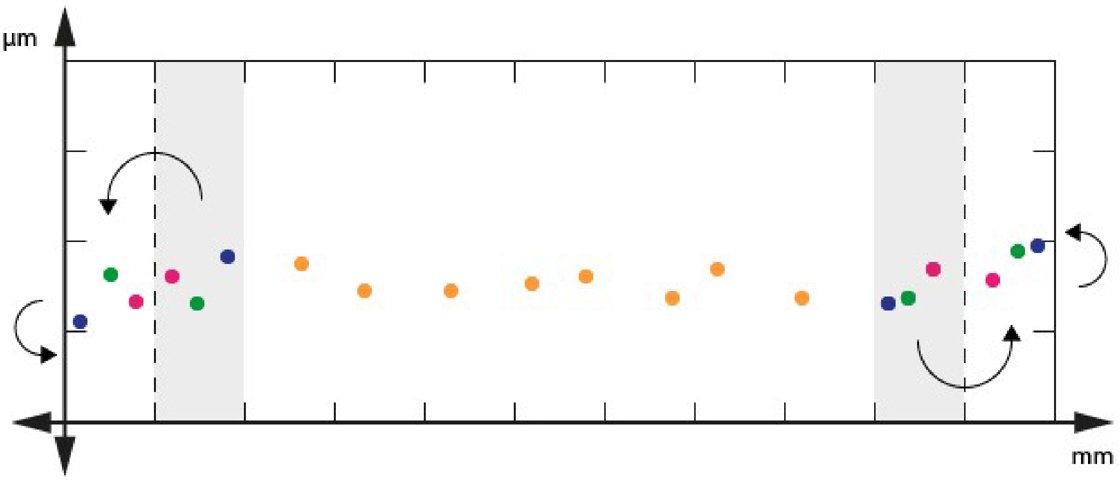
7.2 Line symmetrical reflection

Adds mirrored data around a vertical line at the start and end positions. Reflected values are placed into the end-effect regions of the profile buffer. It is as though a vertical mirror were placed at the start of each end-effect region (colours are used for the points in the end-effect regions to show how the mirroring works).



7.3 Point symmetrical reflection

Adds mirrored data around vertical and horizontal lines at the same time at the start and end positions. Rotationally symmetric reflected values are placed into the end-effect regions of the profile buffer. It is as though a vertical and a horizontal mirror were placed at the start of each end-effect region (colours are used for the points in the end-effect regions to show how the mirroring works).



7.4 Moment retainment

Adds values into the end-effect regions of the profile buffer based on adjacent data. This method seeks to give the best general performance and consequently is the default method in MODUS 1.12. With this method, the filtered curve follows the shape of the data best.

7.5 Zero padding

The simplest method. Adds zeroes into the end-effect regions. Performance is poor. Not recommended.

8 Appendix B - Effect of mechanical filtering on scanned profile

Figure 8.1 shows the effect of mechanical filtering on scanned profile. The important points to note are that:

- Filter transmission = 0 is 0% transmission, none of the signal is seen in the result
- Filter transmission = 1 is 100% transmission, the entire signal is seen in the result
- 1 mm stylus attenuates shorter wavelengths in the key 0.1 mm - 0.5 mm range (waviness) by more than a 3 mm stylus
- 3 mm stylus attenuates shorter wavelengths in the key 0.1 mm - 0.5 mm (waviness) range by more than a 6 mm stylus
- The greater the expected surface profile height (amplitude), the greater the mechanical filtering effect

Figure 8.1 - Effect of mechanical filtering on scanned profiles: