

Article

A Multisensor Data Fusion Method Based on Gaussian Process Model for Precision Measurement of Complex Surfaces

Ji Ding ^{1,2}, Qiang Liu ^{1,2,*}, Mingxuan Bai ^{1,2} and Pengpeng Sun ^{1,2}

¹ School of Mechanical Engineering & Automation, Beihang University, Beijing 100083, China; dingji@buaa.edu.cn (J.D.); baimx@buaa.edu.cn (M.B.); speng@buaa.edu.cn (P.S.)

² Beijing Engineering Technological Research Center of High-efficient & Green CNC Machining Process and Equipment, Beijing 100083, China

* Correspondence: qliusmea@buaa.edu.cn; Tel.: +86-10-8231-6298

Received: 28 November 2019; Accepted: 31 December 2019; Published: 3 January 2020



Abstract: As multisensor measurement technology is rapidly applied in industrial production, one key issue is the data fusion procedure by combining several datasets from multiple sensors to obtain the overall geometric measurement. In this paper, a multisensor data fusion method based on a Gaussian process model is proposed for complex surface measurements. A robust surface registration method based on the adaptive distance function is firstly used to unify the coordinate systems of different measurement datasets. By introducing an adjustment model, the residuals between several independent datasets from different sensors are then approximated to construct a Gaussian process model-based data fusion system. The proposed method is verified through both simulation verification and actual experiments, indicating that the proposed method can fuse multisensor measurement datasets with better fusion accuracy and faster computational efficiency compared to the existing method.

Keywords: data fusion; data registration; adaptive distance function; complex surface measurement; Gaussian process model

1. Introduction

With the development of advanced manufacturing technology, many complex surfaces such as freeform surfaces and structured surfaces can be machined with high precision [1,2]. These surfaces require a complete 3D characterization, with a large measurement range, high resolution and precision, and high measurement efficiency, which poses a challenge to current measurement technology [3,4].

Currently, measuring instruments have been developed to meet specific measuring requirements. Although these instruments have their advantages, no single instrument can simultaneously meet the high requirements in terms of precision, efficiency, resolution, and measuring range [5–7]. For example, coordinate measuring machines (CMMs) usually have high accuracy, but their measurement efficiency is low. Non-contact measuring devices, e.g., structured light scanners, could generate dense measurement data efficiently, which can capture the overall shape of the product well. Yet their measurement accuracy is much lower than the CMMs. Therefore, the combination of multiple measurement sensors could be a better solution to solve complex measurement tasks, which maximizes the advantages of individual measurement sensors [8–10]. By integrating multiple sensors, these instruments can provide multiple measurement results according to the user's choice and adapt to the complexity of product geometry measurement. For example, the contact-type CMMs can be optionally equipped with corresponding non-contact measuring sensors. Data fusion is a key problem following multisensor measurement, which can be used in quality inspection, surface reconstruction, and other fields. The aim of data

fusion is to integrate multiple datasets from different sources into a unified and improved output. The quality of this output data depends heavily on the data fusion method. Multisensor datasets are usually in different coordinate systems with different measurement ranges, accuracy, and resolution, which brings great challenges to multisensor data fusion [11,12].

In most situations, the multisensor data fusion process is divided into three parts: pre-processing, data registration, and data fusion [4]. Data registration and fusion is the key to realize the whole process. In the data registration, multisensor datasets in different reference coordinate systems are unified into a common coordinate system by finding a rigid body transformation matrix. The Iterative Closest Point (ICP) algorithm proposed by Besl and McKay [13] is the most widely used surface registration algorithm. Its key contribution is to put forward an iterative registration method. The method firstly finds the closest point, and then establishes an optimization model to minimize the error function, thus obtaining updated transformation parameters. When the iteration number reaches a given value, or the error reaches a threshold, the iteration process stops. However, the point-point distance function adopted by ICP has linear convergence according to the literature [14], so ICP and related algorithms have a slow convergence rate. The squared distance minimization (SDM) and tangent-squared distance minimization (TDM) are introduced, which both are the point-tangent distance functions and converge faster than ICP [14–16]. However, they need a good initial value of rigid body transformation, otherwise, it is easy for the registration to fail. So, the adaptive distance function (ADF), which is a point-surface distance function, is proposed [17,18]. The ADF maintains a fast convergence rate and has a low requirement for the initial value and good convergence stability [17].

The fusion step mainly deals with the multisensor datasets in a common reference coordinate system after registration, and research on this issue is still inadequate. Jamshidi et al. [19] proposed a high-resolution and low-resolution data fusion method, which realized data fusion by eliminating outliers and gridding all data points after the data registration. However, such methods only substitute the data from different sources locally and cannot improve the accuracy of the overall data. Therefore, the data fusion methods based on modern statistical theory have received much attention [20–23]. The core idea of such methods is to use statistical methods to model the measurement datasets obtained from different sensors, and then provide the final prediction of each position. Senin et al. [24] compared three multisensor point set augmentation methods, including linear interpolation, locally weighted scatterplot smoothing (LOWESS), and a Gaussian process (GP). Ren et al. [25] presented a weighted least squares based fusion method for precision measurement of free surfaces, which used B-spline surface to fit a linear surface model of different sensor datasets. This method depends on the accuracy of the surface model fitting, and is not suitable for several complex surfaces. Xia et al. [26] introduced a fusion algorithm for high and low accuracy point cloud, which is based on the Bayesian Hierarchical (BH) model. This method could establish a multi-layer Bayesian parameter estimation model for corresponding points of different datasets, and obtain the optimal estimation value after fusion by solving the BH model. Colosimo et al. [27–29] proposed a multisensor data fusion method based on the GP model. Reconstruction models of high-accuracy datasets and low-accuracy datasets were established by using the GP model to obtain the optimized fusion dataset. However, these methods rarely consider the effect of different registration methods on fusion quality.

In this paper, a multisensor data fusion algorithm based on GP is presented for precision measurement of freeform surfaces. It can be applied to the data fusion process of 3D point clouds obtained from the same complex surface by multiple sensors. The proposed algorithm utilizes a robust registration algorithm, which is based on the ADF, to unify different coordinate systems, and then uses the GP model to combine multiple datasets from different sensors. The main content of the rest in this paper is divided into three sections. The optimization model and its solution of the ADF based robust registration method is proposed in the second section, and then a GP-based data fusion system is established to generate a surface fusion dataset. The third section gives the verification results of the proposed method under simulation and the actual measurements. The conclusions are presented in the fourth section.

2. The Multisensor Data Fusion Method

2.1. Summary of the Multisensor Fusion Method

Datasets measured by multisensor systems generally have different resolutions and uncertainties, and they are in different coordinate systems. The main goal of this study is to improve local trueness of the dataset (reduce local bias) by integrating the dataset obtained by the structural light scanner or similar high-speed measurement technology (high point density, low accuracy, low cost) with the dataset obtained by the “more real” point obtained by a CMM or similar technology (slow sampling speed, high accuracy, high price). Therefore, it is assumed that there are two datasets available in this data fusion method:

- one type of dataset with high accuracy, low density, which is generated by CMM or high-precision microscope. This high-accuracy dataset is called the HA set for short.
- another type of dataset with low accuracy, high density, generated by the structured light scanner, line scanner, or similar technology. This low-accuracy dataset is referred to as the LA set.

In order to solve the key problems in data fusion, this paper puts forward a method that divides the whole fusion process into preprocessing, data registration, and data fusion. The flow diagram of the proposed data fusion method is illustrated in Figure 1. The first step is to preprocess the HA and LA data, including data unification, outliers removal, and other operations. The second step is to introduce a robust registration method based on the ADF to unify the coordinate systems of HA and LA datasets. The third step is to establish a GP based multi-layer fusion model to realize the fusion process, in which HA data points are taken as reference points so as to correct the local deviation of the LA dataset.

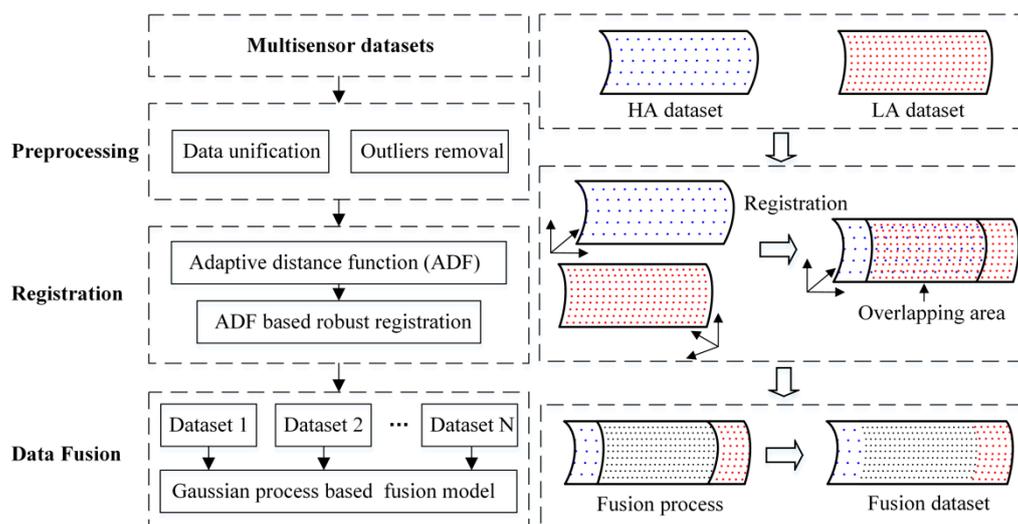


Figure 1. The flow diagram of the multisensor data fusion method.

2.2. ADF-Based Robust Data Registration

In this study, the ADF is used to describe the basis error metric, and the M-estimation is used to improve the registration robustness and eliminate noise effects.

As is shown in Figure 2, we refer to the point set $P = \{p_i : i = 1, \dots, n\}$ as the LA dataset and the point set $Q = \{q_j : j = 1, \dots, m\}$ as the HA dataset. The robust registration of point clouds in \mathbb{R}^3 is to find a rigid body transformation $g = (R, t) \in \mathbb{R}^3$ (including the rotation matrix $R \in \mathbb{R}^3$ and translation matrix $t \in \mathbb{R}^3$) that fits the mobile point set P to the reference point set Q . Point p_{i+} is generated from point p_i through rigid body transformation g . Point q_j is the nearest to point p_i , and the tangent plane at point q_j is called the plane T_j . The point q_{j+} at point set Q is obtained in the line $p_{i+}o_j$ where the

point o_j is the center of the curvature circle at q_j . The normal vector to the surface point set Q at point q_j is denoted by the notation n_j .

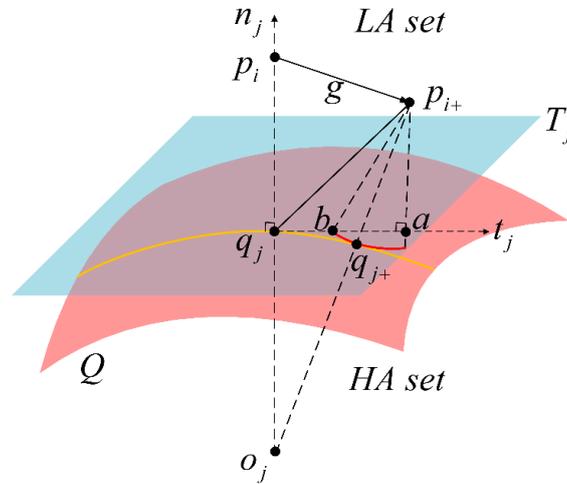


Figure 2. The distance $p_{i+}q_{j+}$ between point p_{i+} and point set Q .

It is notable that $\|p_{i+}q_{j+}\|$ is an approximate expression of the shortest distance between point p_{i+} and surface Q . However, the distance $\|p_{i+}q_{j+}\|$ depends on the curvature change in the region around point q_j , and the direct calculation of $\|p_{i+}q_{j+}\|$ is complicated and susceptible to noise. To calculate the transformation g , a new expression for the shortest distance is studied. Point a is obtained by projecting point p_{i+} onto plane T_j . The arc bq_{j+} is generated by the center p_{i+} and the radius $\|p_{i+}q_{j+}\|$, then $\|p_{i+}q_{j+}\| = \|p_{i+}b\| = \sqrt{\|p_{i+}a\|^2 + \|ab\|^2}$. The tangent distance $\|ab\|$ is calculated by $\|ab\|^2 = \mu\|aq_j\|^2$ ($\mu \in [0, 1]$), where μ denotes the modified coefficient. Then, the ADF is defined as

$$\begin{aligned} d_{\text{ADF}}(g, Q) &= \|p_{i+}b\| = \sqrt{\|p_{i+}a\|^2 + \mu\|aq_j\|^2} \\ &= \sqrt{\|n_j^T(p_{i+} - q_j)\|^2 + \mu\|t_j^T(p_{i+} - q_j)\|^2}, \mu \in [0, 1] \end{aligned} \quad (1)$$

where n_j represents the normal vector, t_j represents a vector parallel to the tangent plane T_j .

As the modified coefficient μ goes to 0, the expression $d_{\text{ADF}}(g, Q)$ becomes $d_{\text{TDF}}(g, Q)$, which turns into the point-tangent distance function $\|p_{i+}a\|$ that is applied to the TDM algorithm. As the modified coefficient μ approaches 1, the expression of ADF becomes $d_{\text{PDF}}(g, Q)$, which corresponds to the point-point distance function $\|p_{i+}q_j\|$ used in the ICP algorithm. Thus, $d_{\text{TDF}}(g, Q)$ and $d_{\text{PDF}}(g, Q)$ both are the special forms of the adaptive function distance (ADF). The ADF selects the appropriate modified coefficient to calculate the tangent distance, so as to reflect the curvature characteristics. In general, the ADF can describe the shortest point-surface distance more accurately than TDM and ICP [17,18].

In order to calculate the rigid body transformation g , an objective function for ADF-based robust registration is established:

$$f(g) = \sum_{i=1}^n \rho(d_{\text{ADF}}(g(p_i), q_i)) \quad (2)$$

where ρ denotes an M-estimate function, d_{ADF} represents the ADF, and the corresponding pairs (p_i, q_i) are selected from the registration point sets. The M-estimate function ρ has the property of reducing the impact of the data points with large deviation. The derivative of the M-estimate function ρ is written as $\psi(e) = \rho'(e)$, and a weight function ω is represented as $\omega(e) = \frac{\psi(e)}{e}$.

The bisquare estimator is a typical example of the M-estimators. The bisquare estimator has the strongest suppression ability to the high impact of noisy data. The weight function for the Bisquare estimator is as follow:

$$\omega(e) = \begin{cases} \left[1 - \left(\frac{e}{k}\right)^2\right]^2 & \text{for } |e| \leq k \\ 0 & \text{for } |e| > k \end{cases} \quad (3)$$

In order to solve the objective function for ADF-based robust registration, the objective function with the Bisquare estimator can be transformed into iterative reweighted least squares (IRLS) minimization, and then a new objective function IRLS-ADF is expressed as:

$$f(g) = \sum_{i=1}^n \omega_i d_{ADF}^2(g(p_i), q_i) \quad (4)$$

with the weight function ω for the Bisquare estimator being a function of the corresponding pairs (p_i, q_i) .

To minimize the objective function for robust registration, a nonlinear optimization model is established:

$$\min f(g) = \sum_{i=1}^n \omega_i d_{ADF}^2(g(p_i), q_i) \text{ s.t. } q_i = \underset{q \in Q}{\text{argmin}} \|p_i - q\| \quad (5)$$

The transformation can be computed by solving the linear equations, similar to [15].

Define $v = [V_x, V_y, V_z]^T$ as the translation vector, and $\omega = [\delta x, \delta y, \delta z]^T$ as the rotation vector, then $\Delta x = p_{i+} - p_i = v + \omega \times p_i$. Define $\xi = [v, \omega]^T$, then

$$\begin{aligned} p_{i+} b &= d_{i1} + n_j^T (v + \omega \times p_i) = d_{i1} + A_i \xi \\ b q_j &= d_{i2} + t_j^T (v + \omega \times p_i) = d_{i2} + B_i \xi \end{aligned} \quad (6)$$

where A_i is a 1×6 matrix and $\xi = [v, \omega]^T$ is a 6×1 matrix.

Therefore, the objective function in (5) can be written as

$$\begin{aligned} \varepsilon(\xi) &: = \sum_i^n \left\{ \omega_i \left[d_{i1} + n_j^T (v + \omega \times p_i) \right]^2 + \omega_i \mu \left[d_{i2} + t_j^T (v + \omega \times p_i) \right]^2 \right\} \\ &= D + 2C^T \xi + \xi^T A \xi + \mu \xi^T B \xi \end{aligned} \quad (7)$$

where $A = \sum_i^n \omega_i A_i^T A_i$ and $B = \sum_i^n \omega_i B_i^T B_i$ are both 6×6 symmetric matrices, $C = \sum_i^n \omega_i d_i A_i^T$ is a column vector, and D is a scalar.

The parameter $\xi = [v, \omega]^T$ is obtained from the solution to the linear equations

$$(A + \mu B) \xi + C = 0 \quad (8)$$

Then, we can get the optimal transformation $g = (R, t)$ from

$$R = e^{\hat{\omega}}, t = v \quad (9)$$

where $\hat{\omega}$ is the anti-symmetric matrix of the rotation vector ω .

The flow chart of the IRLS-ADF method is given in Figure 3.

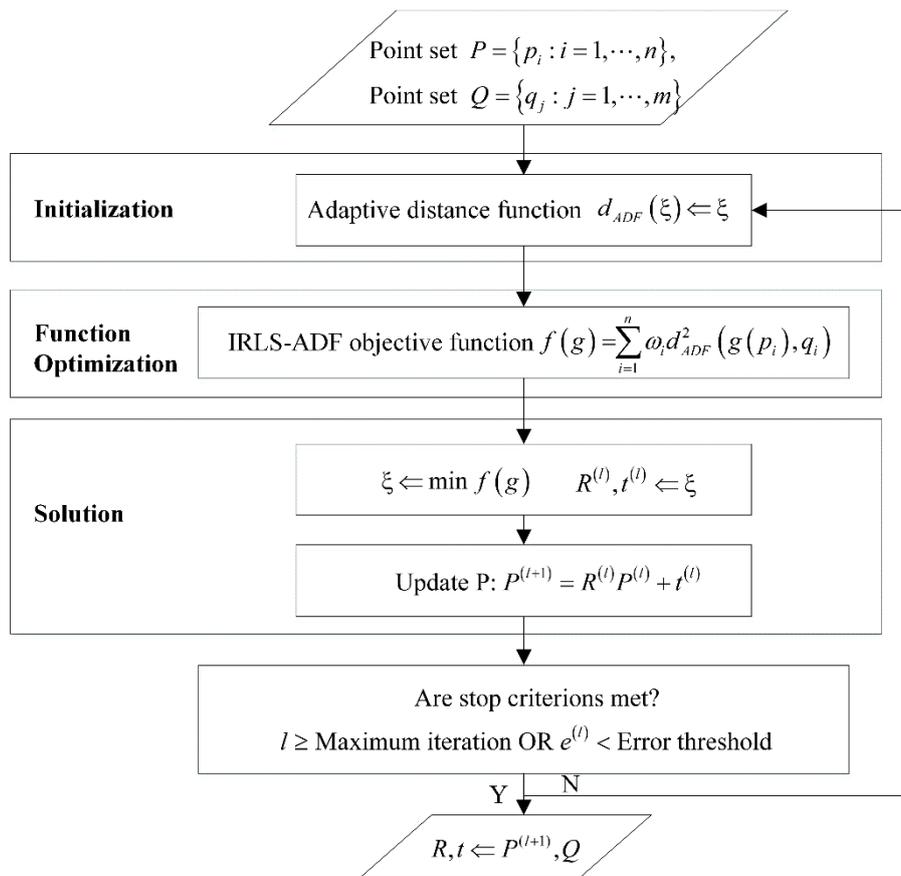


Figure 3. The flow chart of the IRLS-ADF method.

In order to comprehensively consider the convergence speed and stability of the IRLS-ADF method, the modified coefficient μ needs to be carefully selected [17,18]. When the iterative process begins, the optimization problem, which becomes a large residual problem, needs to select a large modified coefficient ($\mu_0 \in [0.5, 0.8]$) to keep a certain ratio of tangent distance in the objective function, thus guaranteeing the initial convergence stability. After a certain number of iterations, the optimization problem needs to select a relatively small modified coefficient μ to speed up the convergence rate, because it becomes a small residual problem.

It is proved that the IRLS-ADF method converges to the global minimum for the Bisquare estimator, when the appropriate transformation parameters are selected at the beginning [30]. In fact, the IRLS-ADF method is a kind of Levenberg–Marquardt method for solving the large residual problem, which has good convergence stability and quadratic convergence rate [17,30].

2.3. GP-Based Data Fusion Method

After registration, a registered LA dataset and HA dataset are integrated into an improved output by a multi-layer Bayesian data fusion method based on GP model. The fusion method mainly consists of two steps: (1) obtaining an alternative model of the LA dataset; (2) introducing the adjustment model to combine LA and HA datasets together.

Assume that the two datasets can be expressed by discrete functions of the type: $z(x_i, y_i)$, that is the z -coordinate of the i -th point is represented as a function of its position on the x, y plane. Then, the LA dataset is denoted as: $z_{LA}(x_i, y_i)$, where $(x_i, y_i) \in V_{LA} \subset \mathbb{R}^2, i = 1, 2, \dots, n_{LA}$ and similarly, the HA dataset is denoted as: $z_{HA}(x_i, y_i)$, where $(x_i, y_i) \in V_{HA} \subset \mathbb{R}^2, i = 1, 2, \dots, n_{HA}$. Generally, the points in LA and HA dataset will not have the exact same position on the plane, i.e., $V_{LA} \cap V_{HA} = \emptyset$.

The LA data $z_{LA}(x_i, y_i)$ can be defined as

$$z_{LA}(x_i, y_i) = f_{LA}(x_i, y_i) + \varepsilon_{LA} \quad (10)$$

where, ε_{LA} represents the error, and it is assumed to follow a normal distribution of 0 mean and $\sigma_{\varepsilon_{LA}}^2$ variance, i.e., $\varepsilon_{LA} \sim N(0, \sigma_{\varepsilon_{LA}}^2)$. In an ideal situation, $f_{LA}(x_i, y_i)$ represents the ideal value of the measured surface and ε_{LA} represents the measurement error of the LA dataset. In this paper, a GP model is considered to represent $f_{LA}(x_i, y_i)$. The GP model is a set of random variables, in which any finite number of random variables and their linear combinations obey the joint Gaussian distribution. Therefore, according to the GP model, it can be expressed as follows:

$$f_{LA}(x_i, y_i) \sim GP(m_{z_{LA}}(v_i), k_{z_{LA}}(v_i, v_j)) \quad (11)$$

where v_i represents (x_i, y_i) and similarly, v_j represents (x_j, y_j) , $m_{z_{LA}}(v_i) = E[z_{LA}(v_i)]$ is the mean function, $k_{z_{LA}}(v_i, v_j) = \text{cov}(z_{LA}(v_i), z_{LA}(v_j))$ is the covariance function. In this paper, the mean function can be expressed with a linear model:

$$m_{z_{LA}}(x_i, y_i) = \alpha_0 + \alpha_1 x_i + \alpha_2 y_i \quad (12)$$

In addition, the square exponential function is chosen to represent the covariance function of the GP model:

$$k_{z_{LA}}(v_i, v_j) = \sigma_{z_{LA}}^2 \exp\left(-\frac{\|v_i - v_j\|^2}{2l_{LA}^2}\right) \quad (13)$$

where $\|v_i - v_j\|$ represents the Euclidean distance between vector v_i and vector v_j , $\sigma_{z_{LA}}^2$ represents the constant variance of the GP model, and l_{LA} is the length parameter of the function.

In the model expressed by Equations (10)–(13), parameters $\{\alpha_0, \alpha_1, \alpha_2, \sigma_{z_{LA}}^2, l_{LA}, \sigma_{\varepsilon_{LA}}^2\}$ are unknown and need to be estimated by actual measurement data. Therefore, as long as the measurement data $z_{LA}(x_i, y_i)$ with $(x_i, y_i) \in V_{LA} \subset \mathbb{R}^2$, $i = 1, 2, \dots, n_{LA}$ is given, the new prediction value $\hat{z}_{LA}(x, y)$ can be computed at any new position (x, y) according to the mean and covariance functions after the completion of the parameter estimation.

However, the alternative model based on the LA data cannot correct the possible local bias in the LA dataset. Due to the high accuracy and unbiasedness of the HA dataset within the measurement range, we introduce the high-quality data of the HA dataset to “correct” the alternative model generated by the LA dataset. In order to describe the difference between the values $\hat{z}_{LA}(x_i, y_i)$ obtained by the alternative model generated from LA dataset and the values of HA dataset $z_{HA}(x_i, y_i)$, the “adjustment model” is introduced as in literature [26,27]:

$$z_{HA}(x_i, y_i) = \beta(x_i, y_i)\hat{z}_{LA}(x_i, y_i) + \delta(x_i, y_i) + \varepsilon \quad (14)$$

where $\beta(x_i, y_i)$ and $\delta(x_i, y_i)$ represent the scaling factor and the shifting factor, respectively, and ε represents the error term. The adjustment model means that, for positions containing HA points, the value $z_{HA}(x_i, y_i)$ can be obtained by scaling and shifting the predicted value $\hat{z}_{LA}(x_i, y_i)$ estimated by the GP model from the LA dataset, plus an error item associated with the HA dataset. Since the adjustment model is statistically established, it should contain the error term ε to indicate residuals between the HA dataset $z_{HA}(x_i, y_i)$ and $\beta(x_i, y_i)\hat{z}_{LA}(x_i, y_i) + \delta(x_i, y_i)$. Assume that the residual term ε obeys a normal distribution $\varepsilon \sim N(0, \sigma_{\varepsilon}^2)$.

The scaling factor β can be expressed by the linear model:

$$\beta(x_i, y_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i \quad (15)$$

The shifting factor δ can be represented by a new GP model composed of the constant mean value δ_0 and covariance function k_δ :

$$\delta(x_i, y_i) \sim \text{GP}(\delta_0, k_\delta) \quad (16)$$

where the square exponential function is used to represent the covariance function $k_\delta = \sigma_\delta^2 \exp(-\frac{\|v_i - v_j\|^2}{2l_\delta^2})$, as the same in Equation (13).

The parameters $\{\beta_0, \beta_1, \beta_2, \delta_0, \sigma_\delta^2, l_\delta, \sigma_\varepsilon^2\}$ in Equations (14)–(16) are all unknown and can be estimated from the actual measurement data $z_{\text{HA}}(x_i, y_i)$ with $(x_i, y_i) \in V_{\text{HA}} \subset \mathbb{R}^2, i = 1, 2, \dots, n_{\text{HA}}$.

After fully estimating all the unknown parameters mentioned above, correction parameters based on the HA dataset (unbiased) can be obtained according to the adjustment model. Therefore, the original GP model (biased) prediction data can be modified to

$$z_{\text{fusion}}(x_i, y_i) = \beta(x_i, y_i) \hat{z}_{\text{LA}}(x_i, y_i) + \delta(x_i, y_i) \quad (17)$$

3. Experimental Verification

3.1. Simulation Verification

In order to analyze the performance of the proposed data fusion algorithm, a freeform surface is defined by Equation (18):

$$z = \sin(0.8x) + \cos(0.5y) \quad (18)$$

where $x, y \in [-5, 5]$ mm. On such a surface, a set of points is sampled uniformly across the whole surface at an interval of 0.4 mm, which is defined as Dataset 1. Another set of points is sampled at 0.15 mm interval over a part $x, y \in [-3, 3]$ of the surface, which is denoted as Dataset 2. Dataset 1 represents the HA dataset, Dataset 2 is equivalent to the LA dataset, and two sets of Gaussian noise with standard deviation 5 μm and 15 μm are added to Dataset 1 and Dataset 2, respectively, to represent measurement errors of the HA and LA datasets. Dataset 2 is moved to a specific location so that the two sets of points are located in different coordinate systems. The given transformation parameters are $t_x = 1$ mm, $t_y = 1$ mm, $t_z = -0.5$ mm (translation parameters) and $r_x = -0.1$ rad, $r_y = 0.3$ rad, $r_z = 0.2$ rad (rotation parameters). Therefore, these two sets of point cloud data have different resolutions, different uncertainties, and different coordinate systems. These two datasets are common in the multisensor measurement of the freeform surface. The generated Dataset 1 and 2 systems are shown in a common coordinate in Figure 4a.

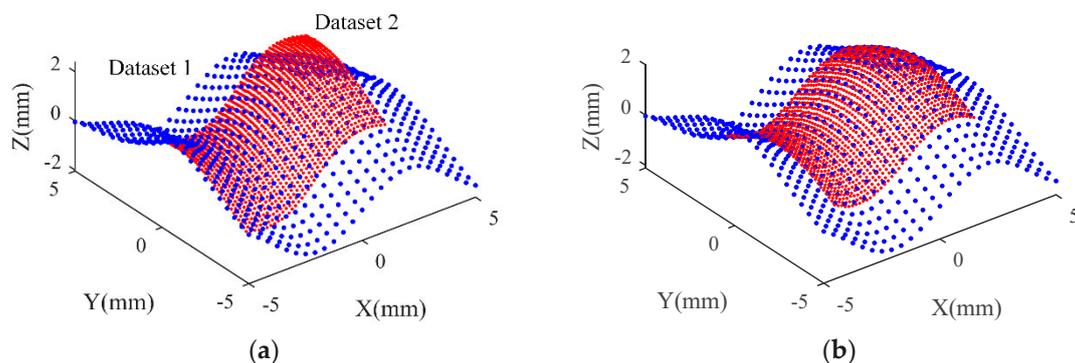


Figure 4. Multisensor data for simulation and data registration results. (a) Two simulation datasets, where blue data represents Dataset 1 and red data represents Dataset 2; (b) Data registration results based on IRLS-ADF method.

Firstly, the proposed robust registration algorithm based on ADF is used to register Dataset 1 and Dataset 2, and unify them into a common coordinate system. Figure 4b shows the data registration results based on IRLS-ADF. In order to evaluate the accuracy and efficiency of the proposed algorithm,

the transformation parameter error and calculation time of the IRLS-ADF and ICP methods are presented in Table 1. Figure 5 shows a comparison of the registration errors between the IRLS-ADF and ICP methods during the iteration process. Both methods are implemented using MATLAB R2016a, with a maximum of 30 iterations, and run on a computer with the Intel® Core™ i7(8 GB RAM) processor. The transformation parameter errors of the IRLS-ADF method is smaller than the ICP method, which indicates that the IRLS-ADF method has higher accuracy. The iteration number of the IRLS-ADF is significantly smaller than that of the ICP method, and the calculation time of IRLS-ADF is also shorter because it is based on the Levenberg–Marquardt method, and achieves faster convergence rates. This result is consistent with the literature [17]. Generally, the IRLS-ADF method has higher precision and faster convergence rates for HA and LA dataset registration, indicating its robustness against the influence of the measurement error.

Table 1. The error of transformation parameters and computation time of the IRLS-ADF and ICP method.

Method	Error of Transformation Parameters						Computation Time (s)
	t_x (μm)	t_y (μm)	t_z (μm)	r_x (mrad)	r_y (mrad)	r_z (mrad)	
IRLS-ADF	1.8	3.7	2.5	0.5	0.2	0.7	1.1
ICP	3.7	5.3	5.2	1.4	5.0	2.1	3.2

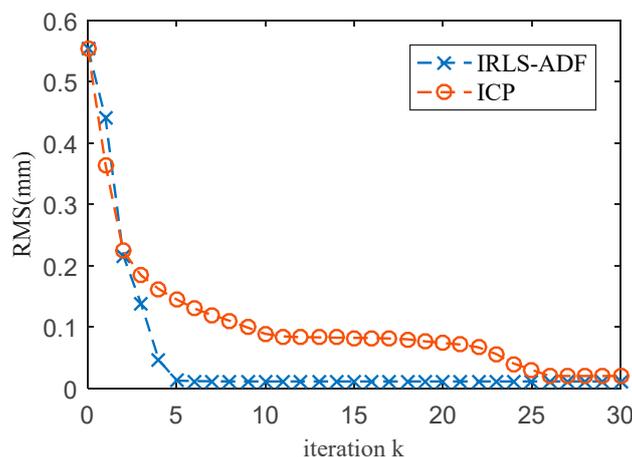


Figure 5. Comparison of registration errors between the IRLS-ADF and ICP methods during the iteration process.

Then, a data fusion model based on the GP model is constructed in the overlapping area according to Section 2.3. Figure 6 shows the fusion data and its estimation uncertainty. Figure 7 shows the form error of the fusion data by comparing it with the designed surface. In order to evaluate the accuracy of the fusion data, the root-mean-square (RMS) error and peak-to-valley (PV) error of the reconstructed model data compared with the designed surface are calculated. Table 2 shows the RMS and PV error comparison between Dataset 1, Dataset 2, and the proposed IRLS-ADF+GP fusion method. It can be seen that the RMS values of Datasets 1 and 2 are 4.3 μm and 14.9 μm , respectively, which is consistent with the measurement noise added in these datasets. Based on the proposed data fusion method, the RMS value of the fusion data is reduced to 1.9 μm , which is significantly lower than the RMS values of Dataset 1 and 2. Similarly, the PV value of the fused data is reduced to 13.8 μm . It shows that the proposed GP-based fusion method is capable of improving the accuracy of the measured datasets. The proposed method is also compared with the widely used data fusion method, which adopts the ICP registration method to align two datasets and integrates them with the weighted mean (WM) method [3]. It is also seen from the results of Table 2 that the proposed method has better fusion accuracy and faster computational efficiency than the ICP-WM method.

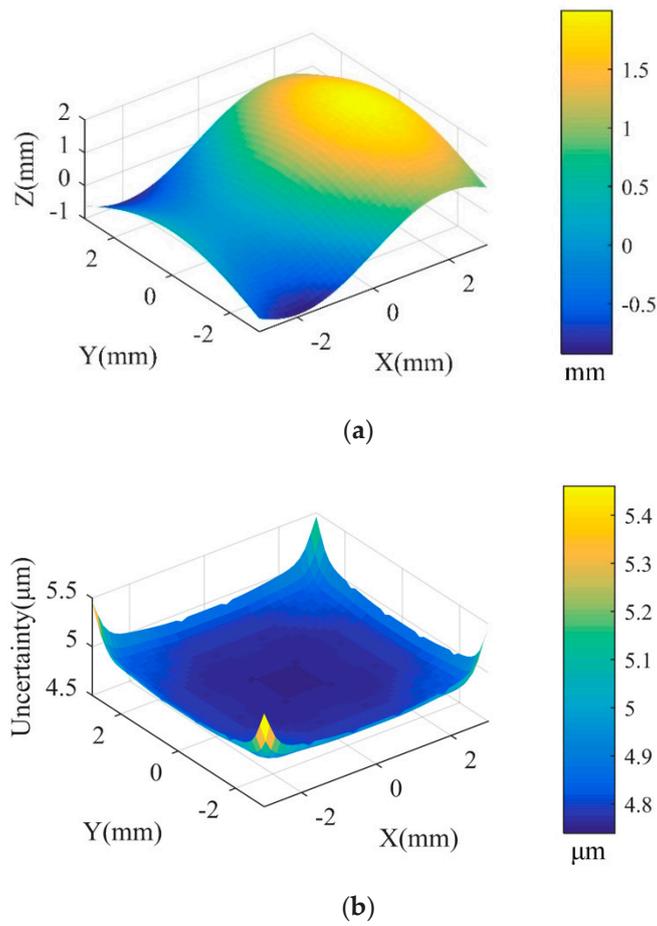


Figure 6. (a) The fusion result of the overlapping area based on the GP model; (b) Uncertainty estimated by the established GP fusion model.

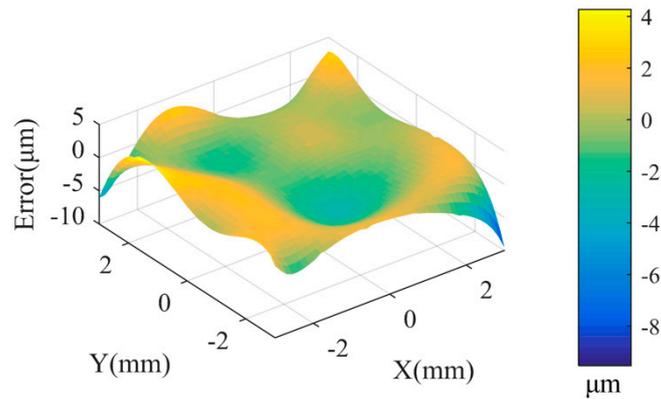


Figure 7. Evaluated error diagram of the fusion data.

Table 2. RMS, PV error, and computation time of different data fusion methods.

	Dataset 1	Dataset 2	Fusion by IRLS-ADF+GP	Fusion by ICP+WM
RMS (μm)	4.3	14.9	<u>1.9</u>	3.4
PV (μm)	19.9	57.4	<u>13.8</u>	17
Computation time (s)	-	-	<u>5.5</u>	8.2

In addition to data fusion performance, measurement efficiency is also a key concern in the multisensor measurement. LA datasets can be obtained through a structured light scanner with high

measurement efficiency. However, HA datasets are usually acquired by CMMs, and the measurement efficiency is low. Only if the number of HA data is small, the measurement time is acceptable. Therefore, multiple datasets (still called Dataset 1), which contain a different number of data points in the overlapping area, are used to evaluate the performance of the proposed algorithm fully. A set of Gaussian noise with standard deviation $5\ \mu\text{m}$ is still added to Dataset 1, and Dataset 2, which is used for fusion with Dataset 1, remains unchanged. Figure 8 shows the final RMS and PV values of fusion results generated by the proposed IRLS-ADF+GP fusion method with the change of HA point number in Dataset 1. As can be seen from Figure 8a, the RMS values of the fusion data do not change much with the decrease of the HA point number in Dataset 1, and remain at a low level ($<2\ \mu\text{m}$). Similarly, as the number of data points in Dataset 1 decreases, the PV values of the fusion data remains low ($<18\ \mu\text{m}$), as shown in Figure 8b. It indicates that the proposed data fusion algorithm can achieve high-quality data fusion performance with a small number of HA data. In actual multisensor measurements, the efficiency of the multisensor measurement can be significantly improved with few HA data points. Therefore, with the LA dataset and a small amount of HA data points, the data fusion algorithm proposed in this paper can complete the data fusion process with a good combination of measurement efficiency and fusion performance. It should be emphasized that too few HA data points are not recommended for multisensor fusion process (for example, <50 in this experiment), because this means that too few “real points” are used to modify the LA dataset, which may lead to a decrease in the registration accuracy and result in poor data fusion performance.

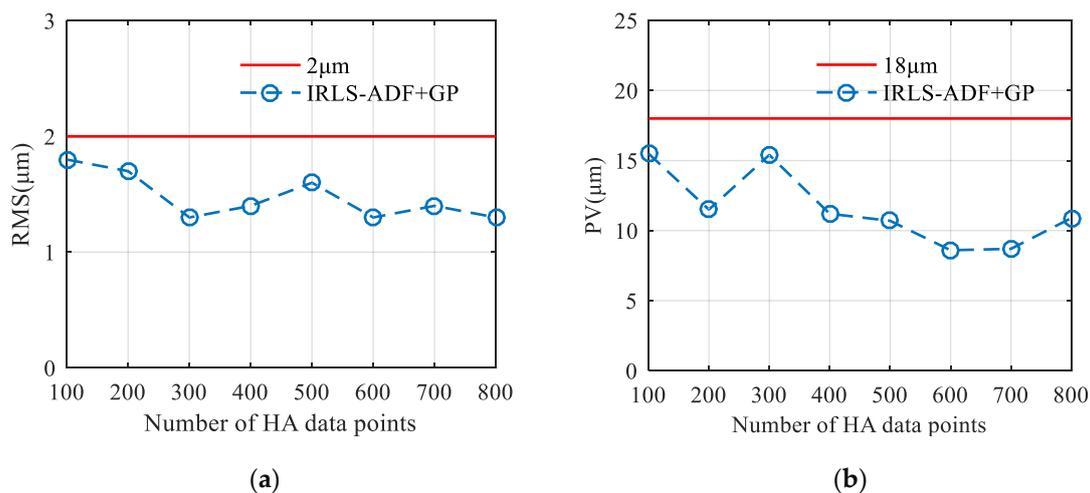


Figure 8. Final RMS and PV values of fusion results when the number of the HA point changes to $N = \{100, 200, 300, 400, 500, 600, 700, 800\}$, respectively. (a) RMS error; (b) PV error.

3.2. Verification in Actual Measurement

The first example of actual measurement is a spherical microstructure surface. The 2D diagram of a single spherical surface is shown in Figure 9a, where the radius of the single spherical surface is $r = 52\ \mu\text{m}$, the depth is $h = 4\ \mu\text{m}$, the diameter of the aperture is $d = 40\ \mu\text{m}$. The spherical microstructure array is distributed on a square grid, and the spacing between the adjacent lines in both the X and Y direction is $\lambda = 100\ \mu\text{m}$. The above microstructure surface has been machined by an ultra-precision machine tool, and the surface topography is captured by a 3D laser microscope Olympus OLS5000, as shown in Figure 9b. The complete geometric information is measured in two steps with $50\times$ objective (zoom is $1.0\times$) and $100\times$ objective (zoom is $3.0\times$), respectively. The field of view of the two objectives is $0.25 \times 0.25\ \text{mm}^2$ and $0.04 \times 0.04\ \text{mm}^2$ respectively. The measurement data are shown in Figure 9c. Due to two different objectives, the two measured datasets have different resolutions and are in different coordinate systems. Then, the proposed method in this paper is used

for the registration and fusion process of the above two datasets to generate a final representation of the microstructure surface.

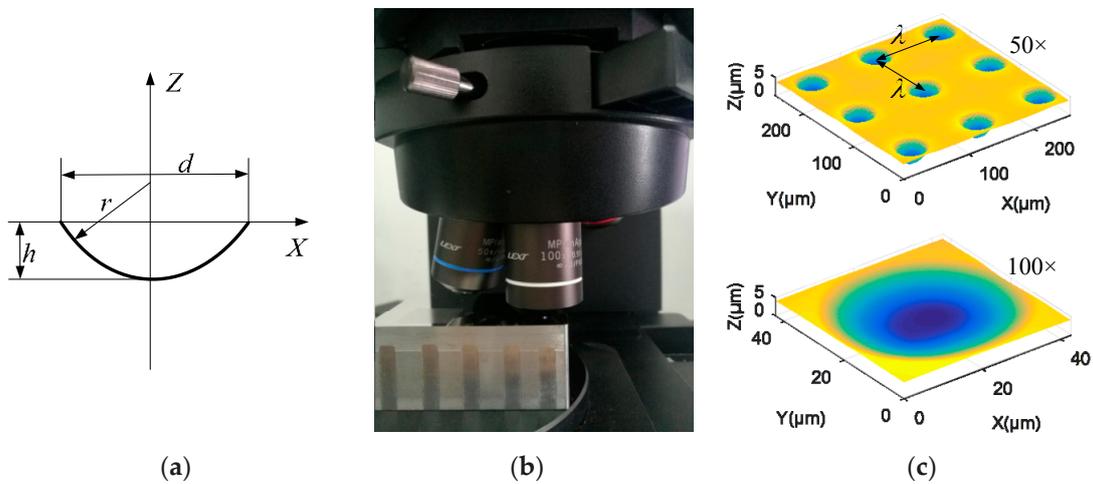


Figure 9. Measurement of a spherical microstructure surface. (a) 2D diagram of a single spherical surface; (b) The measurement process of the machined surface on a 3D laser microscope; (c) Measurement data obtained by 50× and 100× objectives, respectively.

Firstly, the dataset obtained by 100× objective is transformed into the coordinate frame of the dataset measured by 50× objective based on the proposed registration algorithm. Figure 10a shows the registration of measured datasets based on the IRLS-ADF registration method, and Figure 10b presents the datasets after registration. Then, the two datasets at the overlapping area are fused by the proposed fusion method. Finally, the 50× dataset, 100× dataset, and fusion dataset are compared with the original CAD model respectively to verify the quality of the above datasets. The evaluated error diagram of the fusion data is given in Figure 10c. The RMS errors of the three datasets are 0.125 μm , 0.119 μm , and 0.103 μm , respectively. From the above error results, we can see that the dataset processed by the registration and fusion method has higher accuracy than the two original measurement datasets.

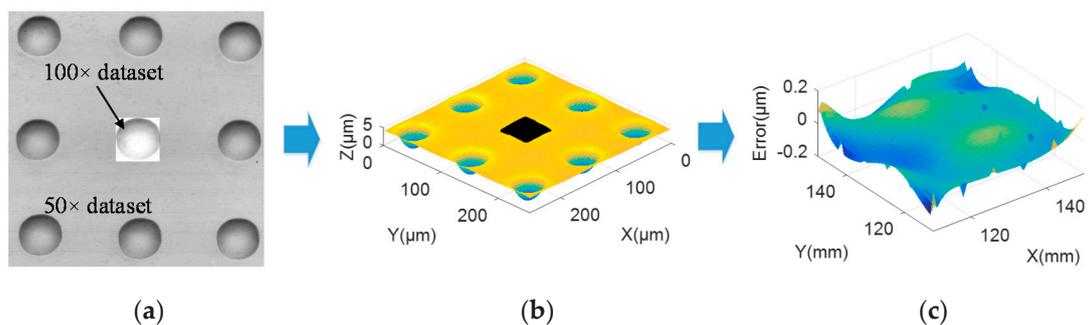


Figure 10. Registration and fusion process of the measured surface data. (a) IRLS-ADF registration; (b) The registered datasets; (c) Evaluated error diagram of the fusion data.

The second example of actual measurement is a machined freeform surface with a geometric dimension of 52 mm \times 50 mm \times 13 mm. The machined surface is measured by two different measurement instruments, namely a high-precision Hexagon Coordinate Measuring Machine (CMM) and a structural light (SL) scanner. Figure 11 shows the measurement process of the machined surface by means of CMM and SL scanner. The length measurement error of the CMM is $U = 1.4 + 3.0 \times L/1000 \mu\text{m}$, and the maximum probing error is $MPEP = 1.2 \mu\text{m}$. The accuracy of the SL scanner is 0.03 mm, and the field of view is 300 \times 240 mm².

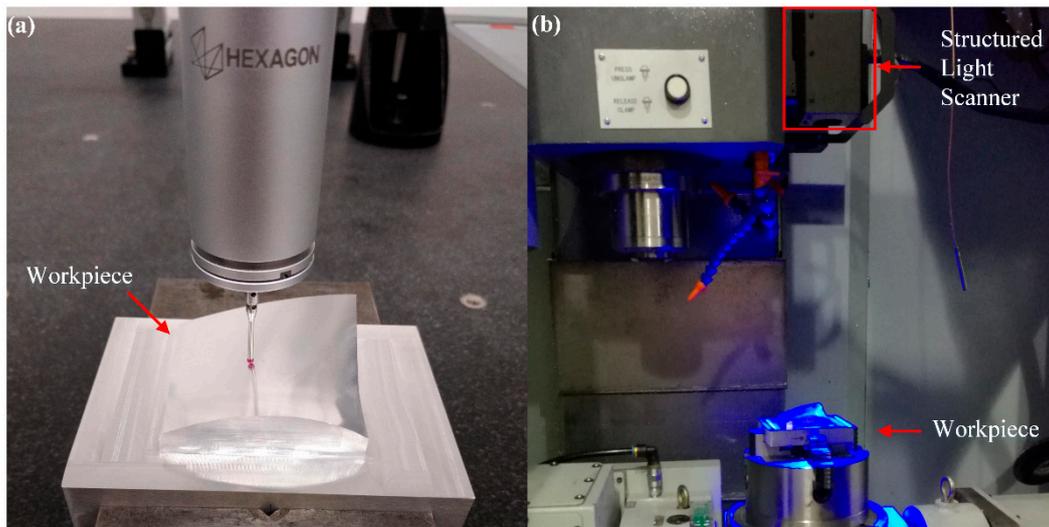


Figure 11. The measurement process of the machined freeform surface. (a) The machined freeform surface on the CMM; (b) The surface under a structured light (SL) scanner.

Figure 12 shows the measurement datasets obtained from the CMM and the structured light scanner. The CMM dataset is a HA dataset, which consists of $n_{\text{CMM}} = 165$ points (Figure 12a) with 4 mm spacing in both X and Y direction. The LA dataset consists of $n_{\text{SL}} = 8131$ SL data points (Figure 12b). The two datasets have different resolutions and uncertainties and are located in different coordinate systems. To verify the fusion performance of the proposed method, we also collect a dataset called a reference dataset, which is composed of $n_{\text{REF}} = 7131$ CMM data points with 0.6 mm spacing over the measured surface.

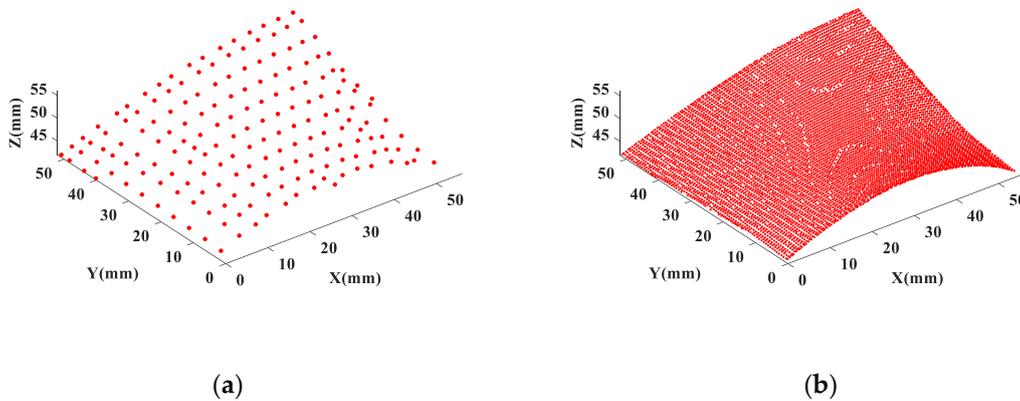


Figure 12. Multisensor measurement datasets. (a) CMM measurement data; (b) SL measurement data.

The proposed algorithm is applied to the fusion of the above two datasets. IRLS-ADF registration algorithm is used to transform two datasets into a common coordinate system. Then a fusion dataset is generated by the GP-based fusion method. Figure 13a shows the fusion dataset. The evaluated error diagram of the fusion data by comparison with the design surface is, as shown in Figure 13b. In order to analyze the performance of the proposed method, the accuracy of the fusion dataset is characterized by RMS and PV error, and the measurement efficiency of the fusion dataset is reflected by the measurement time. Table 3 shows the comparison results of different measurement data, including the CMM dataset, SL dataset, fusion dataset, and reference dataset.

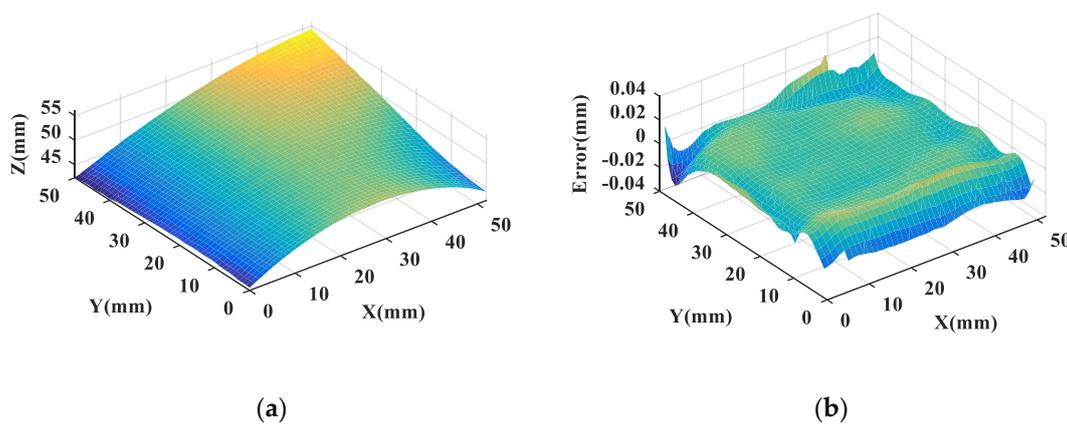


Figure 13. (a) Fusion dataset of the measured surface; (b) Evaluated error diagram of the fusion data.

Table 3. RMS, PV error, and time comparison of different measurement datasets.

Dataset	CMM	SL	Fusion	Reference
RMS (μm)	13.2	17.6	<u>14.1</u>	13.9
PV (μm)	60.4	77.4	<u>66.5</u>	65.4
Time (h)	~0.4	<0.1	<u>~0.4</u>	>2

In terms of accuracy, the RMS value of the fusion dataset generated by the proposed fusion method is $14.1 \mu\text{m}$, and the PV value is $66.5 \mu\text{m}$. Compared with the SL dataset, the accuracy of the fusion dataset is greatly improved. At the same time, the RMS and PV values of the fusion dataset are slightly worse than that of the reference dataset, indicating that the data quality of the fusion dataset is close to that of the reference dataset. In terms of measurement efficiency, the measurement time of the multisensor fusion method based on CMM and SL data is about 0.4 h, which is far less than the measurement time of the reference dataset. It can be seen from the above results that the accuracy of the fusion dataset is significantly improved by means of registering and fusing datasets from CMM and SL, while the measurement efficiency is maintained. In general, a large number of measurement points in the CMM lead to long sampling time, while few CMM measurement points result in insufficient surface sampling, and the error of the machined surface may be underestimated. And the structured light scanner has a fast sampling speed but relatively low accuracy. Therefore, multisensor combination measurement is of great importance to improve the efficiency and precision of complex surface measurement. The above experimental results indicate that the proposed data fusion method can fuse multisensor datasets of complex geometry, with good fusion accuracy and high measurement efficiency.

4. Conclusions

This paper proposes a data fusion method for multisensor combination measurement of complex surfaces. In order to solve the key issues of multisensor fusion, a surface robust registration method is presented to unify the coordinate systems of multisensor datasets, which utilizes the ADF as the basic error metric, and uses M-estimation method to limit the influence of noise points. Then, a GP model-based data fusion method is presented to generate a high-quality fusion dataset. Moreover, the proposed method is verified by simulation and actual experiments. Experimental results show that the proposed method has the capability to fuse multisensor datasets of complex geometry with good fusion performance and fast computational efficiency. In addition, the proposed data fusion algorithm only needs a small amount of HA data points and an LA dataset to complete the data fusion process, achieving a good combination of measurement efficiency and fusion performance.

How to speed up the data fusion method through parallel hardware (FPGA or GPU) is a task that needs to be solved in the future. In addition, the data fusion process may involve more than two

datasets in practical application. The proposed data fusion method can only be used for the fusion between two different datasets, and further research is needed. Furthermore, the prior knowledge of the uncertainty associated with each sensor could be utilized in the future research.

Author Contributions: Conceptualization—J.D. and Q.L.; Methodology—J.D.; Validation—J.D., M.B. and P.S.; Formal Analysis—J.D.; Investigation—J.D.; Resources—P.S.; Data Curation—J.D.; Writing-Original Draft Preparation—J.D.; Writing-Review and Editing—J.D.; Supervision—Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key Development Program of China, grant number No. 2016YFB1102503 and National Natural Science Foundation of China, grant number No.11290144.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jiang, X.J.; Whitehouse, D.J. Technological shifts in surface metrology. *CIRP Ann.* **2012**, *61*, 815–836. [[CrossRef](#)]
- Fang, F.Z.; Zhang, X.D.; Weckenmann, A.; Zhang, G.X.; Evans, C. Manufacturing and measurement of freeform optics. *CIRP Ann.* **2013**, *62*, 823–846. [[CrossRef](#)]
- Wang, J.; Leach, R.K.; Jiang, X. Review of the mathematical foundations of data fusion techniques in surface metrology. *Surf. Topogr. Metrol. Prop.* **2015**, *3*, 023001. [[CrossRef](#)]
- Weckenmann, A.; Jiang, X.; Sommer, K.D.; Neuschaefer-Rube, U.; Seewig, J.; Shaw, L.; Estler, T. Multisensor data fusion in dimensional metrology. *CIRP Ann.* **2009**, *58*, 701–721. [[CrossRef](#)]
- Rak, M.B.; Wozniak, A.; Mayer, J.R.R. The use of low density high accuracy (LDHA) data for correction of high density low accuracy (HDLA) point cloud. *Opt. Lasers Eng.* **2016**, *81*, 140–150. [[CrossRef](#)]
- Peng, J.; Xu, W.; Yuan, H. An Efficient Pose Measurement Method of a Space Non-Cooperative Target Based on Stereo Vision. *IEEE Access* **2017**, *5*, 22344–22362. [[CrossRef](#)]
- Peng, J.; Xu, W.; Liang, B.; Wu, A. Virtual Stereovision Pose Measurement of Noncooperative Space Targets for a Dual-Arm Space Robot. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 76–88. [[CrossRef](#)]
- Sładek, J.; Błaszczuk, P.M.; Kupiec, M.; Sitnik, R. The hybrid contact–optical coordinate measuring system. *Measurement* **2011**, *44*, 503–510. [[CrossRef](#)]
- Rak, M.B.; Mayer, R.; Wozniak, A. Proximity weighted correction of high density high uncertainty (HDHU) point cloud using low density low uncertainty (LDLU) reference point coordinates. *Opt. Lasers Eng.* **2015**, *68*, 160–165. [[CrossRef](#)]
- Peng, J.; Xu, W.; Liang, B.; Wu, A. Pose Measurement and Motion Estimation of Space Non-Cooperative Targets Based on Laser Radar and Stereo-Vision Fusion. *IEEE Sens. J.* **2019**, *19*, 3008–3019. [[CrossRef](#)]
- Qian, Z.; Seepersad, C.C.; Joseph, V.R.; Allen, J.K.; Wu, C.F.J. Building Surrogate Models Based on Detailed and Approximate Simulations. *J. Mech. Des.* **2006**, *128*, 668. [[CrossRef](#)]
- Zhao, Y.; Wang, Y.; Ye, X.; Wang, Z.; Fu, L.; Liu, C.; Wang, Z. Multi-Sensor Registration in High-Precision CMM Based on a Composite Standard. *Sensors* **2018**, *18*, 1220. [[CrossRef](#)] [[PubMed](#)]
- Besl, P.J.; McKay, N.D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [[CrossRef](#)]
- Pottmann, H.; Huang, Q.X.; Yang, Y.L.; Hu, S.M. Geometry and convergence analysis of algorithms for registration of 3D shapes. *Int. J. Comput. Vis.* **2006**, *67*, 277–296. [[CrossRef](#)]
- Pottmann, H.; Leopoldseder, S.; Hofer, M. Registration without ICP. *Comput. Vis. Image Underst.* **2004**, *95*, 54–71. [[CrossRef](#)]
- Wang, W.P.; Pottmann, H.; Liu, Y. Fitting B-spline curves to point clouds by curvature-based squared distance minimization. *ACM Trans. Graph.* **2006**, *25*, 214–238. [[CrossRef](#)]
- Li, W.L.; Yin, Z.P.; Huang, Y.A.; Xiong, Y.L. Three-dimensional point-based shape registration algorithm based on adaptive distance function. *IET Comput. Vis.* **2011**, *5*, 68. [[CrossRef](#)]
- Li, W.L.; Yin, Z.P.; Xiong, Y.L. Adaptive Distance Function and its Application in Free-form Surface Localization. In Proceedings of the Icia: 2009 International Conference on Information and Automation, Zhuhai, Macau, China, 22–24 June 2009; Volume 1–3, pp. 19–23.
- Jamshidi, J.; Owen, G.W.; Mileham, A.R. A New Data Fusion Method for Scanned Models. *J. Comput. Inf. Sci. Eng.* **2006**, *6*, 340. [[CrossRef](#)]

20. Qian, P.Z.G.; Wu, C.F.J. Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments. *Technometrics* **2008**, *50*, 192–204. [[CrossRef](#)]
21. Xia, H.; Ding, Y.; Mallick, B.K. Bayesian hierarchical model for combining misaligned two-resolution metrology data. *IIE Trans.* **2011**, *43*, 242–258. [[CrossRef](#)]
22. Kong, L.B.; Ren, M.J.; Xu, M. Development of Data Registration and Fusion Methods for Measurement of Ultra-Precision Freeform Surfaces. *Sensors* **2017**, *17*, 1110. [[CrossRef](#)]
23. Gong, M.; Zhang, Z.; Zeng, D.; Peng, T. Three-Dimensional Measurement Method of Four-View Stereo Vision Based on Gaussian Process Regression. *Sensors* **2019**, *19*, 4486. [[CrossRef](#)] [[PubMed](#)]
24. Senin, N.; Colosimo, B.M.; Pacella, M. Point set augmentation through fitting for enhanced ICP registration of point clouds in multisensor coordinate metrology. *Robot. Comput. Integr. Manuf.* **2013**, *29*, 39–52. [[CrossRef](#)]
25. Ren, M.J.; Sun, L.J.; Liu, M.Y.; Cheung, C.F.; Yin, Y.H.; Cao, Y.L. A weighted least square based data fusion method for precision measurement of freeform surfaces. *Precis. Eng.* **2017**, *48*, 144–151. [[CrossRef](#)]
26. Xia, H.; Ding, Y.; Wang, J. Gaussian process method for form error assessment using coordinate measurements. *IIE Trans.* **2008**, *40*, 931–946. [[CrossRef](#)]
27. Colosimo, B.M.; Pacella, M.; Senin, N. Multisensor data fusion via Gaussian process models for dimensional and geometric verification. *Precis. Eng.* **2015**, *40*, 199–213. [[CrossRef](#)]
28. del Castillo, E.; Colosimo, B.M.; Tajbakhsh, S.D. Geodesic Gaussian Processes for the Parametric Reconstruction of a Free-Form Surface. *Technometrics* **2015**, *57*, 87–99. [[CrossRef](#)]
29. Colosimo, B.M.; Cicorella, P.; Pacella, M.; Blaco, M. From Profile to Surface Monitoring: SPC for Cylindrical Surfaces Via Gaussian Processes. *J. Qual. Technol.* **2014**, *46*, 95–113. [[CrossRef](#)]
30. Ding, J.; Liu, Q.; Sun, P. A robust registration algorithm of point clouds based on adaptive distance function for surface inspection. *Meas. Sci. Technol.* **2019**, *30*, 075003. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).