Aggregation-Value-Based Active Sampling Method for Multi-sensor Freeform Surface Measurement and Reconstruction

CHEN Gengxiang[1,2], *LI Yingguang*[1*], *MEHDI-SOUZANI Charyar*[2], *LIU Xu*[3]

1. College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China; 2. Université Paris-Saclay, Gif-Sur-Yvette 91190, France; 3. School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211816, P. R. China

(Received 25 October 2024; revised 25 November 2024; accepted 2 December 2024)

Abstract: Freeform surface measurement is a key basic technology for product quality control and reverse engineering in aerospace field. Surface measurement technology based on multi-sensor fusion such as laser scanner and contact probe can combine the complementary characteristics of different sensors, and has been widely concerned in industry and academia. The number and distribution of measurement points will significantly affect the efficiency of multi-sensor fusion and the accuracy of surface reconstruction. An aggregation-value-based active sampling method for multi-sensor freeform surface measurement and reconstruction is proposed. Based on game theory iteration, probe measurement points are generated actively, and the importance of each measurement point on freeform surface to multi-sensor fusion is clearly defined as Shapley value of the measurement point. Thus, the problem of obtaining the optimal measurement point set is transformed into the problem of maximizing the aggregation value of the sample set. Simulation and real measurement results verify that the proposed method can significantly reduce the required probe sample size while ensuring the measurement accuracy of multi-sensor fusion.

Key words: multi-sensor fusion; multi-sensor measurement; data sampling; active learning; Shapley value; intelligent sampling

CLC number: TN925 **Document code:** A **Article ID:** 1005-1120(2024)06-0700-10

0 Introduction

Geometry digitising of freeform surfaces is essential for ensuring quality control and facilitating reverse engineering in critical fields such as aerospace, automatics, optics, etc[1-3]. The advanced development of information technology and sensory advancements has significantly enhanced the accuracy and real-time capabilities of geometry digitisation, thereby playing a crucial role in digital twins and Industry 4.0. With the increasing demand for high precision and efficiency surface measurement in modern manufacturing industries, the integration of measurement data from multiple sensors, including laser scanners and touch probes, becomes increasingly important[4-5]. For example, touch probes with coordinate measuring machines (CMM) have high accuracy but are time-consuming and constrained by environmental factors. In contrast, laser scanners and other non-contact measurement technologies, although less precise, can generate high-resolution point clouds efficiently. Multi-sensor measurement can leverage complementary characteristics of different sensors, thus not only reducing the measurement costs but also improving the quality of final measurement results[6-7].

For multi-sensor measurement, the integration of touch probe points aims to compensate for low accuracy of laser scanner points[8]. Therefore, the number and distribution of touch probe sampling

*Corresponding author, E-mail address: liyingguang@nuaa.edu.cn.

How to cite this article: CHEN Gengxiang, LI Yingguang, MEHDI-SOUZANI Charyar, et al. Aggregation-value-based active sampling method for multi-sensor freeform surface measurement and reconstruction[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2024, 41(6): 700-709.

http://dx.doi.org/10.16356/j.1005-1120.2024.06.003

points on the surface significantly influence the efficiency of the multi-sensor measurement and the accuracy of the surface reconstruction[9]. The one-shot sampling can provide the pre-defined measurements points, namely coresets, based on the distribution, surface characteristics, or surface reconstruction criteria[10-11], where active sampling can generate subsequent measurement points iteratively based on the given query criteria, like the uncertainty, informativeness, or surface reconstruction accuracy.

From the perspective of sampling strategies, the current sampling methods for surface measurement can be divided into representative sampling and adaptive sampling. Representative sampling involves selecting a core set that represents the distribution of overall data or the feature space, with common techniques including random sampling, uniform sampling, Hammersley random sampling, and Latin hypercube sampling[12]. Adaptive sampling focuses on placing more points in areas of interest guided by task-related information. A typical example of adaptive sampling is curvature-based sampling for surface reconstruction, where higher density samples are assigned to the region with higher curvature[13-14]. The measurement points can also be designed based on the mesh structure of the surface, namely mesh-based sampling[15]. Some other researchers also developed adaptive sampling methods considering both the arc length and curvature[16]. However, the classical sampling methods mentioned above typically offer offline sampling strategies for surface measurements, and the sampling points cannot be modified or optimized based on the measurement results.

Intelligent sampling, or active sampling, aims to select samples iteratively from the most beneficial spots on the surface being measured[17-18]. This approach helps ensure that the reconstructed surface is as accurate as possible, achieving the desired level of precision with the fewest samples necessary. The iterative selecting criteria can be defined by reducing the uncertainty[19], reducing the maximum sample deviation[20], etc. In the machine learning field, learning a model through actively adding new samples can be defined as active learning problem, which have been well investigated and applied in many engineering scenarios[21-22]. However, for surface measurement problems, the performance of existing iterative selecting criteria heavily relies on the stability of the surface reconstruction algorithm used in evaluating the uncertainty or deviation. Therefore, how to develop a more general and robust iterative selecting criteria to realize active sampling still deserves further investigation.

In this research, we propose an aggregation value (AV) based active sampling method to iteratively determine the distribution of touch probe points for multi-sensor measurement. The AV sampling is a method recently developed based on combinational game theory[9,23], which uses the aggregation value as a general criterion for evaluating a dataset's value and has been applied to various engineering problems characterised by data scarcity. We utilize the AV sampling to guide the iterative generation of touch probe points for multi-sensor measurement. The general procedure of the proposed method is depicted in Fig.1. For a given freeform sur-

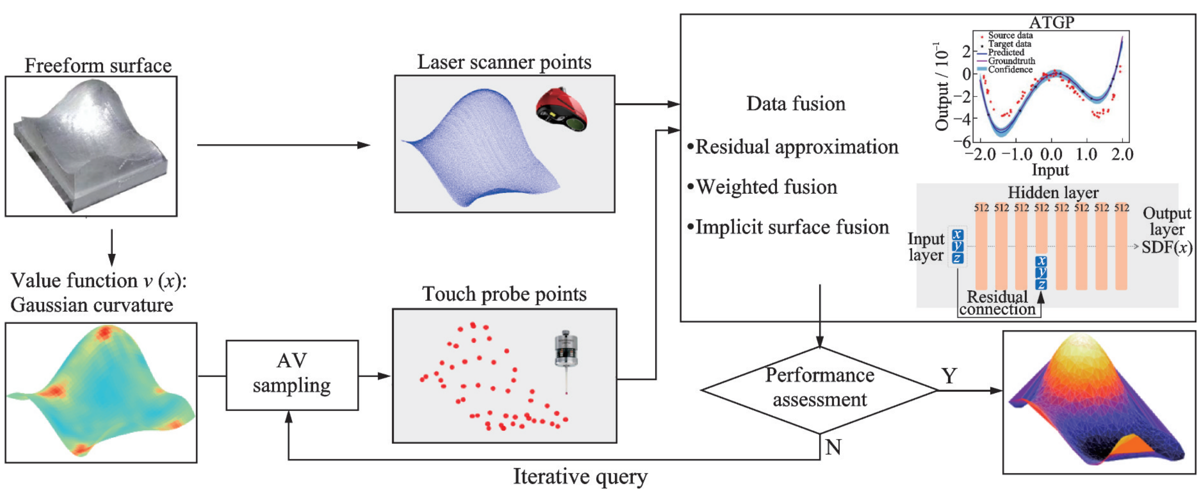


Fig.1 Aggregation-value-based active sampling method for multi-sensor freeform surfaces measurement and reconstruction

face, the Gaussian curvature is calculated to create a value function $v(x)$, which informs the AV sampling and determines the optimal placement of touch probe points. These sampled touch probe points and the laser scanner data are then processed through data fusion algorithms, including residual approximation, weighted fusion, or implicit surface fusion, to accurately reconstruct the surface. Finally, the reconstructed surface undergoes a performance assessment to verify the accuracy and effectiveness. We can iteratively add new touch probe points based on the AV sampling until the measurement and reconstruction results meet the requirements. Section 1 introduces the proposed method in detail, Section 2 reports the experimental verification of the proposed method, and some conclusion are drawn in Section 3.

1 Method

1.1 Problem definition

Multi-sensor measurement in this research aims to reconstruct a CAD model of a freeform surface based on the measurement data from both a touch probe and a laser scanner. The high-accuracy touch probe data, denoted as $\mathcal{D}_h$, consists of n points, $\mathcal{D}_h=\{\boldsymbol{p}_h^1,\boldsymbol{p}_h^2,\cdots,\boldsymbol{p}_h^n\}$, where each point is represented as $\boldsymbol{p}_h^i=(x_h^i,y_h^i,z_h^i), i=1,2,\cdots,n$. Meanwhile, the lower-accuracy laser scanner data $\mathcal{D}_l$, consists of m points, namely $\mathcal{D}_l=\{\boldsymbol{p}_l^1,\boldsymbol{p}_l^2,\cdots,\boldsymbol{p}_l^m\}$, where each point is given as $\boldsymbol{p}_l^i=(x_l^i,y_l^i,z_l^i), i=1,2,\cdots,m$.

The challenge of the active sampling problem in this context can be defined as how to determine the next touch probe point, $\boldsymbol{p}_h^{n+1}$. The multi-sensor data fusion refers to establishing a mathematical representation of the target freeform surface using both $\mathcal{D}_h$ and $\mathcal{D}_l$. The following sections will introduce the aggregation-value-based active sampling strategy and the subsequent multi-sensor data fusion method for surface reconstruction.

1.2 Aggregation-value-based active sampling

From the perspective of combinational game theory, reconstructing a surface with multiple points can be treated as a cooperative game involving multiple players, where each point can be defined as a player[9]. The objective of this game is to reduce the surface reconstruction error. As a result, sampling a touch probe point means finding the most valuable "player" in this game. When an initial touch probe point dataset is obtained, active sampling refers to finding the potential next point that can bring added benefits to the existing dataset. This section will introduce aggregation-value-based active sampling, explaining how to define the value of a dataset and how to calculate the added value of a given point.

The function that quantitatively measures the value of a sample x in a given learning task can be defined as the value function $v(x)$. For general machine learning tasks, the value function $v(x)$ can be evaluated using Shapley theory. According to the analysis in Ref.[9], the value function can also be defined based on domain knowledge, as long as it can offer information that is positively correlated with the task[24]. The curvature is the widely used prior knowledge for traditional measurement sampling methods. Besides, the previous research in surface reconstruction also revealed that the points with higher curvature are more valuable for different reconstruction algorithms[15,25]. Therefore, the value function in this research can be directly defined as the Gaussian curvature of the surface, which can be solved from either the CAD model or the laser scanner data.

Although value function $v(x)$ (Gaussian curvature function) can represent the value of each point for the surface reconstruction, sampling multiple high-value points cannot correspondingly bring multiple values because similar or close high-value points contain redundant information for surface reconstruction. In response to this problem, the concept aggregation value is proposed to represent the actual value of a datasets by considering the neighbouring influence.

The value aggregation function (VAF) of a sample x_* is defined as adding a kernel function to the value function as

$$v'(x,x_*)=v(x)k(x,x_*) \tag{1}$$

where $k(x, x_*)$ is the kernel function for aggregating the neighbouring values. The radial basis function (RBF) is adopted in this research because of its simplicity. The RBF is defined as

$$k(x, x_*) = \exp\left(-\frac{(x - x_*)^2}{\sigma}\right) \tag{2}$$

where σ is the kernel width and can determine the neighbouring influence range. Note that, other kernel functions, like Laplace kernel function, or inverse multiquadric kernel, can also be applied to the definition of VAF. Ref.[9] has investigated the value aggregation performance of different value functions.

Figs.2(a—c) show how the different bandwidth parameters of the RBF kernel function influence the value aggregation ranges of VAF. As shown in Fig.2(a), it is clear that the VAF $v'(x, x_*)$ will converge to $v(x_*)$ when the bandwidth σ approaches zero. By comparison, when the bandwidth σ becomes too large, the VAFs for all samples converge to the value function $v(x)$ itself (as shown in Fig.2(c)), i.e., the values at each point are no longer distinguishable.

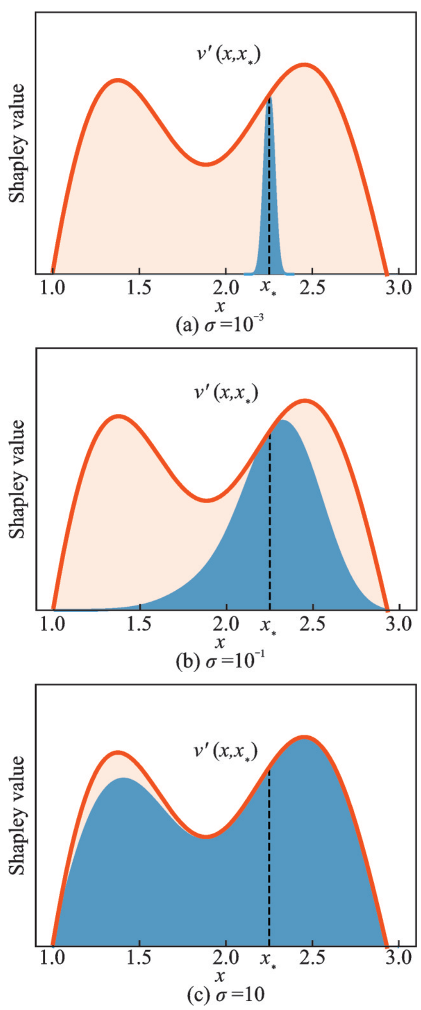


Fig.2 Influence of kernel function in value aggregation

The value function and VAF are illustrated in Fig.3, where the orange curve represents the value function $v(x)$. The VAFs are designed to represent the values of neighbouring points rather than the point itself. Since the VAF of each sample is defined on the entire feature space, the VAFs of different samples may overlap, which can represent the redundant information explicitly. The two samples in Fig.3(a) are far from each other, so there is only a little overlap of VAFs. In contrast, the two samples close to each other in Fig.3(b) have greater overlaps of their VAFs, namely more redundant information.

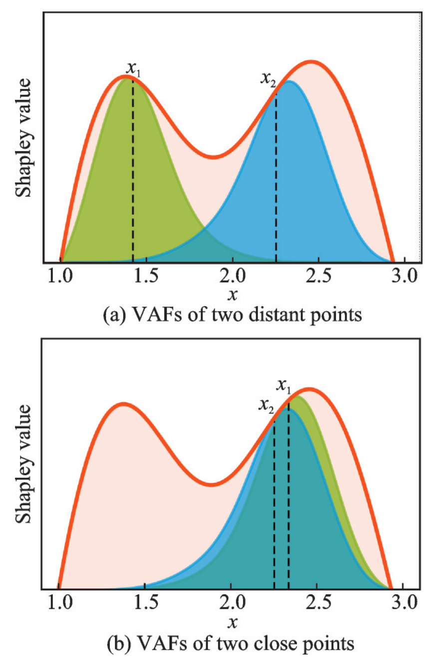


Fig.3 Value aggregation function examples

Therefore, the actual value of a dataset can be simply defined as the area of the "union" VAFs, like the union in Boolean geometry operations. For a dataset $S = \{x_1, x_2, \cdots, x_n\}$, the "union" of VAFs can be defined as

$$v'(x, S) = v(x)\max\{k(x, x_1), k(x, x_2), \cdots, k(x, x_n)\} \tag{3}$$

Then the actual value, defined as aggregation value, can be defined as the integral of VAFs shown as

$$v_{\text{agg}}(S)=\int p(x)v'(x,S)\mathrm{d}x \tag{4}$$

The estimation of aggregation values under finite number of samples is

$$\hat{v}_{\text{agg}}(S)=\frac{1}{n}\sum_{x\in S}v'(x,S) \tag{5}$$

By this definition, the optimal iterative sampling problem can be defined as finding the new sample x_{n+1} that can maximize the increment of aggregation value of the entire sample set. The optimization target can be represented as

$$x_{n+1}=\arg\max_{x}\{\hat{v}_{\text{agg}}(S\cup x_{n+1})-\hat{v}_{\text{agg}}(S)\} \tag{6}$$

The above problem is a classical sub-modularity optimization problem. The greedy algorithm can provide an approximation to optimal results[22]. Therefore, the AV sampling can provide new touch probe points iteratively until the multi-sensor data fusion results satisfy the measurement and reconstruction requirements.

1.3 Multi-sensor data fusion

For each iteration, after obtaining laser scanner points $\mathcal{D}_\text{l}$ and touch probe points $\mathcal{D}_\text{h}$, data fusion techniques can leverage these two groups of points to reconstruct the surface more accurately. The existing classical multi-sensor fusion techniques include residual approximation (RA), weighted fusion, and implicit fusion[25]. The RA aims to establish a residual function between the two datasets to compensate for the low accuracy points. Weighted fusion adjusts the surface by assigning different weights to different datasets based on the uncertainty levels of sensors. The recently proposed implicit fusion method redefines the multi-sensor fusion problem as a transfer learning problem, using low accuracy points to train an implicit neural network, which is then fine-tuned with high accuracy points[26]. In this research, the RA is adopted to fuse the two datasets because of its simplicity and generalisability.

The first step for RA is establishing the prediction model of the low-accuracy dataset with a learning model, such as neural networks (NN) or Gaussian process (GP) regressors[27]. For example, given a data set $\mathcal{D}=\{\boldsymbol{X},\boldsymbol{y}\}$, the GP model will provide predictions for a new sample $\boldsymbol{x}_*$ with both mean value and the corresponding covariance, and the predicted results can be represented as

$$f(\boldsymbol{x}_*|\boldsymbol{X},\boldsymbol{y})\sim N(\boldsymbol{\mu}_*,\boldsymbol{\Sigma}_*)$$
$$\boldsymbol{\mu}_*=\boldsymbol{K}_*^\mathrm{T}\boldsymbol{K}^{-1}\boldsymbol{y},\ \boldsymbol{\Sigma}_*=\boldsymbol{K}_{**}-\boldsymbol{K}_*^\mathrm{T}\boldsymbol{K}^{-1}\boldsymbol{K}_* \tag{7}$$

where $\boldsymbol{\mu}_*$ and $\boldsymbol{\Sigma}_*$ are the predicted mean value and covariance, respectively; $\boldsymbol{K}_*=\kappa(\boldsymbol{X},\boldsymbol{x}_*)$, $\boldsymbol{K}_{**}=\kappa(\boldsymbol{x}_*,\boldsymbol{x}_*)$ and $\boldsymbol{K}=\kappa(\boldsymbol{X},\boldsymbol{X})$ are covariance matrices, κ is a predefined kernel function. In the subsequent text, we will denote the predicted mean of the GP model trained using dataset $\mathcal{D}$ as $\mathrm{GP}(\mathcal{D})$ for simplicity.

For laser scanner points $\mathcal{D}_\text{l}$, the corresponding surface can be represented as a z-direction GP prediction function on the (x, y) coordinates as

$$f_\text{l}(x,y)=\mathrm{GP}(\mathcal{D}_\text{l}) \tag{8}$$

Since touch probe points are more accurate than laser scanner points, there will be systematic errors between touch probe points and the laser scanner prediction model $f_\text{l}(x,y)$. The residual of each touch probe point $\boldsymbol{p}_\text{h}^i$ can be represented as

$$r_i=z_\text{h}^i-f_\text{l}(x_\text{h}^i,y_\text{h}^i) \tag{9}$$

Then, the systematic error function between the two datasets can be represented by training a residual GP model on the n residual values as

$$r(x,y)=\mathrm{GP}([r_1,r_2,\cdots,r_i]) \tag{10}$$

Thus, the final reconstructed surface can be obtained by combining the prediction surface from the low-accuracy dataset with the residual model as

$$f(x,y)=f_\text{l}(x,y)+r(x,y) \tag{11}$$

The RA fusion method can leverage high-accuracy touch probe points and high-resolution laser scanner points to obtain a more accurate reconstructed surface. In addition, the data-fusion process can be iteratively repeated after active sampling new touch probe points until the results satisfy the measurement and reconstruction requirements.

2 Experiments

This section will validate the proposed aggregation-value-based active sampling method for multi-sensor freeform surface measurement through a toy case study and actual measurements.

2.1 Toy case study

The MATLAB Peaks surface is chosen as the toy case study to assess the effectiveness of both active sampling and data fusion procedures. The formula of MATLAB Peaks surface is

$$z=6\left(1-\frac{x}{16}\right)^2\cdot\exp\left(-\left(\frac{x}{16}\right)^2-\left(\frac{y}{16}+1\right)^2\right)-20\left(\frac{x}{5\times 16}-\left(\frac{x}{16}\right)^3-\left(\frac{y}{16}\right)^5\right)\cdot\exp\left(-\left(\frac{x}{16}\right)^2-\left(\frac{y}{16}\right)^2\right)-\frac{2}{3}\exp\left(-\left(\frac{x}{16}+1\right)^2-\left(\frac{y}{16}\right)^2\right)\quad x,y\in[-40,40] \tag{12}$$

As shown in Fig.4, the Gaussian curvature function, evaluated through the python library Pyntcloud, is used as the value function for AV sampling. The potential sampling space consists of 900 points uniformly sampled from the Peaks surface. To evaluate the performance of point sampling, we first compare the surface reconstruction results from points sampled by the AV method to those obtained through random sampling. The remaining points in the sample space are selected as test points to calculate the mean absolute error (MAE) of the reconstructed surfaces.

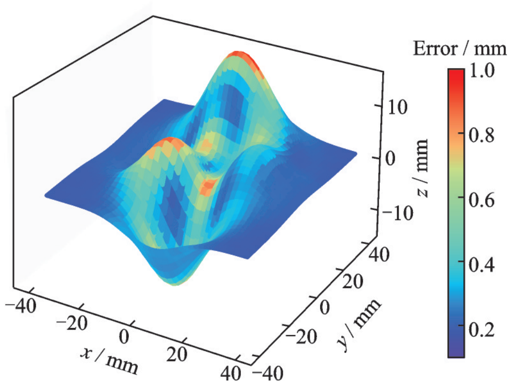


Fig.4 The absolute Gaussian curvature function of MATLAB Peaks surface

Fig.5 shows the comparison of reconstructed MAE with different sample sizes from 20 to 180. AV sampling achieves a significantly smaller MAE compared to random sampling, particularly with a smaller number of samples. This result indicates that points from AV sampling provide more valuable information for surface reconstruction, which could also enhance further multi-sensor data fusion.

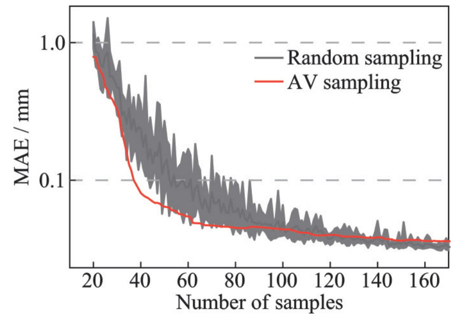


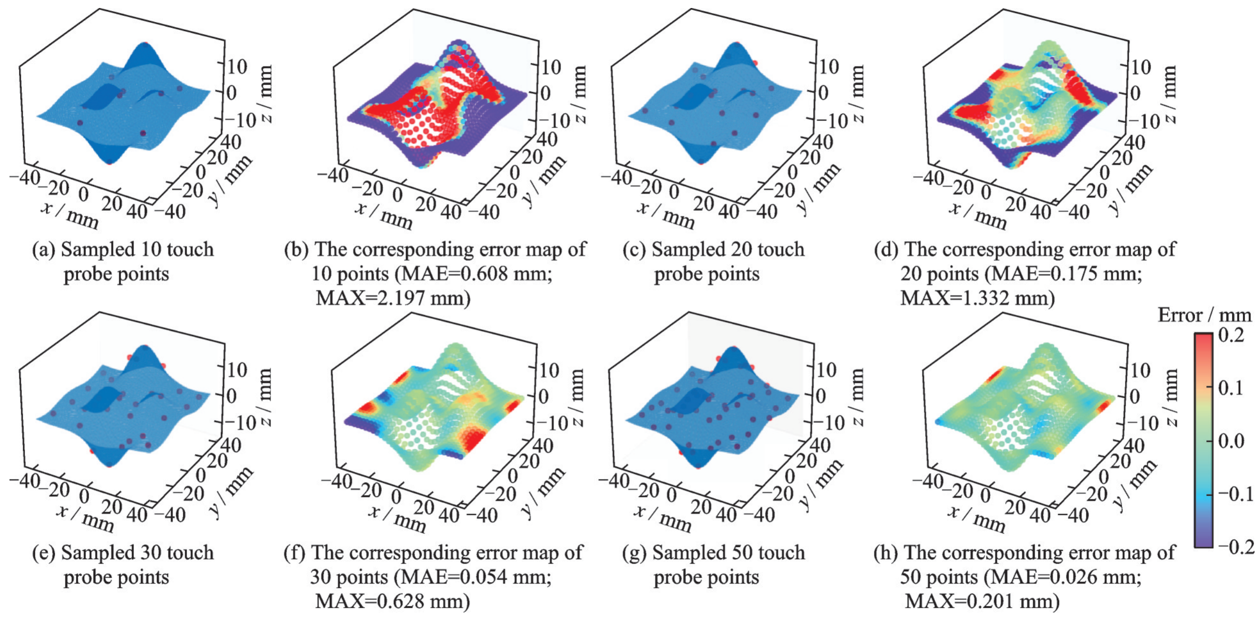
Fig.5 Comparison of reconstructed MAE with different number of samples

To evaluate the multi-sensor fusion performance of the proposed method, both touch probe points and laser scanner points are simulated based on the MATLAB Peaks surface. A Gaussian noise with standard deviation of 0.005 mm is added to the sampled points and the test points to simulate touch probe measurement results. 1 600 points are uniformly sampled from the Peaks surface to simulate the laser scanner points. A larger Gaussian noise of 0.02 mm is added to the laser scanner points as the random error. In addition, the following residual function is added to the laser scanner points to simulate the systematic error.

$$h(x,y)=\sin\left(\frac{x}{8}+2\right)+\sin\left(\frac{y}{8}+2\right) \tag{13}$$

Figs. 6(a, c, e, g) display the sampled touch probe points with sizes of 10, 20, 30, and 50. The newly added points are located in positions that can bring the most incremental information based on the existing point clouds. Subsequently, the RA multi-sensor fusion method reconstructs the surface using both the laser scanner data and the sampled touch probe data. The performance of the surface is characterized by the error across 900 test points. The corresponding MAE and the maximum absolute error (marked as MAX) are reported in Figs.6(b,d,f,h), respectively.

When using only 10 touch probe points, the multi-sensor reconstructed surface exhibits significant errors across the entire surface, with MAE of 0.608 mm and MAX of 2.197 mm. This indicates



(a) Sampled 10 touch probe points (b) The corresponding error map of 10 points (MAE=0.608 mm; MAX=2.197 mm) (c) Sampled 20 touch probe points (d) The corresponding error map of 20 points (MAE=0.175 mm; MAX=1.332 mm)

(e) Sampled 30 touch probe points (f) The corresponding error map of 30 points (MAE=0.054 mm; MAX=0.628 mm) (g) Sampled 50 touch probe points (h) The corresponding error map of 50 points (MAE=0.026 mm; MAX=0.201 mm)

Fig.6 Sampled touch probe points with different sizes and the corresponding error maps of surface from RA multi-sensor reconstruction

that the touch probe points are insufficient to accurately represent the systematic errors in the laser scanner data. However, as the number of touch probe points increases to 20, the error in central region of the surface decreases rapidly. The error continues to decrease with the addition of more samples. Finally, using 50 touch probe samples achieves MAE of 0.026 mm and MAX of 0.201 mm.

Fig.7 shows that the MAE and MAX of the multi-sensor reconstructed surface decrease as the number of touch probe points iteratively increases. The MAE can drop to less than 0.05 mm with more than 30 points. These results demonstrate that the proposed active sampling-based multi-sensor reconstruction method can achieve more accurate results with a limited number of points.

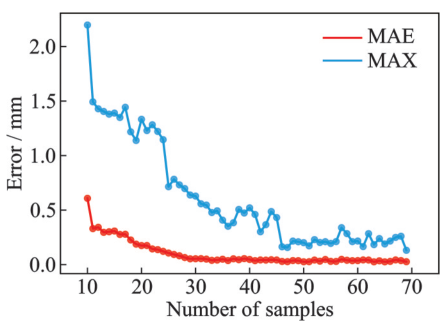
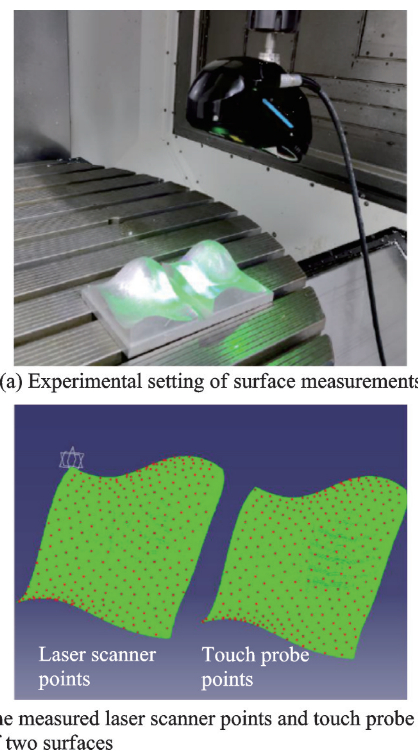


Fig.7 Change of MAE and MAX of multi-sensor reconstructed surface with the number of touch probe points

2.2 Measurement case study

This section verifies the proposed method in the actual measurement experiment of two freeform surfaces. Two surfaces and the measured point clouds are shown in Fig.8. The low-accuracy laser scanner data contains more than 100 000 points, and the high-accuracy touch probe points contain



(a) Experimental setting of surface measurements

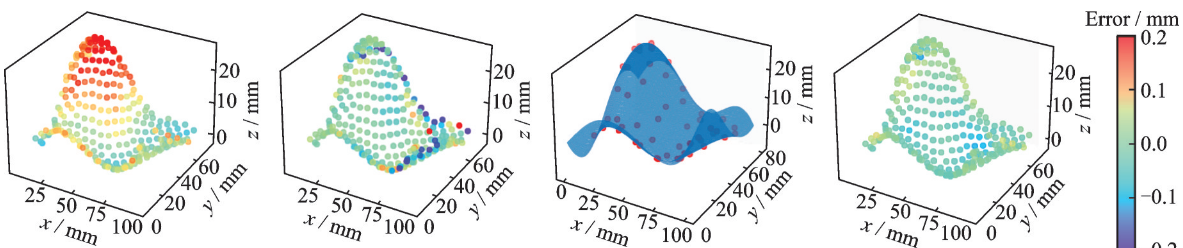
(b) The measured laser scanner points and touch probe points of two surfaces

Fig.8 Experimental settings and measured points of two surfaces

290 points. The details of the measurement experiment can be found in Ref.[27]. The multi-sensor measurement experiment is carried out from 10 touch probe points, actively sampling until 40 touch probe points.

This case emphasizes the maximum reconstruction error of the surface, as this criterion is of significant concern in real engineering applications. Fig.9 presents comparison error maps of reconstructions using only low-accuracy data (Fig.9(a)), only high-accuracy data (Fig.9(b)), and the active sampling multi-sensor reconstruction (Fig.9(d)). The error maps are derived from 260 touch probe test points. The distribution of the 40 sampled touch probe points is shown in Fig.9(c).

Fig.9(a) shows an overall error at the top of the surface, indicating either the systematic error of the laser scanner or the general learning error of the GP model. The maximum reconstruction error can be reduced from 0.438 mm to just 0.114 mm, demonstrating that the proposed active sampling-based multi-sensor fusion can achieve more accurate surface measurements and reconstructions than using only laser scanner data or touch probe data.



(a) D_l: MAX=0.226 mm (b) D_h: MAX= 0.438 mm (c) Sampled points (d) Fusion: MAX=0.114 mm

Fig.9 Comparison of experimental results for multi-sensor measurement for one of the experimental surfaces

3 Conclusions

The measurement of freeform surfaces plays a crucial role in quality control, reverse engineering, and various industrial applications. The integration of data from laser scanners and touch probes in multi-sensor measurement has garnered considerable interest because of the complementary features of these sensors. Considering that the distribution of sampling points affects the accuracy and efficiency of multi-sensor fusion, this study introduces an aggregation-value-based active sampling method for multi-sensor measurement, which actively generates touch probe sampling points on the surface based on combinational game theory. The main contributions of this research include:

(1) Introducing the concept of aggregation value to describe the added value of each point in surface reconstruction, guiding the active sampling of touch probe points.

(2) Combining the classical residual approximation method with the active sampling method to iteratively sample new points during multi-sensor measurement.

(3) Demonstrating through a toy case study and actual experiments that the proposed active sampling-based multi-sensor fusion can achieve more accurate surface measurements and reconstructions.

While this research focuses solely on the combination of active learning with residual approximation, future studies could explore active sampling techniques with weighted fusion, implicit fusion, and other multi-sensor fusion methods.

References

[1] QIE Y, SCHLEICH B, ANWER N. Generative adversarial networks and hessian locally linear embedding for geometric variations management in manufacturing[J]. Journal of Intelligent Manufacturing, 2023, 1: 1-30.

[2] VLAEYEN M, HAITJEMA H, DEWULF W. Digital twin of an optical measurement system[J]. Sensors, 2021, 21(19): 6638.

[3] QIE Y, NICQUEVERT B, ANWER N. Functional specification of complex assemblies using projective geometric algebra[J]. CIRP Annals, 2024, 73(1): 105-108.

[4] LI F, LONGSTAFF A P, FLETCHER S, et al.

Rapid and accurate reverse engineering of geometry based on a multi-sensor system[J]. The International Journal of Advanced Manufacturing Technology, 2014, 74: 369-382.

[5] ZHU Z, AREZKI Y, CAI N, et al. Data fusion-based method for the assessment of minimum zone for aspheric optics[J]. Computer-Aided Design and Applications, 2020, 18(2): 309-327.

[6] HE G, SANG Y, PANG K, et al. An improved adaptive sampling strategy for freeform surface inspection on CMM[J]. The International Journal of Advanced Manufacturing Technology, 2018, 96: 1521-1535.

[7] SADAOUI S E, MEHDI-SOUZANI C, LARTIGUE C. Multisensor data processing in dimensional metrology for collaborative measurement of a laser plane sensor combined to a touch probe[J]. Measurement, 2022, 188: 110395.

[8] LU K, WANG W. A multi-sensor approach for rapid and precise digitization of free-form surface in reverse engineering[J]. The International Journal of Advanced Manufacturing Technology, 2015, 79: 1983-1994.

[9] LIU X, CHEN G, LI Y, et al. Sampling via the aggregation value for data-driven manufacturing[J]. National Science Review, 2022, 9(11): nwac201.

[10] MIRZASOLEIMAN B, BILMES J, LESKOVEC J. Coresets for data-efficient training of machine learning models[C]//Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2020: 6950-6960.

[11] ELHAMIFAR E, SAPIRO G, SASTRY S S. Dissimilarity-based sparse subset selection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(11): 2182-2197.

[12] SUN L, ZHANG W, REN J, et al. A Gaussian process guided super resolution sampling strategy for the efficient recovery of complex surfaces[J]. Measurement Science and Technology, 2023, 34(6): 065013.

[13] TANG Y, FENG J. Multi-scale surface reconstruction based on a curvature-adaptive signed distance field[J]. Computers & Graphics, 2018, 70: 28-38.

[14] ABDULHAMEED O, MIAN S H, AL-AHMARI A, et al. Patch and curvature specific estimation of efficient sampling scheme for complex surface inspection[J]. The International Journal of Advanced Manufacturing Technology, 2020, 110: 3407-3422.

[15] YI B, QIAO F, HUANG N, et al. Adaptive sampling point planning for free-form surface inspection under multi-geometric constraints[J]. Precision Engineering, 2021, 72: 95-101.

[16] PAGANI L, SCOTT P J. Curvature based sampling of curves and surfaces[J]. Computer Aided Geometric Design, 2018, 59: 32-48.

[17] ZHANG Y, HUANG N, HOU L, et al. Intelligent sampling strategy for freeform surfaces based on on-machine measurement through real-time interactive communication[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-8.

[18] THARWAT A, SCHENCK W. A survey on active learning: State-of-the-art, practical challenges and research directions[J]. Mathematics, 2023, 11(4): 820.

[19] YIN Y, REN M J, SUN L. Dependant Gaussian processes regression for intelligent sampling of freeform and structured surfaces[J]. CIRP Annals, 2017, 66(1): 511-514.

[20] YU M, ZHANG Y, LI Y, et al. Adaptive sampling method for inspection planning on CMM for free-form surfaces[J]. The International Journal of Advanced Manufacturing Technology, 2013, 67: 1967-1975.

[21] XU P, JI X, LI M, et al. Small data machine learning in materials science[J]. npj Computational Materials, 2023, 9(1): 42.

[22] KULICHENKO M, BARROS K, LUBBERS N, et al. Uncertainty-driven dynamics for active learning of interatomic potentials[J]. Nature Computational Science, 2023, 3(3): 230-239.

[23] ZHANG J, SUN Q, LIU J, et al. Efficient sampling approaches to Shapley value approximation[J]. Proceedings of the ACM on Management of Data, 2023, 1(1): 1-24.

[24] ROZEMBERCZKI B, WATSON L, BAYER P, et al. The Shapley value in machine learning[EB/OL]. (2022-02-11). https://doi. org/10.48550/arXiv. 2202. 05594.

[25] OBEIDAT S M, RAMAN S. An intelligent sampling method for inspecting free-form surfaces[J]. The International Journal of Advanced Manufacturing Technology, 2009, 40: 1125-1136.

[26] NEMHAUSER G L, WOLSEY L A, FISHER M L. An analysis of approximations for maximizing submodular set functions—I[J]. Mathematical Programming, 1978, 14: 265-294.

[27] WANG J, PAGANI L, LEACH R K, et al. Study of weighted fusion methods for the measurement of surface geometry[J]. Precision Engineering, 2017, 47: 111-121.

Acknowledgements This work was supported by the National Key R&D Program of China (No.2022YFB3402600), the National Science Fund for Distinguished Young Scholars (No. 51925505), the General Program of National Natural Science Foundation of China (No. 52275491), and Joint Funds of the National Natural Science Foundation of China (No.U21B2081).

Authors Dr. **CHEN Gengxiang** received his Ph.D. degree in Aeronautics and Astronautics Manufacturing Engineering from Nanjing University of Aeronautics and Astronautics (NUAA), and his Ph.D. degree in Mechanical Engineering from Université Paris-Saclay in 2023. He is currently a post-doctoral researcher at Université Paris-Saclay. His research interests include metrology, smart manufacturing and data-driven predictive modeling.

Prof. **LI Yingguang** received the B. S. degree in Materials Engineering and the Ph.D. degree in Aeronautics and Astronautics Manufacturing Engineering from NUAA in 1999 and 2004, respectively. Since 2008, he has been a Professor with NUAA. His research interests include data-driven intelligent manufacturing and advanced composite curing technology.

Author contributions Dr. **CHEN Gengxiang** developed the methodology, conducted the experiments, and drafted the manuscript. Prof. **LI Yingguang** contributed to the conceptualization and research funding. Prof. **MEHDI-SOUZANI Charyar** assisted with data analysis and experiment design. Prof. **LIU Xu** performed the analysis and revised the manuscript. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: WANG Jing)

面向多传感器自由曲面测量与重构的聚合价值主动采样方法

陈耿祥[1,2]，李迎光[1]，MEHDI-SOUZANI Charyar[2]，刘 旭[3]

（1.南京航空航天大学机电学院，南京 210016，中国；2.巴黎萨克雷大学，伊维特河畔吉夫 91190，法国；3.南京工业大学机械与动力工程学院，南京 211816，中国）

摘要：自由曲面测量是航空航天领域产品质量控制与逆向工程等的关键基础技术。基于激光扫描仪、接触式探针等多传感器融合的曲面测量技术能够结合不同传感器的互补特性，目前在工业界与学术界广受关注，测量点的数量与分布会显著影响多传感器融合的效率和曲面重建的精度。提出了一种面向多传感器融合自由曲面测量与重构的聚合价值采样方法，基于博弈论迭代式主动生成探针测量点，将自由曲面上每个测量点对多传感器融合的重要性显示地定义为测量点的Shapley值，从而将最优测量点集合获取问题转换为样本集合的聚合价值最大化问题。仿真和真实测量结果验证了本方法能够在保证多传感器融合测量精度的情况下显著降低所需的探针样本量。

关键词：多传感器融合；多传感器测量；数据采样；主动学习；Shapley值；智能采样