Wang, Jian and Pagani, Luca and Leach, Richard K. and Zeng, Wenhan and Colosimo, Bianca M. and Zhou, Liping (2015) Study of weighted fusion methods for the measurement of surface geometry. Precision Engineering, 47 . pp. 111-121. ISSN 0141-6359

Title: Study of weighted fusion methods for the measurement of surface geometry

Author: Jian Wang Luca Pagani Richard K Leach Wenhan Zeng Bianca M Colosimo Liping Zhou

PII: S0141-6359(16)30126-X
DOI: http://dx.doi.org/doi:10.1016/j.precisioneng.2016.07.012
Reference: PRE 6433

To appear in: *Precision Engineering*

Received date: 29-5-2016
Revised date: 21-7-2016
Accepted date: 26-7-2016

Please cite this article as: Wang Jian, Pagani Luca, Leach Richard K, Zeng Wenhan, Colosimo Bianca M, Zhou Liping.Study of weighted fusion methods for the measurement of surface geometry.*Precision Engineering* http://dx.doi.org/10.1016/j.precisioneng.2016.07.012

Study of weighted fusion methods for the measurement of surface geometry

Jian Wang[1*], Luca Pagani[2], Richard K Leach[3], Wenhan Zeng[2], Bianca M Colosimo[4], Liping Zhou[1]

[1] State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, PR China, jianwang@hust.edu.cn (J Wang), zhoulp@hust.edu.cn (L Zhou)
[2] EPSRC Centre for Innovative Manufacturing in Advanced Metrology, University of Huddersfield, Huddersfield HD1 3DH, UK, l.pagani@hud.ac.uk (L Pagani), z.wenhan@hud.ac.uk (W Zeng)
[3] Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, UK, richard.leach@nottingham.ac.uk
[4] Politecnico di Milano, Dip. di Meccanica, via G. La Masa 34, 20156 Milano, Italy, biancamaria.colosimo@polimi.it
[*] The corresponding author

Highlights:

- Four weighted fusion methods for surface measurement were classified and analysed.
- The uncertainty propagation and the relationship with Kalman filter were analysed.
- Advanced models compatible with different fusion methods were described.
- Experiments verified the effectiveness of weighted fusion on accuracy improvement.
- Weighted fusion complements residual approximation fusion in many fusion scenarios.

Abstract

Four types of weighted fusion methods, including pixel-level, least-squares, parametrical and non-parametrical, have been classified and theoretically analysed in this study. In particular, the uncertainty propagation of the weighted least-squares fusion was analysed and its relation to the Kalman filter was studied. In cooperation with different fitting models, these four weighted fusion methods can be applied to a range of measurement challenges. The experimental results of this study show that the four weighted fusion methods compose a computationally efficient and reliable system for multi-sensor measurement problems, especially for freeform surface measurement. A comparison of weighted fusion with residual approximation-based fusion has also been conducted by providing the input datasets with different noise levels and sample sizes. The results demonstrated that weighted fusion and residual approximation-based fusion are complementary approaches applicable to most fusion scenarios.

Keywords: weighted fusion, multi-sensor measurement, surface reconstruction, uncertainty

1. Introduction

Different sensors have different levels of accuracy, speed, resolution and measurement range. By combining the data from different sensors, a more holistic synthetic measurement can be achieved in an efficient and accurate manner. Such a synthetic measurement method using the integration of multiple sensors is often referred to as multi-sensor measurement, and is currently a focus of attention in surface metrology [1]. For example, structured surfaces, which can be widely found in mechanical/electrical applications [2], usually need sub-aperture stitching to ensure a large measuring range and high spatial resolution. Freeform surfaces [3], widely used in mechanical/optical engineering and aesthetics, usually need to be holistically measured with high accuracy, spatial resolution and without a prohibitive time cost. Single sensors or instruments, such as structured light scanners or coordinate measuring machines, cannot always solve these metrological challenges [4].

Data fusion [5], which is usually a computationally-intensive process, is one of the essential processes in multi-sensor measurement. Data fusion combines data from several information sources into a common representational format, hence the metrological evaluation can benefit from all available sensor information and data [1]. Data fusion usually requires a series of pre-processes before fusion, such as denoising, outlier removal, registration and interpolation. Among these pre-processes, registration, which associates one dataset to another so that the data from different sources are represented in the same coordinate system, is a key challenge for data fusion. Reviews about registration techniques can be found elsewhere [6, 7]. In this study, data fusion was carried out by assuming all the pre-process work has been completed. In the final data fusion process, pre-processed data from different sources are fused to produce a unified output which has improved properties over those from any individual sources.

There has only been limited research into multi-sensor data fusion for surface metrology in recent years. Most fusion methods rely on advanced surface fitting techniques under a residual approximation (RA) framework [5]. For example, in the research work [8, 9], Gaussian process (GP) RA fusion (GPRA), which approximates the residuals between two inputs using a GP model, has been demonstrated to be a successful solution to the fusion of heterogeneous data. Other advanced fitting models, such as B-spline and wavelet models [10, 11], may also be able to provide satisfactory fusion results based on the RA fusion framework.

In contrast to the RA fusion framework, weighted fusion has been used in most studies but has received less attention. It is claimed that weighted fusion can work effectively for homogeneous data, such as data from same or similar level accuracy sensors. Pixel-level weighting [12] and weighted least-squares fusion [13] are two of the most popular weighted fusion methods and have been widely used in research [14, 15]. In addition, weighted fusion may include parametric weighted fusion and non-parametric weighted fusion. These four weighted fusion methods share the common fusion idea, *i.e.* weighting. However, the differences among them have rarely been systematically reported, to the authors' knowledge.

The following sections begin with a theoretical analysis from the classical weighted least-squares fusion and its simplified case, *i.e.* pixel-level weighting. In particular, the uncertainty propagation of weighted least-squares fusion is analysed and its relationship with the Kalman filter is studied. Parametric weighted fusion and non-parametric weighted fusion methods are presented in section 3. Some common reconstruction models based on which the fusion can be processed are introduced in section 4 and in section 5 case studies of the proposed fusion methods are tested on the measurement of three typical surfaces. By providing the fusion input with different noise levels or sample sizes, conclusions about the advantages and disadvantages of weighted fusion methods are given.

2. Weighted least-squares fusion and pixel-level weighting

2.1. The methodology

Weighted least-squares (WLS) fusion is a parametric fitting process. WLS fusion constructs a linear system with finite control parameters to approximate multiple data inputs with different weights [13]. Linear fitting makes the fusion process simple when uncertainty needs to be controlled and fast to compute. WLS fusion is the most popular method and is a fundamental approach with well-developed mathematical foundations. WLS will be introduced in this section.

Given a linear measuring system,

$$\boldsymbol{z} = H\boldsymbol{x} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{x}$ is an n-vector comprised of the model parameters to be measured, H is an m-by-n measurement matrix or model basis function matrix with $m > n$, $\boldsymbol{z}$ is an m-vector representing the measurement result, and $\boldsymbol{\varepsilon}$ is an independent and identically distributed normal noise vector with $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \Sigma) = N(\mathbf{0}, \sigma^2 I)$. This linear system is usually used to approximate measurement problems that do not have systematic error. Given K sample sets from different sources, say $\{\boldsymbol{z}_k\}_{k\in K}$ with noise levels $\boldsymbol{\varepsilon}_k \sim N(\mathbf{0}, \sigma_k^2 I)$, the model parameter vector $\boldsymbol{x}$ can be estimated by minimising the weighted squares cost function

$$\sum_{k\in K} w_k \|\boldsymbol{z}_k - H_k\boldsymbol{x}\|^2, \tag{2}$$

where w_k are designed scalar weights. By setting the weights $w_k = \frac{1}{\sigma_k^2}$, the best linear unbiased estimation (BLUE) of the model parameters can be achieved with the minimum estimation variance (minimum uncertainty in this study) [13].

The minimisation of equation (2) can be achieved by forcing its partial differential equal to $\mathbf{0}$, *i.e.*

$$\sum H_k{}^{\mathrm{T}} W_k H_k \hat{\boldsymbol{x}} = \sum H_k{}^{\mathrm{T}} W_k \boldsymbol{z}_k, \tag{3}$$

where $W_k = w_k I$ and I is the identity matrix. Equation (3) can be conveniently written in the form:

$$H^{\mathrm{T}} W H \hat{\boldsymbol{x}} = H^{\mathrm{T}} W \boldsymbol{z}, \tag{4}$$

where $H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_K \end{bmatrix}$, $W = \mathrm{diag}(w_1 I, w_2 I, \ldots, w_K I)$ and $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \vdots \\ \boldsymbol{z}_K \end{bmatrix}$.

Hence, the weighted least-squares fusion produces the BLUE of the model parameter $\boldsymbol{x}$

$$\hat{\boldsymbol{x}} = \left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}} W \boldsymbol{z} = \left(\sum H_k^{\mathrm{T}} W_k H_k\right)^{-1} \sum H_k^{\mathrm{T}} W_k \boldsymbol{z}_k. \tag{5}$$

The measurement noise in equation (1) may be correlated, *i.e.* $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \Sigma_i)$, where Σ_i is a non-diagonal matrix, and equation (5) is also the BLUE of the model parameters for the generalised least-squares problem, provided $W = \mathrm{diag}(\Sigma_1^{-1}, \Sigma_2^{-1}, \ldots, \Sigma_K^{-1})$. Based on the fused estimation of the model parameters, the model predictions at the original observing positions of the input datasets can be expressed as

$$\hat{\boldsymbol{z}} = H\left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}} W \boldsymbol{z}. \tag{6}$$

The expectation and the variance of the model prediction are respectively given by:

$$E(\hat{\boldsymbol{z}}) = H\left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}} W H \boldsymbol{x} = H\boldsymbol{x} = E(\boldsymbol{z}), \tag{7}$$

$$V(\hat{\boldsymbol{z}}) = H\left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}} W V_{\boldsymbol{z}} W H \left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}} = H\left(H^{\mathrm{T}} W H\right)^{-1} H^{\mathrm{T}}. \tag{8}$$

It is clear that the expectation in equation (7) is unbiased. The principal diagonal elements of the variance-covariance matrix $V(\hat{\boldsymbol{z}})$ correspond to the variance of the prediction $\hat{\boldsymbol{z}}_k$ and hence the prediction variances are usually varying at different observing positions.

In some situations, the input datasets have the same sample size and sampling positions so that the datasets from different sensors naturally have point-wise correspondence. For example, two same-resolution range images may be captured for the same scene under different illumination conditions. The linear fusion problem in equation (1) can be simplified by setting the modelling matrix to an identity matrix, *i.e.* $H_k = I$, and hence $\boldsymbol{z}_k = \boldsymbol{x} + \boldsymbol{\varepsilon}_k, \boldsymbol{\varepsilon}_k \sim N(\boldsymbol{0}, {\sigma_k}^2 I)$. In this case, the fusion estimator $\hat{\boldsymbol{x}}$, which is equal to the fusion predictor $\hat{\boldsymbol{z}}$ at the same sample positions for each of the input datasets**,** can be expressed as the following point-wise weighting or pixel-level weighting form

$$\hat{\boldsymbol{z}} = \hat{\boldsymbol{x}} = (\sum W_k)^{-1} \sum (W_k \boldsymbol{z}_k) = \sum \frac{w_k}{\sum w_k} \boldsymbol{z}_k. \tag{9}$$

The pixel-level weighting is a weighted fusion of input datasets (pixel-wise) at each sample position. This is a very useful method for fusion of range images, in particular when the images have the same sampling resolution and are captured from the same scene. For example, Ramasamy *et al.* [16] applied pixel-level weighted fusion for the fusion of multi-scale range images with specific weight designs.

2.2. Uncertainty analysis

In this section, the uncertainty propagation of WLS fusion (and pixel-level weighting) is analysed and it is demonstrated that WLS fusion can produce improved accuracy reconstruction results.

The estimation and prediction uncertainty of un-weighted linear fusion systems has a simple analytical form [17]. However, the uncertainty propagation of weighted least-squares fusion has not been well investigated, although it is known that equation (6) is a BLUE [13]. It can be shown from equation (8) that, the variances of the predictions (principal diagonal elements of the covariance matrix), or the standard uncertainties, vary at different observing positions. An arithmetical mean of the variance values can usually be taken as a typical indicator representing general changes of uncertainty or accuracy level.

We consider the uncertainty propagation for a simple (un-weighted) linear fitting problem. Given a dataset $\boldsymbol{z}_k$ with the fitting model as in equation (1), *i.e.* $\boldsymbol{z}_k = H_k \boldsymbol{x} + \boldsymbol{\varepsilon}_k$, $\boldsymbol{\varepsilon}_k \sim N(\boldsymbol{0}, {\sigma_k}^2 \boldsymbol{I})$, the least-squares estimation

$$\hat{\boldsymbol{x}}_k = \left({H_k}^{\mathrm{T}} H_k\right)^{-1} {H_k}^{\mathrm{T}} \boldsymbol{z}_k, \tag{10}$$

which results in the maximum Gaussian likelihood [17], provides the BLUE of the model parameters. Then, the model prediction can be computed as

$$\hat{\boldsymbol{z}}_k = H_k \hat{\boldsymbol{x}}_k = H_k \left({H_k}^{\mathrm{T}} H_k\right)^{-1} {H_k}^{\mathrm{T}} \boldsymbol{z}_k. \tag{11}$$

Because a QR factorisation of the modelling matrix has $H_k = QR = [Q_1, Q_2] \begin{bmatrix} R_1 \\ \boldsymbol{0} \end{bmatrix} = Q_1 R_1$, where Q is an orthogonal matrix, and R_1 is an upper triangular matrix, equation (11) can be simplified to

$$\hat{\boldsymbol{z}}_k = Q_1 Q_1^{\mathrm{T}} \boldsymbol{z}_k. \tag{12}$$

Since it is known from equation (10) that $E(\hat{\boldsymbol{x}}_k) = \boldsymbol{x}$, the prediction errors at all the observing positions and their variances can be given by

$$\hat{\boldsymbol{r}}_k = \hat{\boldsymbol{z}}_k - H_k\boldsymbol{x} = Q_1 Q_1^{\mathrm{T}}\boldsymbol{\varepsilon}_k, \tag{13}$$

$$V(\hat{\boldsymbol{r}}_k) = \sigma_k^2 Q_1 Q_1^{\mathrm{T}}. \tag{14}$$

Because Q_1 is a matrix with orthogonal column vectors, the mean squared error (MSE) of the model prediction results, can be calculated with the concise expression:

$$\mathrm{MSE}(\hat{\boldsymbol{z}}_k) = \frac{1}{m_k}E\left(\sum_{i=1}^{m_k}\hat{r}_i{}^2\right) = \frac{1}{m_k}tr\left(V(\hat{\boldsymbol{r}}_k)\right) = \frac{\sigma_k^2}{m_k}tr(Q_1^{\mathrm{T}}Q_1) = \frac{n}{m_k}\sigma_k^2, \tag{15}$$

where $\hat{r}_i$ are the entries of the prediction error vector $\hat{\boldsymbol{r}}_k$, m_k is the size of the measurement point set $\boldsymbol{z}_k$, and n is the number of model parameters. The square root of equation (15) corresponds to the mean standard uncertainty of the least-squares fitting prediction $\hat{\boldsymbol{z}}_k$, and is usually smaller than the sampling noise σ_k .

Now, we demonstrate how weighted least-squares fusion reduces the prediction uncertainty. Given two sets of independent measurement $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ of sizes m_1 and m_2 with different random measurement uncertainties (σ_1 and σ_2), by applying the least-squares fitting given by equation (10), we have the squares of the standard uncertainty associated with the predictions from each individual sets as

$$\begin{aligned} u^2(\hat{\boldsymbol{z}}_1) &= \frac{n}{m_1}\sigma_1^2, \text{ and} \\ u^2(\hat{\boldsymbol{z}}_2) &= \frac{n}{m_2}\sigma_2^2. \end{aligned} \tag{16}$$

According to equation (6), the typical WLS fused prediction at the original sample positions is

$$\hat{\boldsymbol{z}} = H\left(H^{\mathrm{T}}WH\right)^{-1}H^{\mathrm{T}}W\boldsymbol{z}, \tag{17}$$

where $\boldsymbol{z} = \begin{bmatrix}\boldsymbol{z}_1\\ \boldsymbol{z}_2\end{bmatrix} = H\boldsymbol{x} + \boldsymbol{\varepsilon} = \begin{bmatrix}H_1\\ H_2\end{bmatrix}x + \begin{bmatrix}\boldsymbol{\varepsilon}_1\\ \boldsymbol{\varepsilon}_2\end{bmatrix}$, $W = V(\boldsymbol{\varepsilon})^{-1} = \mathrm{diag}(\sigma_1^{-2}I, \sigma_2^{-2}I)$ and $\hat{\boldsymbol{z}} = \begin{bmatrix}\hat{\boldsymbol{z}}_{1f}\\ \hat{\boldsymbol{z}}_{2f}\end{bmatrix}$. $\hat{\boldsymbol{z}}_{1f}$ and $\hat{\boldsymbol{z}}_{2f}$ are respectively the fused prediction results observed on the sample positions of dataset $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. With $V(\boldsymbol{\varepsilon})^{-1} = W$, the prediction error $\hat{\boldsymbol{r}} = \hat{\boldsymbol{z}} - H\boldsymbol{x}$ has a variance-covariance matrix given by

$$V(\hat{\boldsymbol{r}}) = H\left(H^{\mathrm{T}}WH\right)^{-1}H^{\mathrm{T}} = W^{-1/2}Q_{w1}{Q_{w1}}^{\mathrm{T}}W^{-1/2}, \tag{18}$$

where $W^{-1/2} = \mathrm{diag}(\sigma_1 I, \sigma_2 I)$, and Q_{w1} is derived from the QR factorisation $W^{1/2}H = [Q_{w1}, Q_{w2}]\begin{bmatrix}R_{w1}\\ \mathbf{0}\end{bmatrix} = Q_{w1}R_{w1}$. Therefore, the MSE of the prediction results $\hat{\boldsymbol{z}}$ can be expressed as

$$\mathrm{MSE}(\hat{\boldsymbol{z}}) = \frac{1}{m_1+m_2}tr\left(V(\hat{\boldsymbol{r}})\right) = \frac{1}{m_1+m_2}\left(\frac{1}{w_1}\sum_{i=1}^{m_1}\sum_{j=1}^{n}q_{i,j}^2 + \frac{1}{w_2}\sum_{i=m_1+1}^{m_1+m_2}\sum_{j=1}^{n}q_{i,j}^2\right), \tag{19}$$

where $q_{i,j}$ are the entries of the column-orthonormal matrix Q_{w1} and satisfied with

$$\sum_{i=1}^{m_1+m_2} q_{i,j}^2 = 1. \tag{20}$$

The entries of Q_{w1} usually rely on the modelling matrix $H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$, the weighting matrix W and the applied QR factorisation algorithm, such as the Gram-Schmidt process [18]. Given an $(m_1 + m_2) \times n$ modelling matrix *H*, for a specific column vector, the ratio of the sum of squares of the upper m_1 entries to the lower m_2 entries is defined as C_H. If the observing points of each measurement are uniformly distributed (in random) and have large enough sizes, C_H is expected to be m_1/m_2, *i.e.* $\mathrm{E}[C_H] = m_1/m_2$. Because Gram-Schmidt orthogonalisation is linearly operated on column vectors, an orthogonalisation process of *H* retains the same row properties, which indicates that $\mathrm{E}\left[C_{otho(H)}\right] = m_1/m_2$. Therefore, Q_{w1}, which is an orthogonalisation of the row-weighted modelling matrix $W^{1/2}H$, has the expected value

$$\mathrm{E}\left[C_{Q_{w1}}\right] = \frac{\sum_{i=1}^{m_1} q_{i,j}^2}{\sum_{i=m_1+1}^{m_1+m_2} q_{i,j}^2} = \frac{m_1 w_1}{m_2 w_2}. \tag{21}$$

Combined with equation (20), we know that

$$\begin{aligned} &\mathrm{E}\left[\sum_{i=1}^{m_1} q_{i,j}^2\right] = \frac{m_1 w_1}{m_1 w_1 + m_2 w_2}, \text{ and} \\ &\mathrm{E}\left[\sum_{i=m_1+1}^{m_1+m_2} q_{i,j}^2\right] = \frac{m_2 w_2}{m_1 w_1 + m_2 w_2}. \end{aligned} \tag{22}$$

Considering $w_k = \sigma_k^{-2}$, it can hence be concluded that the MSE, *i.e.* the mean squared standard uncertainty in this study, of the fusion predictions has the approximation given by

$$u^2(\hat{\boldsymbol{z}}) = \mathrm{MSE}(\hat{\boldsymbol{z}}) \approx \frac{n}{m_1 w_1 + m_2 w_2} = \frac{n}{\frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2}}. \tag{23}$$

For the fusion of over two sets of data (say *K* sets with individual sizes in m_k), it can be deduced using the same procedures as above that:

$$u^2(\hat{\boldsymbol{z}})_K \approx \frac{n}{\sum_{k \in K} \frac{m_k}{\sigma_k^2}}. \tag{24}$$

Compared to equation (16), it can be simply shown that $u^2(\hat{\boldsymbol{z}}) < u^2(\hat{\boldsymbol{z}}_k)$. In particular, if $m_k = m$ for $k \in K$ and $K = 2$, the reduction of the fusion error from the best individual fitting error has

$$\Delta = \min u^2(\hat{\boldsymbol{z}}_k) - u^2(\hat{\boldsymbol{z}}) = \frac{n \cdot \min \sigma_k^4}{m(\sigma_1^2 + \sigma_2^2)}, \tag{25}$$

which indicates that when σ_k are equal, the error reduction reaches the maximum. If $\sigma_k = \sigma$ for $k \in K$ and $K = 2$, the reduction of the fusion error from the best individual fitting error has the form

$$\Delta = \min u^2(\hat{\boldsymbol{z}}_k) - u^2(\hat{\boldsymbol{z}}) = \frac{n\sigma^2 \cdot}{\sum m_k} \cdot \frac{\min m_k}{\max m_k}, \tag{26}$$

which indicates that when m_k are equal, the error reduction reaches a maximum. When m_k are differentiated, the error reduction decreases.

The following conclusions about WLS fusion can be summarised. (1) WLS fusion can provide improved accuracy results with high probability, and the probability is determined by the sample size and the distributions of observing points. (2) The larger the size of the sample or the smaller the number of modelling parameters, the smaller the fusion uncertainty that can be achieved. (3) If the observing points of individual datasets are not distributed in a uniform or random pattern, or the prediction positions are different from the input datasets, the prediction uncertainty is undetermined.

Figure 1 presents a WLS fusion result and individual least-squares reconstructions of two datasets with different uncertainty from a cubic B-spline curve with the knots [−0.5, −0.4, 0, 0.4, 0.6]. It can be shown that the fused curve is close to the higher accuracy dataset with a small amount of shift towards the lower accuracy dataset. With the WLS fusion, the prediction uncertainty is reduced (see the confidence intervals of Figure 1). With 2000 runs of simulation in which the noise was randomly generated, the mean of the root-mean-squared error (RMSE) of each reconstruction method from the design model was calculated. The results show a steady reduction of the prediction error by approximately 10 % from that of a lower noise set, which coincides with the results in equation (16), (23) and (25).

2.3. The relationship with Kalman filters

Kalman filters (KFs) [19] have been known as an effective tool to fuse new data with previously estimated parameters when multiple datasets are obtained in a time sequence. With KFs, sequentially-obtained data points or sets can be dynamically integrated without requiring all previous data kept in storage. Successful KF applications in surface reconstruction can be found elsewhere [20-22]. In this section, we show that WLS fusion and KFs lead to the same fusion results.

For a linear measuring system, a KF predicts the posterior distribution of $\widehat{\boldsymbol{x}}_k$ based on a prior estimation $\widehat{\boldsymbol{x}}_{k-1}$ and a current observation $\boldsymbol{z}_k$. The KF can be written in the form

$$\widehat{\boldsymbol{x}}_k = \widehat{\boldsymbol{x}}_{k-1} + K_k(\boldsymbol{z}_k - H_k\widehat{\boldsymbol{x}}_{k-1}), \tag{27}$$

where K_k is the so-called Kalman gain and H_k is the measurement matrix for the k^{th} observation. The terms $\boldsymbol{z}_k - H_k\widehat{\boldsymbol{x}}_{k-1}$ are known as the *a priori* prediction residuals, found by subtracting the current (k^{th}) prediction based on the (k-1)$^{\text{th}}$ estimation from the current observations. The optimal Kalman gain can be found by minimising the estimation error, *i.e.* the trace of the estimation covariance, $tr(E[(\widehat{\boldsymbol{x}}_k - \boldsymbol{x}_k)(\widehat{\boldsymbol{x}}_k - \boldsymbol{x}_k)^T])$ [23, 24]. A typical form of the optimal Kalman gain can be written as

$$K_k = V_{\widehat{\boldsymbol{x}}_{k-1}} H_k^{\mathrm{T}} \left(H_k V_{\widehat{\boldsymbol{x}}_{k-1}} H_k^{\mathrm{T}} + V_{\boldsymbol{z}_k}\right)^{-1}, \tag{28}$$

where $V_{\boldsymbol{z}_k}$ and $V_{\widehat{\boldsymbol{x}}_{k-1}}$ are respectively the covariance matrix of the k^{th} measurement and the (k-1)$^{\text{th}}$ estimation. Hence, the variance-covariance matrix of the k^{th} estimation $\widehat{\boldsymbol{x}}_k$ can be updated as [20]:

$$V_{\widehat{\boldsymbol{x}}_k} = (I - K_k H_k) V_{\widehat{\boldsymbol{x}}_{k-1}}. \tag{29}$$

From equation (28), it can be shown that when the noise of the current measurement is extremely low, *i.e.* $V_{\boldsymbol{z}_k} \to \mathbf{0}$, there are $K_k \to {H_k}^{-1}$ (or pseudo inverse ${H_k}^{+}$), which means the fused estimation $\widehat{\boldsymbol{x}}_k$ is mainly determined

by $\boldsymbol{z}_k$ because it is highly reliable. If *a priori* estimation has the covariance matrix $V_{\hat{\boldsymbol{x}}_{k-1}}$ approaching zero, $\hat{\boldsymbol{x}}_k$ is mainly determined by the previous estimation $\hat{\boldsymbol{x}}_{k-1}$ [25]. With equations (27) to (29), the estimation of the model parameters can be recursively updated without referring to all the previous measurement datasets.

Alternatively, the KF model estimation can be computed in a batch mode [20] in the following manner:

$$\hat{\boldsymbol{x}}_k = \left(V_{\hat{\boldsymbol{x}}_0}^{-1} + H^{\mathrm{T}} V_{\boldsymbol{z}}^{-1} H\right)^{-1} \left(V_{\hat{\boldsymbol{x}}_0}^{-1} \boldsymbol{x}_0 + H^{\mathrm{T}} V_{\boldsymbol{z}}^{-1} \boldsymbol{z}\right), \tag{30}$$

$$V_{\hat{\boldsymbol{x}}_k} = \left(V_{\hat{\boldsymbol{x}}_0}^{-1} + H^{\mathrm{T}} V_{\boldsymbol{z}}^{-1} H\right)^{-1}, \tag{31}$$

where $H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_k \end{bmatrix}$, $V_{\boldsymbol{z}}^{-1} = \mathrm{diag}\left(V_{\boldsymbol{z}_1}^{-1}, V_{\boldsymbol{z}_2}^{-1}, \ldots, V_{\boldsymbol{z}_k}^{-1}\right)$ and $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \vdots \\ \boldsymbol{z}_k \end{bmatrix}$.

Normally, $\boldsymbol{x}_0$ is an *a priori* estimation based on null observation. This implies that $\boldsymbol{x}_0$ is equals to $\mathbf{0}$ and $\left\| V_{\boldsymbol{x}_0} \right\| = +\infty$. Hence, equation (30) can be simplified to the WLS fusion form, as in equation (5), where $W = V_{\boldsymbol{z}}^{-1}$. In other words, KF fusion and WLS fusion are essentially equivalent.

KFs were reported to effectively save computation space as they do not require storage for all the previously measured data [20]. By observing equation (5), it is known that WLS fusion has the same advantage because it only requires the storage for historical accumulation results. Regarding the computation complexity, KF fusion is $O(n^3 m)$ if new data of size m is point-wise integrated, and WLS fusion (or the KF fusion in batch mode) is $O(n^2 m + n^3)$, where m is the size of the new dataset and n is the number of model parameters. Therefore, WLS fusion (or KF fusion in batch mode) is preferable for high computation speed in a majority of situations when $m > 1$.

3. Parametric and non-parametric weighted fusion

Some advanced fitting or fusion models for surface modelling or reconstruction are parametric models; but the model parameters cannot be calculated using the linear method described in equation (5). For example, the multilevel B-spline approximation [10] and the least-squares B-spline approximation [26] determine the model parameters using an iterative approximation method. For fusion with these advanced parametric models, one can simply weight the individually-obtained fitting parameters from each dataset for the model estimation.

For example, given each individual dataset $\boldsymbol{z}_k = H_k \boldsymbol{x} + \boldsymbol{\varepsilon}_k, \boldsymbol{\varepsilon}_k \sim N(\mathbf{0}, {\sigma_k}^2 I)$, the model parameters $\hat{\boldsymbol{x}}_k$ can be individually estimated, say $\hat{\boldsymbol{x}}_k = f(\boldsymbol{z}_k, H_k, \sigma_k)$. An appropriate weighted mean of the individual parameters can be taken as the fusion result, *i.e.*

$$\hat{\boldsymbol{x}} = \sum \frac{w_k}{\sum w_k} \hat{\boldsymbol{x}}_k, \tag{32}$$

where w_k are the weights of the k^{th} estimation. Assuming $m_k = m$ and $H_k = H$ for $k \in K$, and given H with the QR decomposition $H = [Q_1 Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1$, the fusion prediction $\hat{\boldsymbol{z}} = H\hat{\boldsymbol{x}}$ has the same unbiased expectation as that from the individual fitting

$$E(\hat{\boldsymbol{z}}) = HE\left(\sum_{k\in K}\frac{w_k}{\sum w_k}\hat{\boldsymbol{x}}_k\right) = H\boldsymbol{x}, \tag{33}$$

and the variance-covariance matrix

$$V(\hat{\boldsymbol{z}}) = \sum_{k=1,\ldots,K}\left(\frac{w_k}{\sum w_k}\right)^{\mathbf{2}}\sigma_k^2 H\left(H^{\mathrm{T}}H\right)^{-1}H^{\mathrm{T}} = \sum_{k\in K}\left(\frac{w_k}{\sum w_k}\right)^{\mathbf{2}}\sigma_k^2\, Q_1 Q_1{}^{\mathrm{T}}. \tag{34}$$

Similar to equation (19), equation (34) can be analysed with the squared mean standard uncertainty of the fusion prediction, thus

$$u^2(\hat{\boldsymbol{z}}) = \mathrm{MSE}(\hat{\boldsymbol{z}}) = \frac{1}{m}\mathrm{tr}\left(V(\hat{\boldsymbol{z}})\right) = \frac{n}{m}\sum_{k\in K}\left(\frac{w_k}{\sum w_k}\right)^{\mathbf{2}}\sigma_k^2, \tag{35}$$

which achieves its minimum with the optimal weight design $w_k = \frac{1}{\sigma_k^2}$. In fact, this minimised prediction variance $u^2_{\min}(\hat{\boldsymbol{z}}) = \frac{n}{m\sum 1/\sigma_k^2}$ is equal to that of the WLS method shown in equation (24). Thus the parametric weighted fusion also has higher accuracy than that from individual datasets. However, it should be noted that the parametric weighting method may only work effectively for the situation when $H_k = H$, *i.e.* when the input datasets have the same sample size and observing position distribution.

Non-parametric models, such as GP models [27], have also been reported recently for surface reconstruction and surface quality assessment, for example in form assessment [28]. Non-parametric models do not have model parameters but use all input data to linearly predict model values. It is, therefore, meaningless to weight model parameters in the same way as with parametric weighting methods. Equation (36) shows a GP model-based linear calculation of the prediction values $\hat{\boldsymbol{z}}_*$ and their variance-covariance

$$\hat{\boldsymbol{z}}_*|\boldsymbol{z} \sim N\left(\boldsymbol{\mu}_* + V_*^{\mathrm{T}}V^{-1}(\boldsymbol{z}-\boldsymbol{\mu}),\ V_{**} - V_*^{\mathrm{T}}V^{-1}V_*\right), \tag{36}$$

where $\hat{\boldsymbol{z}}_*$ is the prediction values of the GP model at $\boldsymbol{x}_*$ positions, $\boldsymbol{z}$ is an input data vector observed at position $\boldsymbol{x}$, $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_*$and are the mean-line vectors of the GP model at $\boldsymbol{x}$ and $\boldsymbol{x}_*$ positions, V and V_{**} are respectively the variance-covariance matrix of the GP model at $\boldsymbol{x}$ and $\boldsymbol{x}_*$ positions, and V_* is covariance of the GP model between $\boldsymbol{x}$ and $\boldsymbol{x}_*$ positions.

By assuming that K input datasets come from several uncorrelated sensors (this is the normal case in practice), the variance-covariance matrix V in equation (36) has the form $V = \begin{bmatrix} V_1 & & 0 \\ & \ddots & \\ 0 & & V_K \end{bmatrix}$. Then, the expectation of the fusion model has the following summation form

$$\begin{aligned} \mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}) &= \boldsymbol{\mu}_* + V_{1*}^{\mathrm{T}}V_1^{-1}(\boldsymbol{z}_1-\boldsymbol{\mu}_1) + \cdots + V_{K*}^{\mathrm{T}}V_K^{-1}(\boldsymbol{z}_K-\boldsymbol{\mu}_K), \\ &= \boldsymbol{\mu}_* + \mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}_1) + \cdots + \mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}_K). \end{aligned} \tag{37}$$

To make the expectation unbiased to that from individual models, we modify the results into the following unbiased weighted summation

$$\mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}) = \boldsymbol{\mu}_* + w_1\mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}_1) + \cdots + w_K\mathrm{E}(\hat{\boldsymbol{z}}_*|\boldsymbol{z}_K), \text{ subject to } \sum_K w_k = 1. \tag{38}$$

In other words, prediction with the non-parametric weighted fusion can be viewed as a weighted summation of individual non-parametric model prediction results.

4. Advanced fusion models

Engineered surfaces often have designed surface geometry [29]. The underlying design models must be known in advance when using the weighted fusion methods described above. However, real-world geometrical products normally have unknown models or have violations from the underlying model due to manufacturing imperfections [29]. Therefore, fusion of the data from such surfaces needs some common models which can be flexible enough to fit an arbitrary shape.

There are a variety of common models applicable to surface metrology, including parametric and non-parametric models. Parametric models such as bivariate polynomials, radial basis functions (RBFs), B-splines, wavelets and NURBS [17] have been widely used in surface reconstruction. None of the models can cover all types of surfaces. For example, RBFs are computationally efficient for scattered data (*i.e.* data not on a uniform grid), but they are difficult to be applied to large datasets due to their global computation characteristic [17]. B-splines and NURBS are computationally efficient but require sophisticated optimisation methods, especially when the number of control points and the knots are unknown [30]. Non-parametric models, such as GP models [28], have recently been proposed for surface reconstruction but they usually have a high computational cost.

In this section, three advanced models with high stability and flexibility are introduced for complex surface fitting and fusion. These three models are respectively applicable to WLS fusion, parametric weighting and non-parametric weighting.

4.1. Tensor product B-spline models

A B-spline is a piecewise polynomial function in degree K (C^K continuity) for argument $x \in \mathbb{R}^1$. A functional surface in 2D form, *i.e.* $z = f(x, y)$, can usually be approximated by a superposition of multiple B-splines in the following tensor product form

$$z = \varphi_{\boldsymbol{a}}(x) \times \phi_{\boldsymbol{b}}(y) = \sum_{k=1}^{n_a} \sum_{l=1}^{n_b} a_k b_l \varphi_k(x) \phi_l(y) = \sum_{k=1}^{n_a} \sum_{l=1}^{n_b} c_{k,l} \psi_{k,l}(x, y), \tag{39}$$

where $c_{k,l} = a_k b_l$ are modelling parameters, $\psi_{k,l}(x, y) = \varphi_k(x)\phi_l(y)$ is the design basis constructed from the tensor product of two sets of orthogonal B-spline basis functions [30, 31], and n_a and n_b are respectively the number of basis functions in the two orthogonal directions. The shape of the B-spline basis functions is dependent on the selection of knots, which determines the modelling accuracy and stability. Generally, regularly distributed knots are preferred choices from a convenience standpoint. Such simple knot settings are flexible enough for most smooth freeform surfaces. In matrix form, equation (39) can be written as

$$\boldsymbol{z} = \Psi \boldsymbol{c}, \tag{40}$$

where $\boldsymbol{z}$ is the size m observation result vector, Ψ is the $m \times n_a n_b$ model matrix, with $\psi_{k,l}$ as the column entries, and $\boldsymbol{c}$ is the $n_a n_b \times 1$ parameter vector with $c_{k,l}$ as the entries. On account of the linear modelling property, B-spline models can be applicable to WLS fusion methods.

B-spline modelling has been found to have many beneficial characteristics for surface metrology. First, B-spline modelling is computationally efficient due to the compact support of the basis functions. B-spline models are also flexible enough to represent complex geometry with varying characteristics and sharp changes by appropriately designing the knots. For example, by assigning dense knots in areas of high slope [17], complex local geometry can be accurately approximated.

4.2. Multilevel B-spline models

Multilevel B-spline approximation (MBA) [10, 32] is a computationally efficient model based on multilevel regular knot setting B-splines, which has improved flexibility over general B-splines. MBA provides an approximation with a sum of multiple model surfaces at different resolution levels, *i.e.*

$$\boldsymbol{z} = \boldsymbol{z}_0 + \boldsymbol{z}_1 + \cdots + \boldsymbol{z}_K, \tag{41}$$

where $\boldsymbol{z}_k \in S_k$ and $\{S_k\}$ are a nested sequence of the tensor product B-spline model (see equation (39)) subspaces and $S_0 \subset S_1 \subset \cdots \subset S_K$. For example, given a set of sample points within a square domain $\Omega = [0, m] \times [0, n]$, a hierarchy of control lattices $\{\Phi_k\}_{k=0,1,\ldots,K}$ can be designed overlaid on the domain Ω as with $\left(2^k + 3\right) \times \left(2^k + 3\right)$ control points, based on cubic B-splines [30]. In other words, $\boldsymbol{z}_k$ corresponds to different spatial frequency components of the source $\boldsymbol{z}$.

In the process of each level control lattice (parameter) estimation, MBA first calculates the neighbour 4×4 control parameters for every sample point in a least-squares manner. Then the independently calculated parameters for each control point are weighted to obtain the final parameter estimation for the whole $\left(2^k + 3\right) \times \left(2^k + 3\right)$ control lattice. Finally, the residuals of the current approximations at the sample data are taken as the input for the parameter estimation of the next level control lattice. The MBA iteratively estimates the control parameters of each level control lattice until an approximation threshold is achieved. On account of the non-linear parameter estimation process, MBA models can cooperate appropriately with the parametric weighted fusion.

MBA models can approximate a discrete surface sample set for smoothing or exact interpolation, by assigning different numbers of approximation levels. The more approximation levels are assigned, the higher the modelling accuracy that can be achieved until the modelled surface passes exactly through every measured point.

4.3. Gaussian process models

Gaussian processes (GPs) are continuous stochastic processes in which any finite sample follows a joint Gaussian distribution, *i.e.*

$$\begin{bmatrix} \boldsymbol{z}_* \\ \boldsymbol{z} \end{bmatrix} \sim \mathbb{N}\left(\begin{bmatrix} \boldsymbol{\mu}_* \\ \boldsymbol{\mu} \end{bmatrix}, \begin{bmatrix} V_{**} & V_*^{\mathrm{T}} \\ V_* & V \end{bmatrix}\right), \tag{42}$$

the notations of which conform to that of equation (36). In contrast to general parametric models, which are usually difficult to calculate due to the problems of under-/over-fit, GP non-parametric models have many advantages, such as ease of use and high flexibility. GP models do not have a finite number of model parameters to estimate, but use all input data to linearly predict model values as in equation (36).

In the last ten years, GP prediction models have been successfully used in geostatistics [33] and surface metrology [28]. As non-parametric models, GP models can work with the proposed non-parametric weighted fusion, *i.e.* weighting individual model prediction results as fusion output.

5. Case studies and discussion

In this section, the four weighted fusion methods presented above are applied to practical measurement challenges, including the measurement of a deterministic surface with explicit functions, a freeform surface with CAD deign model and a freeform surface with an unknown design model.

5.1. F-theta lens

F-theta lenses [34], which are designed to provide a flat field at an image plane, are important devices used in laser scanners and printers. The form accuracy of such F-theta lens surfaces is critical to the quality of the image. The F-theta lens surfaces in this example have a deterministic shape with the following explicit function:

$$\begin{cases} z = -\frac{1}{250}x^2 + \frac{1}{92000}x^4 - \frac{1}{25}y^2 \\ -20 \le x \le 20 \text{ mm},\ -7.5 \le y \le 7.5 \text{ mm} \end{cases}. \tag{43}$$

Figure 2a presents two sets of simple random sample points from the design model within a rectangular patch area. Both of the datasets (blue dots for Set 1 and green dots for Set 2) have the same size with 500 points. To imitate a practical measuring procedure, different Gaussian noise (different noise levels and the same correlation length) [27] were randomly generated and superimposed on the extracted samples. An example of the Gaussian noise can be seen in Figure 2b and 2c. Experiments were then carried out for 200 runs in which the sample noise changes every time. Hence, the performance of the weighted fusion methods could then be statistically analysed.

An example view of the reconstruction results from individual datasets and their weighted (WLS) fusion is presented in Figure 3, in which the design model and a substitute (regular knots) B-spline model are tested. The design model has three natural parameters while the selected B-spline model has 6×6 control parameters, which was demonstrated in this case to provide higher accuracy reconstruction than other options. With the WLS fusion, Figure 3 shows a typical example that the fusion provided higher accuracy reconstruction results than the individual fitting results. Fitting or fusion with the design model provided higher accuracy than that with a non-design model.

The error of reconstructed surfaces from their design models are usually characterised by parameters. There are many well-designed metrics for reconstruction error evaluation, such as root-mean-squared error (RMSE) and peak-to-valley values widely used in surface quality assessment [35, 36], universal quality index and structural similarity index for image comparison used in image processing [12], and diverse application specific parameters [37]. Considering that the popularity and robustness, RMSE is used in this study as the sole metric for the reconstruction error evaluation.

Since fusion error is usually smaller than individual fitting error, the reduced amount of fusion error from the smallest individual fitting error, *i.e.*

$$\Delta e_f = \min(e_{k\in K}) - e_f \tag{38}$$

exhibits the fusion effectiveness, where e_f and e_k are respectively the reconstruction RMSE of the fusion and individual fitting processes. A positive Δe_f indicates a successful fusion with reduced reconstruction error from the best individual fitting results. Figure 3c and 3f show that the fusion with the design model and the B-spline model respectively lead to the reduction of error in 32 % and 24 %. Statistics of the experimental results based on more reconstruction models are presented in Figure 4. In Figure 4a, the input datasets have the same level of noise. In Figure 4b, one input dataset has noise that is three times that of the other.

The presented five groups of results in Figure 4 correspond respectively to the reconstruction with the design model, the selected tensor product B-splines (6×6 parameters), a MBA (11×11 parameters, three levels) model and two optimised GP models. The former four groups employed the proposed weighted fusion methods; while the last group employed the GPRA fusion technique highlighted in previous research work [8]. Within each group of tests, the mean value of the reconstruction errors e_1, e_2, e_f and Δe_f are presented, and the standard deviation of Δe_f is also given to show the stability of the fusion performance.

The results demonstrate that the use of weighted fusion results in a higher accuracy reconstruction with high probability (see the positive Δe_f and its standard deviation bars). Different fitting models lead to different accuracy levels of reconstruction. Reconstruction with the design model can provide the highest reconstruction accuracy. Among the substitute models, GP non-parametric reconstruction shows better performance over the others due to its high flexibility. As a comparison, however, the GPRA fusion failed with this example as it provided a negative mean Δe_f. Also, it should be noted that weighted fusion reduces the reconstruction error mostly when the input datasets have equivalent sample noise.

5.2. Turbine blade surfaces

Turbine blades used in aeroengines usually have tortile and thin-wall freeform designs which provide optimised aero-/thermo-dynamic performance [38]. Inaccuracies in the blade surface shape will cause energy loss or blade failure. Figure 5b shows a part of working area of a blade surface with a set of extracted sample points (yellow dots). A holistic measurement of the blade requires several scans at different areal locations (see Figure 5a). Hence, fusion of overlapped scanning data for a universal reconstruction is an important task in this holistic measuring process.

We followed the same procedure (*i.e.* sample size 500 and different Gaussian noise superimposed) as in section 4.1 to obtain two sets of input samples. Then the proposed weighted fusion methods above with different reconstruction models were applied for error analysis with respect to a designed CAD model. The tested turbine blade surface has no explicit design function. Hence, only substitute reconstruction models were tested. In this case, the selected B-spline model had 7×10 control parameters and the MBA model had 35×35 parameters (six levels), both of which provided the highest reconstruction accuracies. The WLS fusion, parametric and non-parametric weighted fusion were respectively applied to the B-spline, MBA and GP reconstruction models. With 200 runs, the reconstruction error of the individual fits and the weighted fusion are presented in Figure 6.

Figure 6 demonstrates the same conclusion with the F-theta lens measurement, *i.e.* weighted fusion methods can provide higher accuracy reconstruction results over individual fits with high probability (see the standard deviation bars of Δe_f). In particular, weighted fusion has relatively stable performance on accuracy improvement

when input datasets have equivalent noise. When input sets have non-equivalent levels of noise, the effectiveness of weighted fusion degenerates.

Among all the substitute reconstruction models, non-parametric individual fits and fusion with GP models usually provides the best reconstructed surfaces when compared to parametric models, such as B-splines and MBAs. The GPRA fusion failed again in this example as it provided a negative mean Δe_f.

5.3. A freeform artefact

In this section, a freeform artefact with designed high accessibility [39] is used for experimental verification. Two sets of measuring results were sequentially obtained by using a structured light scanner (SL) and a Zeiss contact coordinate measuring machine (CMM) [39], both with 9635 sample points at the same positions. For visual convenience, the freeform artefact and the SL data are presented in Figure 7. Because the CMM has higher measuring precision than the SL scanner, the CMM data is regarded as the reference in the later error analysis. In other words, the measuring uncertainty of the SL data is much larger than that of the CMM data, *i.e.* $\sigma_{SL} \gg \sigma_{CMM}$.

The GPRA method has been shown to be a successful approach to fuse the 9635 SL sample points with randomly selected 100 CMM sample points [8]. In consideration of computing time and comparability with the former research in [8], the proposed weighted fusion methods were tested here by using simple randomly selected 100 SL points and 100 CMM points as the reconstruction input. With 200 tests for each fusion method, the statistics of the RMS individual fitting errors (e_1 for SL data and e_2 for CMM data), fusion error e_f and the reduction of the fusion error Δe_f were shown as in Figure 8. The selected B-spline model had 5×8 control parameters and the MBA model had 67×67 parameters (seven levels), both of which provided the highest reconstruction accuracies. In the weighting process with B-splines, MBA and GP models, the inverse of squared individual fitting errors $1/e_k^2$ were used as the weights instead of inverse of the squared measuring noise $1/\sigma_k^2$ [*], as described in section 1.

The results in this example show that the weighted fusion methods produce slight improvements in reconstruction accuracy, which coincides with the conclusions in the two previous examples and the theoretical argument in equation (25), *i.e.* when input datasets have unequal noise levels, the effectiveness of the weighted fusion degenerates. In this extreme case with $\sigma_1 \gg \sigma_2$, weighted fusion methods are nearly invalid in contrast to the GPRA fusion, which provided the highest accuracy improvement to high probability (see the standard deviation bars). By analysing the performance trend of the weighted fusion and the GPRA fusion under different input noise (equal noise in Figure 4a and Figure 6a, threefold difference of noise in Figure 4b and Figure 6b, and extreme difference of noise in Figure 8), it can be summarised that the weighted fusion and the GPRA fusion have complementary performance. Weighted fusion effectively works when input datasets have equal noise levels; while the GPRA fusion performs best when input datasets have significantly different levels (usually over threefold) of measurement noise.

[*] Because $\sigma_1 >> \sigma_2$, weighting with the inverse of squared noise as proposed will cause the fusion results same as the set 2 (the CMM set) fitting results, *i.e.* $e_f = e_2$ which means no accuracy improvement with fusion. In the situation when sample size is small and fitting (and interpolation) error dominates the reconstruction error over sampling noise, replacement of the fusion weights by the inverse of squared RMS fitting errors of individual sets lead to better fusion results.

Among the weighted fusion tests (with the B-spline, MBA and GP models), the MBA models with the parametric weighting have the best reconstruction accuracy in this example. Combined with the results in the previous two examples, it is found that different reconstruction models have their advantages for different surfaces. Looking for a universal model flexible enough for all surfaces is still a challenge in metrology. The B-splines model produces much larger reconstruction error and standard deviations because the limited number (100) sample points may easily lead to poor conditioning [17] of the design matrix with forty (5×8) control parameters.

Because a high-precision instrument usually has slow measuring speed (such as CMMs), the obtained sample size of a high-precision instrument is usually smaller than a low-precision instrument within the same sampling duration. A further test was carried out to verify the performance of weighted fusion by given two inputs with different sample sizes. 1000 SL (low-precision) sampling points were randomly selected for reconstruction and fusion with a randomly-selected 100 CMM (high-precision) points to imitate practical fusion cases as in [8, 40]. For 200 random tests, the statistics of the reconstruction errors were shown in Figure 9.

Weighted fusion in this example presented nearly no improvement in reconstruction accuracy. By observing the magnified standard deviation bars of the error reduction from fusion, it can be shown that the desired error reduction was unstable. By comparing with Figure 8, the performance of the weighted fusion was found to further degenerate, which coincides with theoretical argument in equation (26), *i.e.* when the sample sizes of the input datasets are different, the effectiveness of weighted fusion degenerates. In contrast, in this extreme case, the GPRA fusion presented good effectiveness and stability on accuracy improvement, which verified the conclusions in previous research work [8].

6. Conclusions

This study summarised four types of weighted fusion methods by giving the theoretical foundations and the analysis of uncertainty propagation. In addition, the advantages of WLS fusion over the Kalman filters were analysed following which, several advanced reconstruction models which are compatible with the weighted fusion methods were introduced. Lastly, three case studies on different measuring specimens were carried out to verify the performance of the proposed methods.

All the theoretical and experimental results demonstrated the effectiveness of weighted fusion for accuracy improvement in surface reconstruction. In particular, weighted fusion methods behave best when input datasets have the same sample sizes and the same levels of noise, in which case residual approximation-based fusion methods such as the GPRA have no effect. The effectiveness of weighted fusion degenerates when the noise levels and sample sizes of input datasets are very different (over threefold). In extreme situations when the noise level of one dataset is far smaller than the other and if the low-precision dataset has a much larger sample size, weighted fusion performed effectively no function; while in this situation, the GPRA fusion performed a good reconstruction with stable improvement of accuracy. Figure 10 presents the performance statistics of the two fusion frameworks in which the fusion error reduction percentages were summarised from the experiments in section 4. The experimental conclusions coincide with the theoretical arguments and previous research work.

In summary, weighted fusion methods combined with different reconstruction models form a good supplement to residual approximation-based fusion methods. The two types of fusion frameworks effectively work for different fusion scenarios as described elsewhere [41].

Among the weighted fusion tests analysed above, it was found that different reconstruction models lead to different reconstruction errors. However, no universal model can be flexible enough to effectively work for all types of surfaces. Looking for universally applicable advanced fitting models is still an important task in surface reconstruction and weighted fusion.

Acknowledgements

This work was supported by the Huazhong University of Science and Technology and the National Natural Science Foundation of China (No. 51275192).

References

[1] Weckenmann, A., et al., Multisensor data fusion in dimensional metrology. CIRP Annals - Manufacturing Technology, 2009. **58**(2): p. 701-721.
[2] Bruzzone, A.A.G., et al., Advances in engineered surfaces for functional performance. CIRP Annals - Manufacturing Technology, 2008. **57**(2): p. 750-769.
[3] Jiang, X.J. and D.J. Whitehouse, Technological shifts in surface metrology. CIRP Annals - Manufacturing Technology, 2012. **61**(2): p. 815-836.
[4] Leach, R., Optical measurement of surface topography. 2011: Springer.
[5] Wang, J., R.K. Leach, and X. Jiang, Review of the mathematical foundations of data fusion techniques in surface metrology. Surface Topography: Metrology and Properties, 2015. **3**(2): p. 023001.
[6] Ez, Y., et al., A Qualitative Review on 3D Coarse Registration Methods. Acm Computing Surveys, 2015. **47**(3): p. 1-36.
[7] Rusinkiewicz, S. and M. Levoy. Efficient variants of the ICP algorithm. in 3dim. 2001.
[8] Colosimo, B.M., M. Pacella, and N. Senin, Multisensor data fusion via Gaussian process models for dimensional and geometric verification. Precision Engineering (In review), 2014.
[9] Qian, P.Z.G. and C.F.J. Wu, Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments. Technometrics, 2008. **50**(2): p. 192-204.
[10] Seungyong, L., G. Wolberg, and S. Sung-Yong, Scattered data interpolation with multilevel B-splines. Visualization and Computer Graphics, IEEE Transactions on, 1997. **3**(3): p. 228-244.
[11] Dıez, D.C., Adaptive Scattered Data Fitting with Tensor Product Spline–Wavelets. 2005, Universität Bonn.
[12] Ramasamy, S.K. and J. Raja, Performance evaluation of multi-scale data fusion methods for surface metrology domain. Journal of Manufacturing Systems, 2013. **32**(4): p. 514-522.
[13] Strutz, T., Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond. 2010: Vieweg and Teubner. 244.
[14] Ramasamy, S.K., Multi-scale data fusion for surface metrology. 2011, The University of North Carolina at Charlotte.
[15] Raol, J.R., Multi-Sensor Data Fusion with MATLAB. 2009: CRC Press.
[16] Ramasamy, S.K., J. Raja, and B.D. Boudreau, Data Fusion Strategy for Multiscale Surface Measurements. Journal of Micro and Nano-Manufacturing, 2013. **1**(1): p. 011004-011004.
[17] Barker, R.M., et al., Best Practice Guide No. 4 Software Support for Metrology: Discrete modelling and experimental data analysis, in Technical report, National Physical Laboratory. 2004: Teddington, UK.
[18] Golub, G.H. and C.F. Van Loan, Matrix Computations. 1996: Johns Hopkins University Press.
[19] Kalman, R.E., A new approach to linear filtering and prediction problems. Journal of basic Engineering, 1960. **82**(1): p. 35-45.
[20] Huang, Y. and X. Qian, A Dynamic Sensing-and-Modeling Approach to Three-Dimensional Point- and Area-Sensor Integration. Journal of Manufacturing Science and Engineering, 2006. **129**(3): p. 623-635.
[21] Szeliski, R. Probabilistic modeling of surfaces. 1991.

[22] Wang, Y.F. and J.-F. Wang, On 3D model construction by fusing heterogeneous sensor data. CVGIP: Image Underst., 1994. **60**(2): p. 210-229.
[23] Thacker, N.A. and A.J. Lacey, Tutorial: The likelihood interpretation of the Kalman filter. 2006, University of Manchester: Manchester, UK.
[24] Brown, R.G. and P.Y.C. Hwang, Introduction to Random Signals and Applied Kalman Filtering with MATLAB Exercises, 4th Edition. 2012: Wiley Global Education.
[25] Welch, G. and G. Bishop, An Introduction to the Kalman Filter. 1995, University of North Carolina at Chapel Hill.
[26] Pagani, L., Data fusion for process Optimization and surface reconstruction, in Department of Mechanical Engineering. 2005, Politechnic University of Milan: Milan.
[27] Rasmussen, C.E. and C. Williams, Gaussian Processes for Machine Learning. 2006: the MIT Press.
[28] Xia, H., Y. Ding, and J. Wang, Gaussian process method for form error assessment using coordinate measurements. IIE Transactions, 2008. **40**(10): p. 931-946.
[29] Evans, C.J. and J.B. Bryan, "Structured", "Textured" or "Engineered" Surfaces. CIRP Annals - Manufacturing Technology, 1999. **48**(2): p. 541-556.
[30] Piegl, L.A. and W. Tiller, The NURBS Book. 1997: Springer. 646.
[31] de Boor, C., A Practical Guide to Splines. Applied Mathematical Sciences. 2001: Springer.
[32] Hjelle, **Ø.**, Approximation of scattered data with multilevel b splines. 2001.
[33] Isaaks, E.H. and M.R. Srivastava, An Introduction To Applied Geostatistics. Technometrics, 1991. **33**(4).
[34] Li, H., et al., Inspection Technology for F-Theta Lens: Practices and Discussions. Key Engineering Materials, 2008. **364**: p. 1191-1196.
[35] ISO 1101:2012, Geometrical product specifications (GPS) -- Geometrical tolerancing -- Tolerances of form, orientation, location and run-out. 2012, International Standard Organisation: Geneva.
[36] ISO 25178-2:, Geometrical product specification (GPS) - Surface texture: areal - part 2: terms, definitions and surface texture parameters. 2012.
[37] Savio, E., L. De Chiffre, and R. Schmitt, Metrology of freeform shaped parts. CIRP Annals - Manufacturing Technology, 2007. **56**(2): p. 810-835.
[38] Li, W.-l., L.-p. Zhou, and S.-J. Yan, A case study of blade inspection based on optical scanning method. International Journal of Production Research, 2015. **53**(7): p. 2165-2178.
[39] Cavallaro, M., S. Petrò, and G. Moroni, Performance evaluation of non contact measuring systems considering bias, in Innovative Developments in Design and Manufacturing. 2009, CRC Press.
[40] Kennedy, M.C. and A. O'Hagan, Predicting the output from a complex computer code when fast approximations are available. Biometrika, 2000. **87**(1): p. 1-13.
[41] Wang, J., R.K. Leach, and X. Jiang, NPL Report ENG 54 - Review of data fusion for surface topography measurement. 2015: UK.

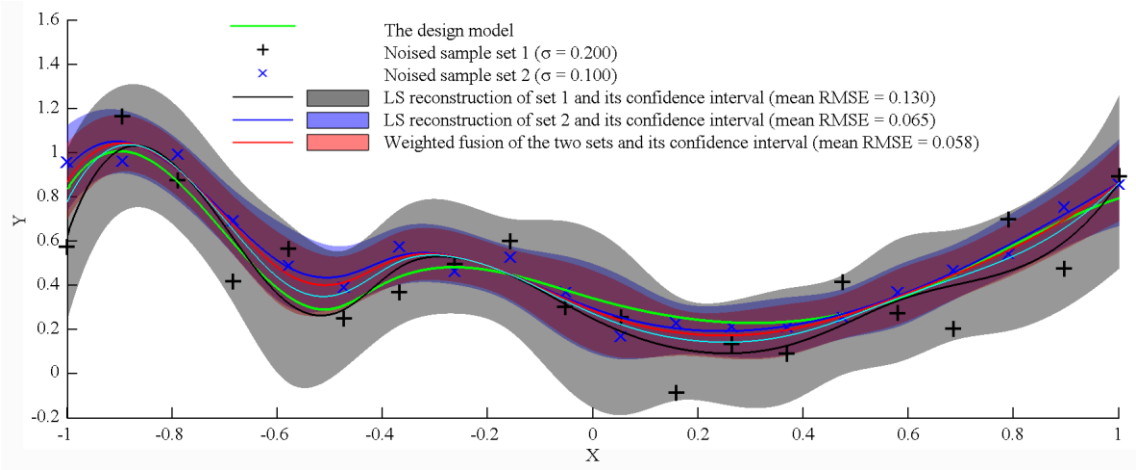
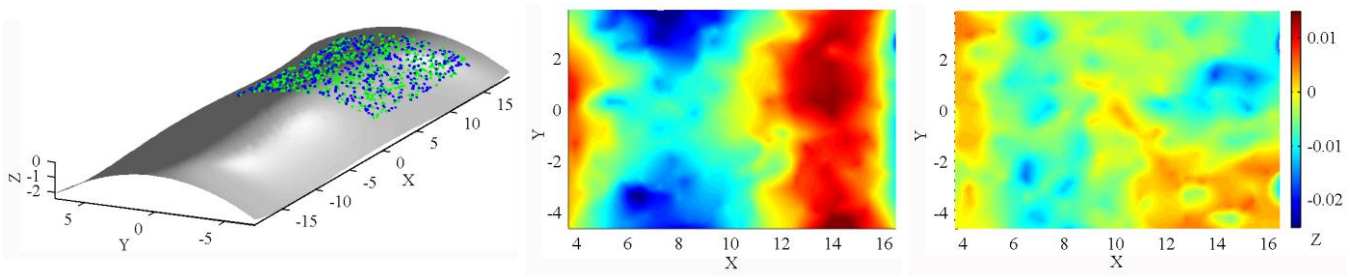
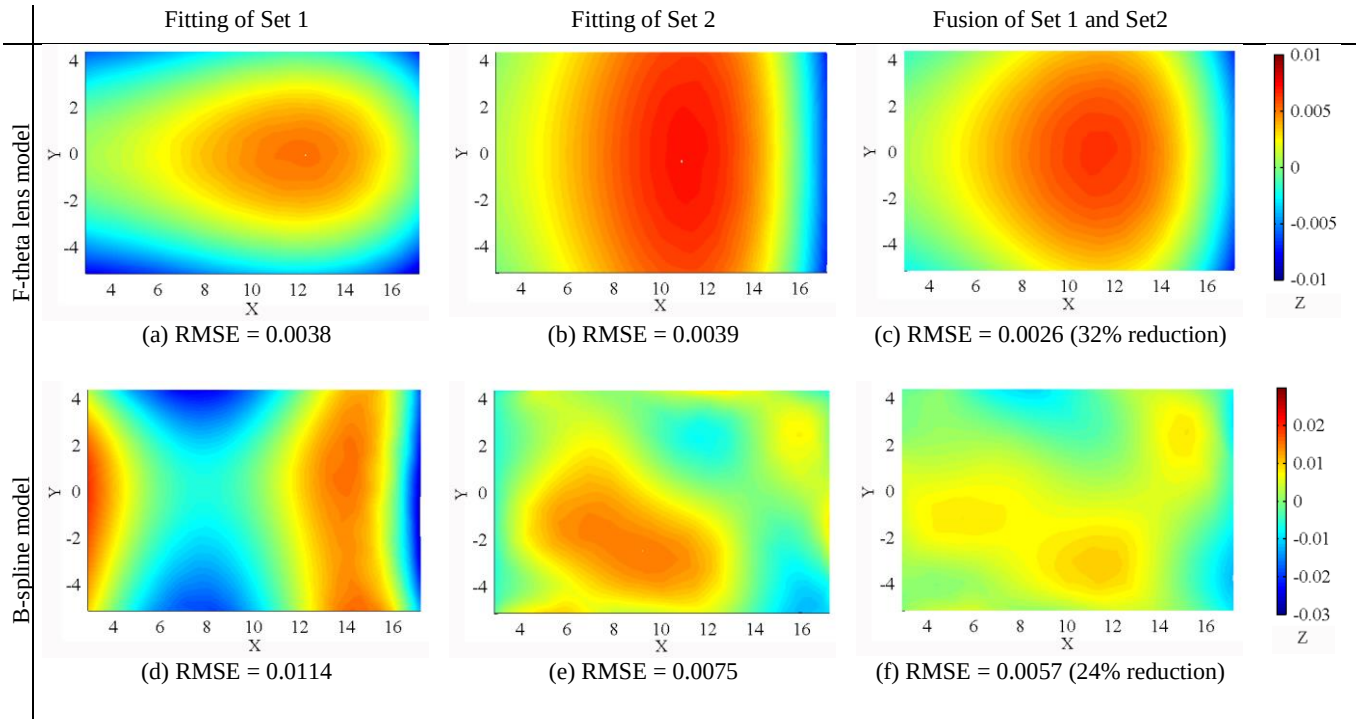


Figure 1. The performance of the WLS fusion and individual least-squares reconstruction techniques.



(a) Set 1 (blue), Set 2 (green) and the CAD model (b) An error map of Set 1 (c) An error map of Set 2

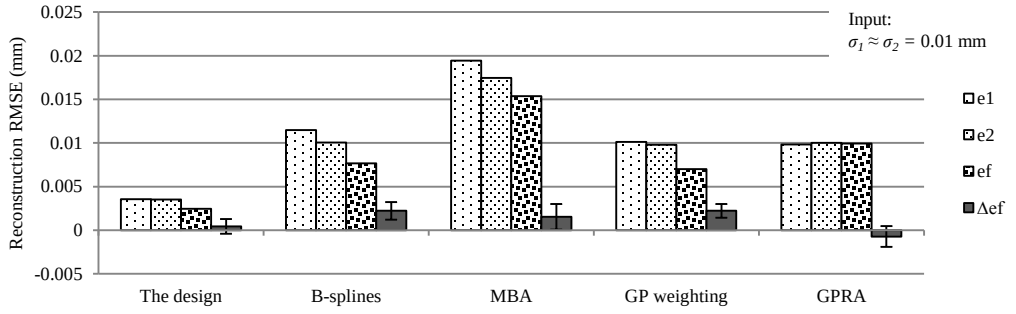
Figure 2. Two example sets of noised sample data extracted from a F-theta lens model (a) and their error maps (b and c) (unit: mm).



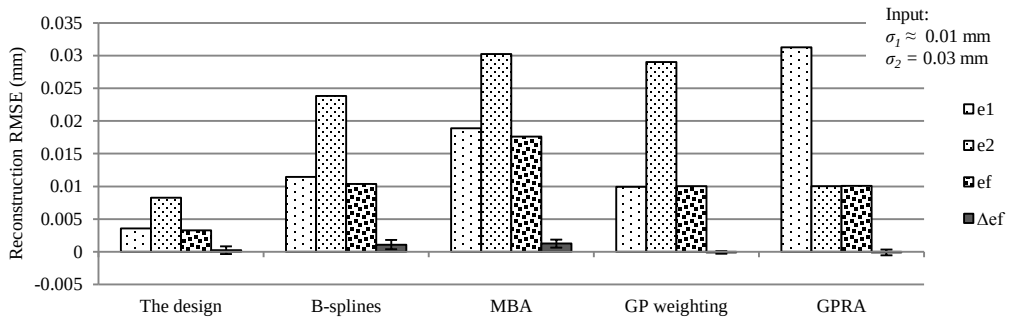
(a) RMSE = 0.0038 (b) RMSE = 0.0039 (c) RMSE = 0.0026 (32% reduction)

(d) RMSE = 0.0114 (e) RMSE = 0.0075 (f) RMSE = 0.0057 (24% reduction)

Figure 3. An example of the reconstruction error maps from individual fitting and the WLS fusion (unit: mm).

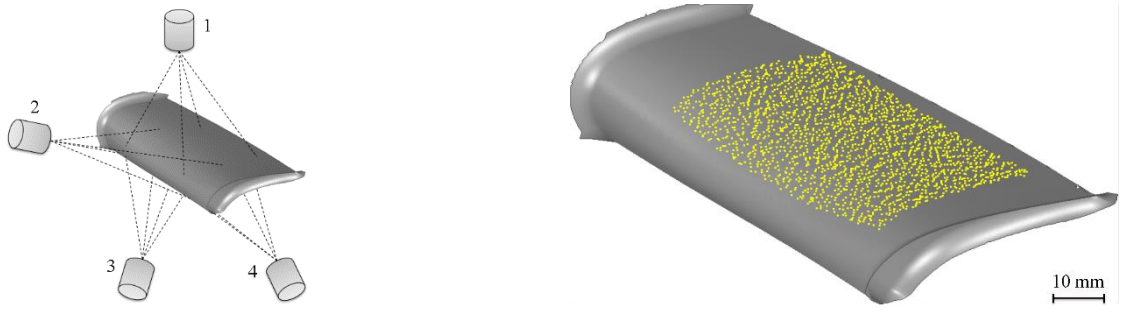


(a) For noise-equivalent input



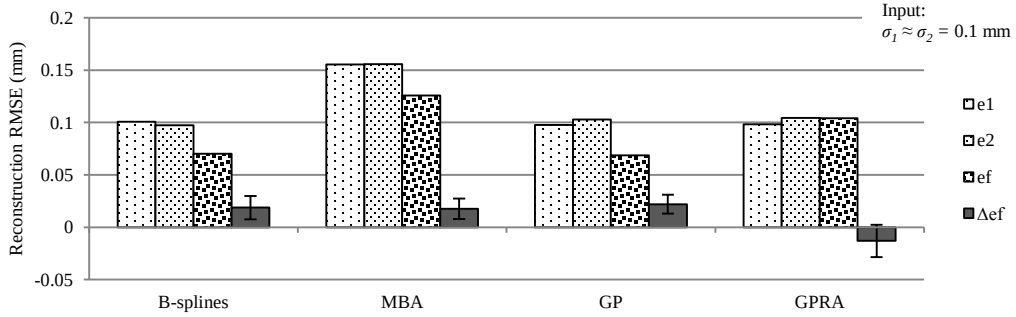
(b) For noise-non-equivalent input

Figure 4. Statistics of reconstruction errors for the F-theta lens measurement with simulated input noise.

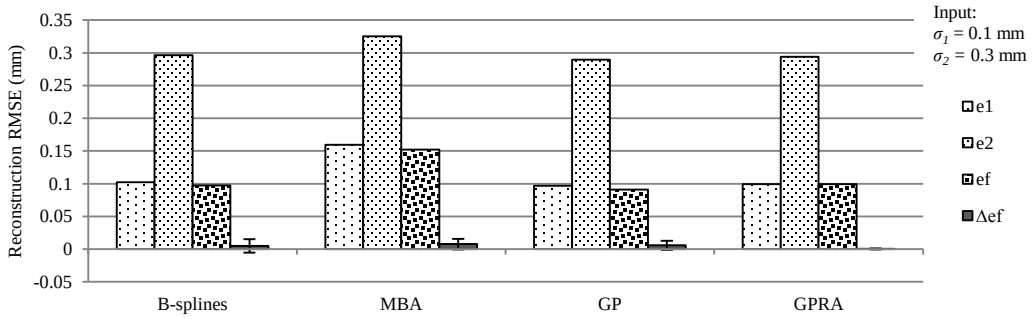


(a) **(b)**

Figure 5. Holistic measurement of a turbine blade with several areal scans (a) and a set of sample points (b).

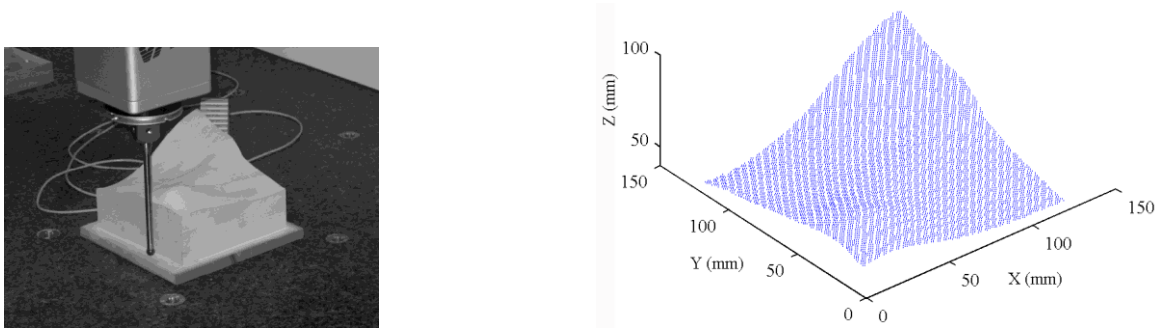


(a) With noise-equivalent input.



(b) With noise-non-equivalent input.

Figure 6. Statistics of the reconstruction errors for the turbine blade measurement with simulated input noise.



(a) **(b)**

Figure 7. The freeform artefact and the 9635 sample points from a structured light scanner.

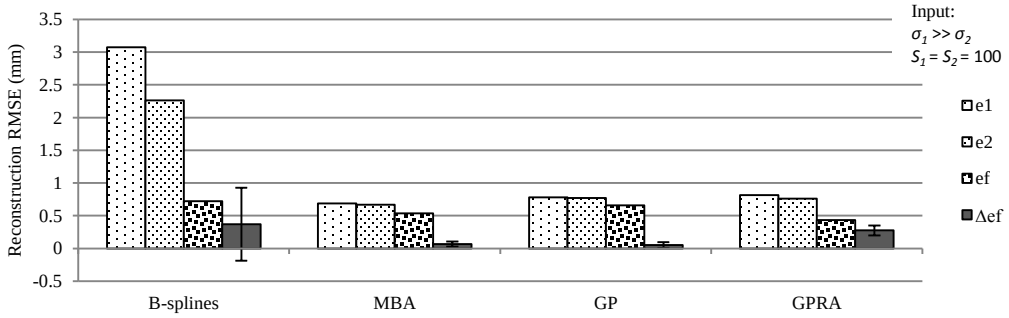


Figure 8. Statistics of the reconstruction errors for the freeform artefact data from a SL and CMM with the same sample size.

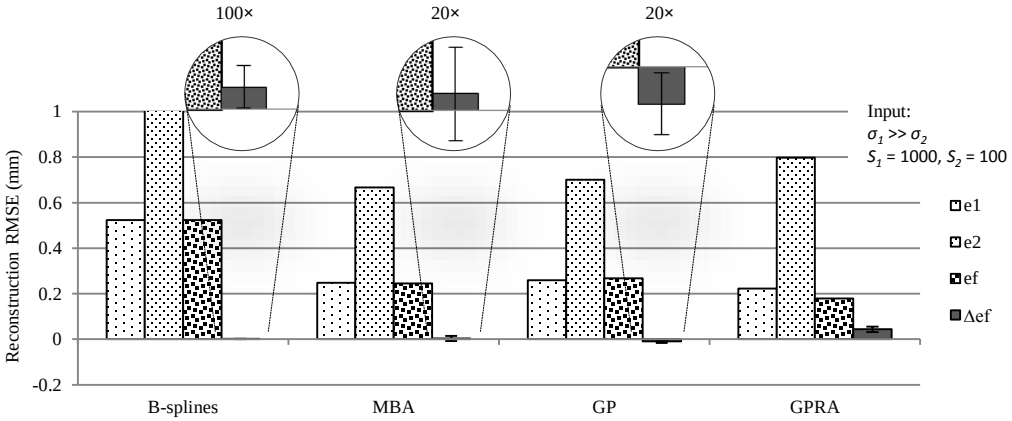


Figure 9. Statistics of the reconstruction errors for the freeform artefact data from a SL and CMM with different sample sizes.

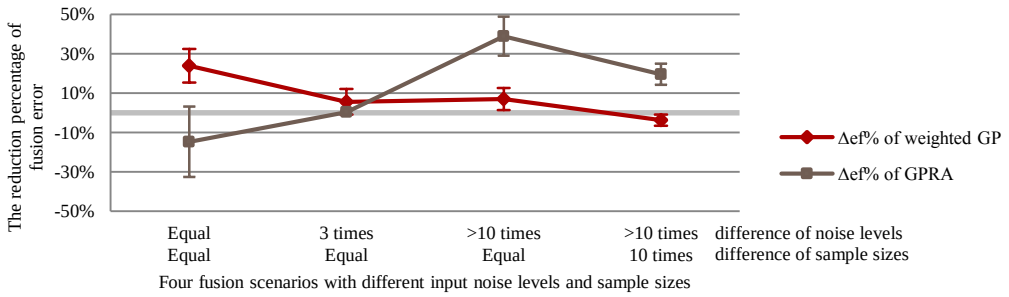


Figure 10. The performance of weighted and RA fusion with GP models for different fusion scenarios.