Precision Engineering 40 (2015) 74–80

Contents lists available at ScienceDirect

Precision Engineering

journal homepage: www.elsevier.com/locate/precision

High-order spline filter: Design and application to surface metrology

Hao Zhang [a,*], Yibao Yuan [b,c], Jin Hua [a], Yuzhu Cheng [a]

[a] College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, China
[b] School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai, China
[c] School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, China

ARTICLE INFO

Article history:
Received 3 April 2014
Received in revised form 2 October 2014
Accepted 9 October 2014
Available online 24 October 2014

Keywords:
Mean line
High-order spline filter
Gaussian filter
Surface metrology

ABSTRACT

The differences between the transmission characteristics of the standard cubic spline filter and the Gaussian filter lead to different evaluation results even for the same profile. There is indeed an adverse impact on the comparison of measurement results and the applications of the related international standards. A novel high-order spline filter is proposed to resolve this practical problem of approximating the Gaussian filtering characteristic. The design of the new filter is based on an improved variational approach by adding the high order derivative terms to the bending energy part whose structural parameters are determined by the aid of the universal Taylor series, so as to realize the convergence to the function of the Gaussian filter. In addition, a cascade algorithm in terms of the low-order filters is also developed in order to ensure stable performance of the high-order filter. Its effectiveness and application were verified by the experiments.

1. Introduction

A profile filter is a mechanical, electrical (analog) or digital device or process which is used to separate the roughness profile from finer fluctuations and from the waviness profile or to separate the waviness profile from the roughness profile. Profile filters with long-wavelength cutoff provide a smooth mean line to a measured profile, thus providing a suitable, modified profile for the calculation of parameters of roughness or waviness with respect to the mean line [1].

In the past 30 years, the profile filters have evolved from mechanical devices or analog processes to digital filters, and finally the international standard 2RC filter has been replaced by the Gaussian filter and the spline filter [2–5]. As a phase correct filter, the Gaussian filter has become the most widely used profile filter. It is recognized as an optimal filter because of its zero-phase characteristic and its minimum product of time width and frequency width [6]. However, the Gaussian filtering algorithm is always disturbed by serious distortions (end effects), because of which the profile ends or area image boundaries cannot be included in an assessment subsequent to the filtering process [7]. In order to overcome this disadvantage, the spline filter was proposed by Krystek as a complementary method for the Gaussian filter [8,9]. Distinct from other filters using the convolution algorithm, the spline filter is implemented in terms of the matrix factorization algorithm [8–11]. The novel algorithm can not only improve the filtering efficiency, but also alleviate significantly end effects with the optional boundary condition applied for different profiles. Owing to these advantages, the spline filter was adopted as a standard profile filter of ISO 16610-22 in 2006 [4].

According to various parts of international standard ISO 16610, the Gaussian filter and the spline filter both can be used to establish the mean line for engineering surface evaluation. However, their different transmission characteristics inevitably determine different filtered results and different mean lines. The matter may also impact the comparison of measurement results and the application of the 16610 standard. To solve this problem, various methods have been developed as substitutes for both spline and Gaussian filters. These include the cascade spline filtering algorithm [12] and the fractional spline algorithm [13]. All of these filters have the similar transmission characteristics to the Gaussian filter's. However, from a proof in Ref. [12], we also know that these algorithms cannot realize the unification of the standard profile filters, because their convergent results always yield an approximation error of at least 0.3% to the standard Gaussian filtering characteristic [2,3]. In addition, these cascade algorithms may exacerbate the end distortion in some way.

In this paper, a novel high-order spline filter based on an improved variational principle is proposed. Its transmission characteristic determined by referring to the standard Taylor series, can

* Corresponding author.
E-mail address: zhanghaowo79@sohu.com (H. Zhang).

http://dx.doi.org/10.1016/j.precisioneng.2014.10.007
0141-6359/

approximate that of the Gaussian filter with high accuracy. Moreover, the filter is carried out with the low-order cascade algorithm [10] so as to ensure its stable application.

2. High-order spline filter

Based on the classical variational principle [10,12], the high-order spline filter for a profile is defined as the following functional minimization procedure:

$$\varepsilon = \sum_{i=1}^{N}[y_i - s(x_i)]^2 + \int_{x_1}^{x_N}\left[\mu_1\left(\frac{ds(x)}{dx}\right)^2 + \mu_2\left(\frac{d^2s(x)}{dx^2}\right)^2 + \cdots + \mu_n\left(\frac{d^ns(x)}{dx^n}\right)^2\right] \times dx \to \text{Min} \quad (1)$$

where y_i are the measured profile data with a constant sampling interval Δx, $s(x)$ is the output profile, i is the index of a point in the dataset and N is the total number of measured data points. In Eq. (1), the first component is supposed to guarantee that the mean line s is to approximate the profile y. The second component called bending energy is to ensure appropriate smoothness of the filtered result [14–16]. Different from the traditional spline filter, the variational function is improved by adding the high-order derivative terms to the bending energy part. Here, $\mu_1, \mu_2, \ldots, \mu_n$ are the regularization parameters to control the compromise between the fidelity and the smoothness of the data. In the case of regularization, the entire solution process of Eq. (1) is equivalent to filtering the data with a low-pass filter. Generally, $n+1$ is defined as the order of the filter.

The digital approximation of Eq. (1) may be written as

$$\varepsilon = \sum_{i=1}^{N}(y_i - s_i)^2 + \sum_{i=1}^{N}[\mu_1(\nabla s_i)^2 + \mu_2(\nabla^2 s_i)^2 + \cdots + \mu_n(\nabla^n s_i)^2] \to \text{Min} \quad (2)$$

where each order derivative of $s(x_i)$ is approximated with finite differences, $\triangledown$ is the difference operator, $s(x_i)$ is denoted by s_i, and

$$\nabla^m s_i = \sum_{k=0}^{m}\{(-1)^k C_m^k \cdot s_{i+\lfloor m/2\rfloor-k}\} \quad (3)$$

where C_m^k are coefficients.

As discussed in Refs. [10,12], the solution to the variational function of Eq. (2) is obtained by the matrix factorization algorithm, and the essential matrix equation can be written as

$$(I+Q)S = Y \quad (4)$$

where I is the identity matrix, Y is the vector of sampled data values, S is the vector of output data values. In practice, the matrix equation is derived by performing the partial derivative operation to s_i. The coefficient matrix Q is different according to the different boundary condition, which may be classified as periodic or non-periodic [4].

For periodic data, the boundary condition is given by

$$s_i = s_{i+N} \quad (5)$$

But, for most instances, the non-periodic boundary is

$$\nabla^n s_1 = \nabla^n s_N = 0 \quad (6)$$

3. The transmission characteristic of the high-order spline filter

Let us now consider Eq. (4) in more detail for the case of a periodic boundary condition, where the partial derivative operation with respect to s_i may be written as

$$\frac{\partial\varepsilon}{\partial s_i} = -2(y_i - s_i) + [-2\mu_1(\nabla^2 s_i) + 2\mu_2(\nabla^4 s_i) + \cdots + (-1)^n 2\mu_n(\nabla^{2n} s_i)] = 0 \quad (7)$$

Using Eq. (7), in the z transform domain [17], we can deduce the transfer function of the high-order spline filter $G(z)$. Here, the z transform of $\triangledown^2$ is equal to $(z-2+z^{-1})$, thus the z transform of $\triangledown^{2n}$ can be expressed as $(z-2+z^{-1})^n$. Deriving the ratio of output to input and replacing the difference operators with $(z-2+z^{-1})^n$ in the z transform domain, we can get the following equation

$$G(z) = \frac{1}{1-\mu_1(z-2+z^{-1})+\mu_2(z-2+z^{-1})^2+\cdots+(-1)^n\mu_n(z-2+z^{-1})^n} \quad (8)$$

Replacing the factor z by $\exp(-j\omega)$, Eq. (8) yields

$$G(\omega) = \frac{1}{1+k_1(1-\cos\omega)+k_2(1-\cos\omega)^2+\cdots+k_n(1-\cos\omega)^n} \quad (9)$$

where $k_1, k_2, \ldots, k_n$ are defined as the structural parameters and

$$\begin{cases} k_1 = 2^1\mu_1 \\ k_2 = 2^2\mu_2 \\ \vdots \\ k_n = 2^n\mu_n \end{cases} \quad (10)$$

Eq. (9) describes the transmission characteristics of the high-order spline filters with arbitrary order. Also, note that the structural parameters are the keys to determine the characteristics. It implies that we have to find a multivariate algorithm to determine these parameters and their corresponding characteristics. Therefore, a comparison method based on the Taylor series is developed.

Firstly, taking into account the function e^x, the Taylor series is given as

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

Similarly, the transmission function of the Gaussian filter can be expanded to

$$e^{-\pi(\alpha(l_c/l))^2} = \frac{1}{e^{\pi(\alpha(l_c/l))^2}} = \frac{1}{(1+\pi\alpha^2(l_c^2/l^2)+\pi^2\alpha^4 l_c^4/2!l^4+\cdots+\pi^n\alpha^{2n}l_c^{2n}/n!l^{2n})} \quad (11)$$

where l and l_c are the numbers of data in the wavelength λ and the cut-off wavelength λ_c respectively. In addition, $\alpha = \sqrt{\ln 2/\pi}$ [1].

Secondly, with the aid of the Taylor series again, we have

$$1-\cos\omega = \frac{\omega^2}{2!} - \frac{\omega^4}{4!} + \cdots - (-1)^n\frac{\omega^{2n}}{(2n)!}$$

where the digital angular frequency ω is equal to $2\pi\Delta x/\lambda$ for the spatial signal. By defining $\lambda = l\Delta x$, then $\omega = 2\pi/l$, Eq. (9) can be written as

$$G(l) = \frac{1}{1+(r_1/l^2)+(r_2/l^4)+\cdots+(r_n/l^{2n})} \quad (12)$$

where r_n are the polynomial coefficients about k_n.