Institute of Physics Publishing Journal of Physics: Conference Series **13** (2005) 254–257
doi:10.1088/1742-6596/13/1/059 7th International Symposium on Measurement Technology and Intelligent Instruments

Linear and robust gaussian regression filters

J Seewig

Institute for Measurement and Control, University of Hanover
Nienburger Straße 17, D-30167 Hanover, Germany

E-mail: joerg.seewig@imr.uni-hannover.de

Abstract. This paper presents a brief overview about gaussian regression filters to extract surface roughness. The mathematical background in the spatial as well as in the frequency domain is discussed. It is shown that gaussian regression filters work without any running in and running out sections and can approximate form up to p^{th} degree. In the industrial world it is well known that linear filters are non robust, which means that any protruding peak or valley (also called "outlier") leads to a distorted roughness topography and effects the calculation of surface parameters directly. In particular plateau like surfaces are good candidates for such critical datasets. In the paper it is shown that such a distortion can be avoided by choosing an appropriate ψ- function. This proceeding leads to the so called robust gaussian regression filter with all the advanced properties of the linear one.

1. Introduction

The gaussian filter according to ISO 11562 [1] is a worldwide used filter to separate roughness from longer waved components like form, form deviation and waviness in a measured profile. In many cases the gaussian filter works satisfactorily. However, looking at modern manufacturing processes, there is a lack of robustness when filtering profiles with function relevant structure elements (for example laser holes or hard particles in metal matrix composites). In other cases the running in and running out sections of the standardised filter process shorten the filter mean line in a not acceptable manner. Also form filtering can be a critical application for the standardised gaussian filter. For this reason new filter methods were developed in recent years. In germany two of these filter methods became very popular, namely the spline filter proposed by Krystek [2] and the gaussian regression filter proposed by Bodschwinna and (last but not least) the author himself [3]. Both filters became integral parts of the ISO 16610 series. Since only german references are available up to now this paper gives a brief overview about the gaussian regression filter technique.

2. A brief overview about gaussian regression filters

2.1. The filter equation and its spatial properties.

The gaussian regression filter technique is based on the following modified Savitzky Golay filter [4] of p^{th} degree:

$$\rho\left(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k\right)^T\cdot\mathbf{s}_k \quad\rightarrow\quad \underset{\boldsymbol{\beta}_k}{Min} \tag{1}$$

with the matrix

$$\mathbf{q}_k^j := \begin{bmatrix} \Delta q_{0,k}^j \\ \Delta q_{1,k}^j \\ \vdots \\ \Delta q_{n-1,k}^j \end{bmatrix},\ \mathbf{X}_k := \begin{bmatrix} \mathbf{q}_k^0 & \mathbf{q}_k^1 & \cdots & \mathbf{q}_k^p \end{bmatrix},\ \boldsymbol{\beta}_k := \begin{bmatrix} w_k \\ \beta_{1,k} \\ \vdots \\ \beta_{p,k} \end{bmatrix},\ \mathbf{s}_k := \begin{bmatrix} \exp\left(-\alpha \cdot \Delta q_{0,k}^2\right) \\ \exp\left(-\alpha \cdot \Delta q_{1,k}^2\right) \\ \vdots \\ \exp\left(-\alpha \cdot \Delta q_{n-1,k}^2\right) \end{bmatrix} \qquad (2)$$

and the parameters (an element of a vector or matrix is indicated by an index and written in italic font)

n	= number of profile data values,	p	= degree of the approximation polynomial,
k,l	$= 0,1,\ldots,n-1$,	j	$= 0,1,\ldots,p$,
$\Delta q_{l,k}$	$= \Delta x\cdot(l-k)$ with Δx the sampling interval,	$\mathbf{e}_n$	$= [1,0,\ldots,0]^T$ vector of dimension $n \times 1$,
$\beta_{r,k}$	$= r^{th}$ derivative at the position k,	α	= smoothing value of the profile filter,
z, w	= vector of dimension $n \times 1$ of the profile data values respectively the filter mean line.		

As a necessary constraint, the first derivative of the minimisation problem in (1) must be zero at the minimum. We yield (the operator diag(**x**) forms a diagonal matrix of vector **x**):

$$\mathbf{X}_k^T \cdot \mathbf{S}_k \cdot \psi\left(\mathbf{z} - \mathbf{X}_k \cdot \boldsymbol{\beta}_k\right) = \mathbf{0} \text{ with } \psi(x) := \frac{\partial \rho(x)}{\partial x} \text{ and } \mathbf{S}_k := \operatorname{diag}(\mathbf{s}_k) \qquad (3)$$

Unfortunately, the unknown parameters $\boldsymbol{\beta}_k$ cannot be calculated directly because equation (3) is nonlinear. To overcome this problem, we choose newtons method, respectively we linearise equation (3). This leads to an iterative procedure for the unknown parameters $\boldsymbol{\beta}_k$:

$$\mathbf{X}_k^T \cdot \mathbf{S}_k \cdot \psi\left(\mathbf{z} - \mathbf{X}_k \cdot \boldsymbol{\beta}_k^m\right) - \mathbf{X}_k^T \cdot \operatorname{diag}\left(\psi'\left(\mathbf{z} - \mathbf{X}_k \cdot \boldsymbol{\beta}_k^m\right)\right) \cdot \mathbf{S}_k \cdot \mathbf{X}_k \cdot \left(\boldsymbol{\beta}_k^{m+1} - \boldsymbol{\beta}_k^m\right) = \mathbf{0} \qquad (4)$$

where m is the iteration step. Taken into account that the derivative ψ' in equation (4) can be nearly expressed as $\psi'(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k^m) \cong \operatorname{diag}(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k^m)^{-1}\cdot \psi(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k^m) := \delta(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k^m)$, we get a weighted least square result from (4):

$$\boldsymbol{\beta}_k^{m+1} = \left(\mathbf{X}_k^T \cdot \operatorname{diag}\left(\delta\left(\mathbf{z} - \mathbf{X}_k \cdot \boldsymbol{\beta}_k^m\right)\right) \cdot \mathbf{S}_k \cdot \mathbf{X}_k\right)^{-1} \cdot \mathbf{X}_k^T \cdot \operatorname{diag}\left(\delta\left(\mathbf{z} - \mathbf{X}_k \cdot \boldsymbol{\beta}_k^m\right)\right) \cdot \mathbf{S}_k \cdot \mathbf{z} \qquad (5)$$

Equation (5) is the exact solution for the mean line **w** and the profile derivatives $\beta_{r,k}$ of the gaussian regression filter. However, the approximation $\operatorname{diag}(\delta(\mathbf{z}-\mathbf{X}_k\cdot\boldsymbol{\beta}_k^m))\cdot\mathbf{S}_k \cong \operatorname{diag}(\delta(\mathbf{z}-\mathbf{w}^m))\cdot\mathbf{S}_k$ simplifies the given formula drastically and leads to an iterative approach for the gaussian regression filter:

$$w_k^{m+1} = \mathbf{e}_{p+1}^T \cdot \left(\mathbf{X}_k^T \cdot \operatorname{diag}\left(\delta\left(\mathbf{z} - \mathbf{w}^m\right)\right) \cdot \mathbf{S}_k \cdot \mathbf{X}_k\right)^{-1} \cdot \mathbf{X}_k^T \cdot \operatorname{diag}\left(\delta\left(\mathbf{z} - \mathbf{w}^m\right)\right) \cdot \mathbf{S}_k \cdot \mathbf{z} \qquad (6)$$

In matrix notation with $\mathbf{M}^m$ of dimension $n \times n$ we get $\mathbf{w}^{m+1} = \mathbf{M}^m\cdot\mathbf{z}$. Each column of the matrix $\mathbf{M}^m$ corresponds to a space variant weighting function at position k.

The filter equation is defined over the range $k = 0,\ldots,n-1$ and has therefore no running in and running out sections. Moreover the filter has a vanishing moment up to p^{th} degree. To show this property, we locally expand an arbitrary profile in its Taylor series $\mathbf{z} = \mathbf{X}_k\cdot\boldsymbol{\beta}_k$. Inserting the profile **z** in equation (6) gives the equality $w_k^{m+1} = w_k$.

2.2. The linear case

In order to derive the linear gaussian regression filter, we have to set $\rho(x) := x^2$ respectively $\psi(x) := 2x$. This leads to $\delta(\mathbf{z}-\mathbf{w}^m) = 2$ and the solution for the linear case:

$$w_k^{m+1} = \mathbf{e}_{p+1}^T \cdot \left(\mathbf{X}_k^T \cdot \mathbf{S}_k \cdot \mathbf{X}_k\right)^{-1} \cdot \mathbf{X}_k^T \cdot \mathbf{S}_k \cdot \mathbf{z} \qquad (7)$$

which doesn't depend on the mean line $\mathbf{w}^m$ and is therefore resolvable in one step. Due to linear operation, we are able to define a transmission characteristic $S(\lambda)$ for a sinusoidal profile of arbitrary wavelength λ, as shown below (2.2.2 and 2.2.3).

2.2.1. *Choosing the smoothing parameter* α. In roughness measurement technique, the filter is characterised by its cutoff wavelength λc (nesting index). That means, the amplitude of a sinusoidal profile with a wavelength λc is damped to 50%. To calculate the cutoff wavelength for the linear gaussian regression filter we take a complex exponential function $\exp(-i\cdot 2\pi\cdot k\cdot\Delta x\cdot\lambda c^{-1})$ as the profile $\mathbf{z}$ with an infinit number n of data values and a sampling interval Δx which tends to zero. This simplifies the numerical evaluation because any sum can be now approximated by an integral.

2.2.2. *The gaussian regression filter of* 0^{th} *and* 1^{st} *degree.* Once the complete gaussian weighting function lies within the profile, the filter of 0^{th} degree and the filter of 1^{st} degree are nearly identical and corresponds to the discrete version of the profile filter according to ISO 11562. However, the advantage of the regression filter technique over the standardised filter is the fact that no running in and running out sections shorten the filter mean line. Following the definition given in chapter 2.2.1 the transmission characteristic for the 0^{th} and 1^{st} degree filter can be written as:

$$S(\lambda)=\exp\left(-\frac{\pi^2}{\alpha\cdot\lambda^2}\right)\quad\text{with}\quad\alpha=\frac{\pi^2}{\lambda c^2}\cdot\frac{1}{\ln 2}\tag{8}$$

For roughness surface measurement purpose it is recommended to use the 1^{st} degree filter. The reason is simple: calculating the mean line of a sloped profile $\mathbf{z} = [\mathbf{q}_k^0\ \mathbf{q}_k^1]\cdot\beta_k$ we get $w_k = w_k + \|\mathbf{s}_k\|^{-1}\cdot\mathbf{s}_k^T\cdot\mathbf{q}_k^1\cdot\beta_{1,k}$ as the response of the 0^{th} degree filter and $w_k = w_k$ as the response of the 1^{st} degree filter. At the marginal area of the profile the term $\|\mathbf{s}_k\|^{-1}\cdot\mathbf{s}_k^T\cdot\mathbf{q}_k^1\cdot\beta_{1,k}$ of the 0^{th} degree filter is $\neq 0$ and the mean line doesn't follow the slope anymore. The 1^{st} degree filter is the better choice !

2.2.3. *The gaussian regression filter of* 2^{nd} *degree.* The gaussian regression filter of 2^{nd} degree is suitable for both roughness and form filter applications where it is required to approximate curvatures as well. The transmission characteristic is

$$S(\lambda)=\left(1+\frac{\pi^2}{\alpha\cdot\lambda^2}\right)\cdot\exp\left(-\frac{\pi^2}{\alpha\cdot\lambda^2}\right)\quad\text{with}\quad\alpha=\frac{\pi^2}{\lambda c^2}\cdot\left(-1-\mathrm{W}\left(-1,-\frac{1}{2\cdot e}\right)\right)^{-1}\tag{9}$$

and $\mathrm{W}(k, x)$ as the "Lambert W" function [5]. The asymptotic weighting function is:

$$s\left(\Delta q_{l,k}\right)=\sqrt{\frac{\alpha}{\pi}}\cdot\left(\frac{3}{2}-\alpha\cdot\Delta q_{l,k}^2\right)\cdot\exp\left(-\alpha\cdot\Delta q_{l,k}^2\right).\tag{10}$$

The weighting function and transmission characteristic is shown in figure 1 (in comparison, ISO 11562 in light grey).

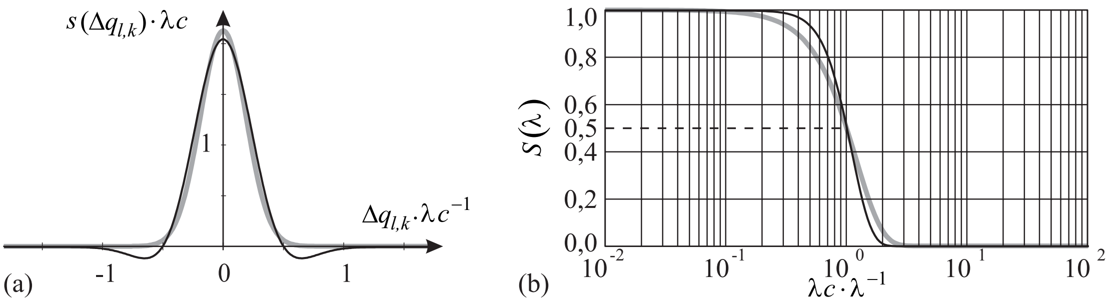

Figure 1. The weighting function (a) and the transmission characteristic (b) of the 2^{nd} degree filter. In light grey ISO 11562.

2.3. The robust case

Typically a linear filter is sensitive against outliers or structure elements like deep pores. For example the grooves of a honed surface will influence the mean line drastically and therefore the roughness profile and its parameters as well. To make the regression filter robust we have to choose an appropriate $\rho(x)$ - respectively $\psi(x)$ - function. In case of the robust gaussian regression filter we take Tukey's biweight function [6] with the definition:

$$\psi(x) = x \cdot \left(1 - \left(\frac{x}{c}\right)^2\right)^2 \text{ for } |x| \le c \text{ and } \psi(x) = 0 \text{ otherwise.} \tag{11}$$

Other functions are defined by Andrews, Huber and Hampel [6]. The constant c is a normalisation factor which ensures the scale independency of the regression filter. Having in mind a least square estimator ($\rho(x) := x^2$) can be seen as the maximum likelihood solution for a gaussian distributed roughness profile superimposed with a long waved component, we define the constant c as three times the standard deviation of the underlying roughness. Than we get $\psi(x) \cong x$ for $x << c$. The standard deviation itself can be estimated using the median absolute deviation which is also known as MAD [6]. For a gaussian distributed roughness profile **z-w** with standard deviation σ we yield

$$\int_{-\mathrm{MAD}\cdot\sigma^{-1}}^{\mathrm{MAD}\cdot\sigma^{-1}} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot z^2\right) \cdot dz = 0{,}5 \quad \Rightarrow \quad \sigma \cong 1{,}4826 \cdot \mathrm{MAD}. \tag{12}$$

In case of robust filtering an iterative approach defined in equation (6) is required. For each step we must calculate a constant $c^m = 4{,}4478 \cdot \mathrm{MAD}$, respectively $c^m = 4{,}4478 \cdot \mathrm{median}(|\mathbf{z}\text{-}\mathbf{w}|)$. The iteration is repeated until the relative change $|c^{m+1} - c^m| < c^m \cdot \varepsilon$ is smaller than the given ε. The initial estimate for the mean line $\mathbf{w}^0$ is the least square approach: $\delta(x)$ = *constant*.

2.3.1. *Choosing the smoothing parameter* α. The robust filter is nonlinear and it makes no sense to determine a transmission characteristic for sinusoidal profiles with arbitrary wavelength. In the robust case we propose to use a feature oriented choice of the cutoff wavelength λc (also known as functional filtering). As a guideline, very good results are obtained when the cutoff wavelength is about three times the feature width being in the profile data set. For example figure 2 shows a profile of an EBT (electrical beam textured) surface with a scratch of 0,8mm width. The cutoff λc equals 2,5mm for both the standardised gaussian filter according to ISO 11562 with running in and running out sections and the robust gaussian regression filter of 2nd degree.

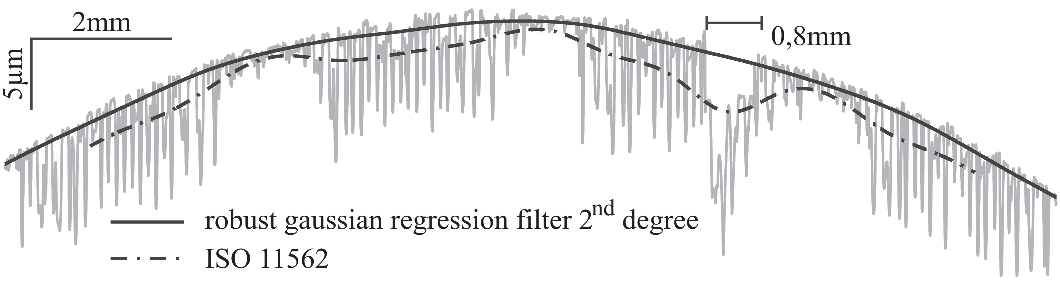

Figure 2. Filtering an EBT profile with the robust gaussian regression filter of 2nd degree and the standardised filter according to ISO 11562. The cutoff wavelength equals λc = 2,5mm.

References

[1] ISO 11562 1996 Geometrical Product Specification (GPS) – surface Texture: Profile method – Metrological characteristics of phase correct filters

[2] Krystek M 1996 Formfiltering by Splines *Measurement* 18 pp 9-15

[3] Seewig J 2000 Praxisgerechte Signalverarbeitung zur Trennung der Gestaltabweichungen technischer Oberflächen *Thesis University of Hanover*

[4] Savitzky A and Golay M J E 1964 *Analytic Chemistry* vol 36 pp 1627-1639

[5] Corless R M, Gonnet G H, Hare D E G, Jeffrey D J and Knuth D E 1996 On the Lambert W Function *Advances in Computational Mathematics* 5 pp 329-359

[6] Huber P J 2004 *Robust Statistics* New York John Wiley & Sons