DOI: 10. 13382/j. jemi. B2205922

基于高斯过程模型的多源点云数据融合方法*

师　乐[1]　罗　钧[1]　何小妹[2]

（1. 重庆大学光电技术及系统教育部重点实验室　重庆　400030；2. 航空工业北京长城计量测试技术研究所　北京　100095）

摘　要：多传感器测量技术被认为是表面计量学中一个很有效的解决方案。针对多源数据的融合问题，本文提出了一种基于高斯过程模型的多源点云数据融合框架。首先，提出一种自适应距离的鲁棒点云配准方法统一不同测量数据集的坐标系；然后，通过引入平差理论，对来自不同传感器的多个独立数据集之间的残差进行近似，构建基于 Matern 核函数的高斯过程模型；最后，通过仿真模拟和实际应用，与现有方法进行了一系列对比实验，验证了该方法的有效性。实验结果表明，该方法能以更高的融合精度和更快的计算效率融合多传感器数据集。

关键词：多传感器测量；数据融合；自适应距离；点云配准；高斯过程

中图分类号：TP274　　**文献标识码**：A　　**国家标准学科分类代码**：460. 40

Multi-source point cloud data fusion method based on Gaussian process model

Shi Le[1]　Luo Jun[1]　He Xiaomei[2]

(1. Key Laboratory of Optoelectronic Technology and System of Ministry of Education, Chongqing University, Chongqing 400030, China; 2. AVIC Changcheng Institute of Metrology & Measurement, Beijing 100095, China)

Abstract: Multi-sensor measurement technology is considered to be a very effective solution in surface metrology. Aiming at the problem of modeling and fusion of multi-scale complex data sets, this paper proposes a multi-source point cloud data fusion framework based on Gaussian process. Firstly, a robust point cloud registration method with adaptive distance is proposed to unify coordinate systems of different measurement datasets. Then, by introducing adjustment theory, the residuals between multiple independent data sets from different sensors are approximated, and a Gaussian process model based on Matern kernel function is constructed. Finally, the method is verified by simulation verification and practical application, and a series of comparative experiments with existing methods are carried out to verify the effectiveness of the method. The results show that the method can fuse multi-sensor datasets with higher fusion accuracy and faster computational efficiency.

Keywords: multi-sensor measurement; data fusion; adaptive distance; point cloud registration; Gaussian process

0　引　言

表面几何尺寸测量是精密工程领域最有影响力的学科之一。随着先进机械加工技术的发展使得复杂曲面能够在机械和光学工程中广泛应用。然而，对这些先进表面的几何复杂性需要完整的三维表征，这就在测量范围、分辨率、测量精度和测量效率等方面对当前的测量设备、测量技术提出了更高的要求[1]。

随着自由曲面的日益复杂，尽管这些仪器有其优点，但很少能够同时满足精度、效率、分辨率和测量范围方面的高要求。例如，接触式测量设备，三坐标测量机（coordinate measuring machine, CMM）通常具有较高的精度，但无法捕捉微观结构信息且测量效率较低[2]；非接触式测量设备，如结构光扫描仪，可以有效快速地生成密集的点云数据，从而很好地捕捉被测件的整体形状，然而不易测量具有尖锐特征的结构化表面，测量精度也远低于三坐标测量机。因此，多个测量传感器的有效组合可能

收稿日期：2022-10-18　　Received Date：2022-10-18

*基金项目：国家科技重大专项(J2019-VIII-0015-0176)项目资助

是解决复杂曲面精确测量任务的最佳方案，从而最大限度地发挥每个测量传感器的优势[3]。数据融合是继多传感器测量之后的一个关键问题，可以应用于质量检测、表面重建等领域。数据融合的目的是通过汇集来自不同传感器所接收到的多个信息，以输出新的数据信息组合[4-5]。输出数据的质量在很大程度上取决于数据融合方法。多传感器数据集通常位于不同的坐标系中，具有不同的测量范围、精度和分辨率，这给多传感器数据融合带来了巨大的挑战。

根据多传感器测量的特点，数据融合的建模过程通常分为两个步骤：点云配准和点云融合。

点云配准，是通过寻找刚体变换矩阵，将不同参考坐标系中的多个点云数据变换到同一坐标系中。Besl 等[6]提出的迭代最近点（iterative closest point, ICP）算法是应用最广泛的点云配准算法，是一种基于迭代的方法。该方法首先找到最近点，然后建立优化模型，使误差函数最小化，从而得到更新后的变换参数。当迭代次数达到给定值或误差小于阈值时，迭代过程停止。然而，ICP 采用的点-点距离函数具有线性收敛性，容易陷入局部最优，对异常值和重叠很敏感，而且存在收敛速度较慢、耗时长的问题。为了解决 ICP 算法存在的问题，一些学者提出了其改进算法。Pottmann 等[7-8]引入了平方距离最小化（squared distance minimization, SDM）方法和切面平方距离最小化（tangent-squared distance minimization, TDM）方法，这两种方法都是基于点-切面距离的函数，有效提高了配准速度。然而，在实际应用中，由于低密度测点的平方距离和在目标函数中所占比例较小，会引起匹配失真问题[9]。Rusinkiewicz[10]在 ICP 算法的基础上使用距离或法线约束来过滤匹配过程中的异常值，但是约束参数很难调整，参数不同对结果的影响也不同。最近，一种快速、鲁棒的 RICP 算法被提出[11]，该方法引入 Welsch 函数以提高 ICP 的鲁棒性，加快了配准速度。然而，基于 Welsch 函数的方法只能抑制噪声和异常值的负面影响，而不能完全消除。因此，本文针对以上配准问题，提出了一种自适应距离（adaptive distance, AD）的方法。该方法以测点距离与均值距离的残差为基础构造目标函数，该方法收敛速度快，稳定性好，可以解决配准向高密度点偏移的问题。

点云融合，主要处理配准后的公共参考坐标系中的多源数据集，将具有不同分辨率和不同不确定性的所有数据集融合成一个高质量数据集，以构建唯一的曲面模型[12]。Jamshidi 等[13]提出了一种在数据配准后通过消除异常值并对所有点云数据进行网格化的方法来实现高、低分辨率数据的融合。然而，这种方法只能替换来自不同来源的局部数据，并不能提高整体数据的准确性。加权融合方法具有简单实现、运算速度快等优点，也已经在许多研究中使用，但是这种方法削弱了细节信息，难以取得满意的融合效果。Ren 等[14]提出了一种基于加权最小二乘的融合方法，用于自由曲面的精确测量，使用 B 样条曲面拟合不同传感器数据集的线性曲面模型。该方法依赖于曲面模型拟合的精度，并不适用于多个复杂曲面。基于现代统计理论的数据融合方法备受研究者的关注[15]，其核心思想是使用统计理论对来自不同传感器的多个测量数据集建立回归模型，然后预测每个点云的最终位置。Xia 等[16]介绍了一种基于贝叶斯层次（Bayesian hierarchical, BH）模型的高、低精度点云融合算法。该算法可以对不同数据集的对应点建立多层贝叶斯参数估计模型，通过求解 BH 模型得到融合后的最优估计值。然而，该模型没有考虑不同的配准方法对融合质量的影响。

最近，高斯过程（Gaussian process, GP）模型也被广泛用于解决多源传感器数据融合问题[17-18]。GP 模型是一种非参数的贝叶斯统计推理模型，是通过均值和协方差函数对各种复杂曲面建模的强大数学工具。GP 模型的统计特性使其能够将测量误差纳入建模过程，并为构建的模型分配可信度。Colosimo 等[19]和 Castillo 等[20]利用 GP 模型建立了高精度数据集和低精度数据集的重建模型，以获得优化的融合数据集。

基于以上研究现状，本文提出了一种基于高斯过程模型的多传感器测量数据融合方法，应用于从同一复杂曲面上获取的多源三维点云数据的融合，以实现复杂曲面的精确测量，其中测量数据来自三坐标测量机和光扫描仪。首先，采用基于 AD 的鲁棒配准算法统一坐标系，然后建立 GP 模型融合来自不同传感器的多个数据集。该方法解决了多源传感器数据融合中的关键问题，包括提高点云配准的准确性和鲁棒性，降低融合过程的不确定性和计算成本。在仿真研究和实验测试的基础上，与现有方法进行了对比，验证了该方法的有效性。

1 基于高斯过程的多源数据融合测量方法

多传感器系统测量的数据集通常具有不同的分辨率和不确定性，并且位于不同的坐标系中。本研究的主要目标是将光扫描仪（高密度、低精度）获得的数据集与低密度、高精度（CMM）获得的“更真实”数据集进行融合，以提高数据集的局部确定性（减少局部偏差）。其中 CMM 测量点充当局部修正点，本质上是将局部偏差修正引入高密度点集中。因此，假设此数据融合方法中有两个可用数据集：

LP 数据集，由光扫描仪生成的低精度、高密度数据集，简称为 LP（low precision）。

HP 数据集，由 CMM 获取的高精度、低密度数据集，简称为 HP（high precision）。

为了解决数据融合中的关键问题,本文提出的融合过程旨在从以下两个步骤提高数据集的质量。

1）点云配准,将不同来源的数据集转换到同一坐标系。本文提出了 AD 的鲁棒配准方法,统一 LP 数据集和 HP 数据集的坐标系。

2）点云融合,减少数据集的不确定性和计算成本。建立基于 GP 的多层融合模型,实现点云数据的融合,以 HP 数据集为参考点,修正 LP 数据集的局部偏差。

具体融合过程如图 1 所示。

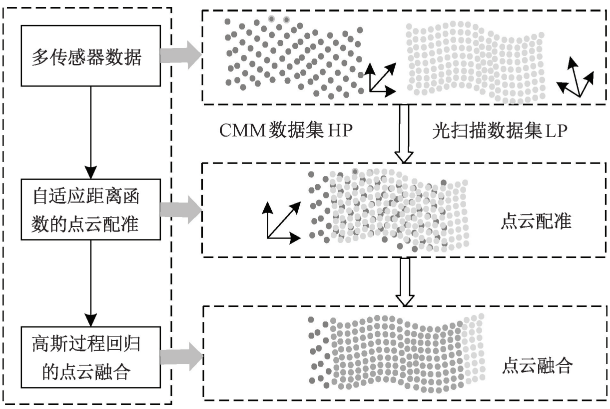


图 1　多传感器数据融合方法框图

Fig. 1　Block diagram of multi-sensor data fusion method

2　基于自适应距离的点云配准

点云配准,是通过求解坐标转换关系,将不同参考坐标系中的多传感器数据集转换到同一坐标系中。本文提出了一种基于 AD 的点云配准方法。

2.1　目标函数建立

如图 2 所示,点集 $P=\{p_i,i=1,2,\cdots,n\}$ 为 LP 数据集,点集 $Q=\{q_j,j=1,2,\cdots,m\}$ 为 HP 数据集。三维空间中,点云的鲁棒性配准是 0 找到一个刚体变换 $g=(\boldsymbol{R},\boldsymbol{t})\in \mathrm{R}^3$（包括旋转矩阵 $\boldsymbol{R}\in \mathrm{R}^3$ 和平移矩阵 $\boldsymbol{t}\in \mathrm{R}^3$）,将移动点集 P 匹配到参考点集 $\boldsymbol{Q}$。点 p_{i+} 是点 p_i 经过刚体变换 g 后的更新点。点 q_j 是点集 Q 中最接近点 p_i 的点,平面 T_j 为点 q_j 处的切平面。点集 Q 在点 q_j 处的法向量用符号 n_j 表示。点 a 为点 p_{i+} 在平面 T_j 上的投影。迭代匹配通过定义一个点-切面最小化距离函数计算三维配准的运动变量,以 $d_i=\|\overrightarrow{p_{i+}a}\|$ 为样本,构造残差距离 $d_i{}'=d_i-\bar{d}=d_i-\frac{1}{n}\sum_{i=1}^{n}d_i$,目标函数如下：

$$\min F(\boldsymbol{R},\boldsymbol{t})=\frac{1}{n}\sum_{i=1}^{n}(d_i-\frac{1}{n}\sum_{i=1}^{n}d_i)^2 \tag{1}$$

2.2　自适应距离函数

在本研究中,本文提出了一种 AD 函数来描述基本

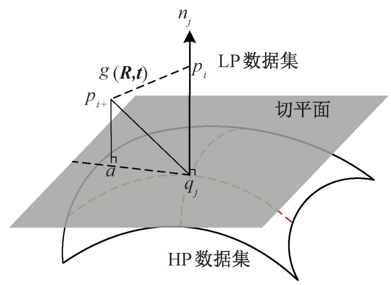


图 2　点 p_{i+} 与点集 Q 的距离 $d=\|\overrightarrow{p_{i+}a}\|$ 的几何表示

Fig. 2　Geometric representation of the distance $d=\|\overrightarrow{p_{i+}a}\|$ between point p_{i+} and point set Q

误差度量,式(1)中引入权重系数 μ：

$$\min F(\boldsymbol{R},\boldsymbol{t})=\frac{1}{n}\sum_{i=1}^{n}(d_i-\frac{\mu}{n}\sum_{i=1}^{n}d_i)^2 \tag{2}$$

当 $\mu=0$ 时,即为应用于 TDM 算法的点-切面平方距离最小化函数;当 $\mu=1$ 时,式(2)可以用方差最小化函数表示：

$$\min F(\boldsymbol{R},\boldsymbol{t})=\sum_{i=1}^{n}(d_i-\bar{d})^2= \sum_{i=1}^{n}d_i^2-\sum_{i=1}^{n}2d_i\bar{d}+\sum_{i=1}^{n}\bar{d}^2= \sum_{i=1}^{n}d_i^2-n\bar{d}^2=\sum_{i=1}^{n}d_i^2-\frac{1}{n}(\sum_{i=1}^{n}d_i)^2 \tag{3}$$

由此可见,以上两种常用的距离函数都设置权重系数 μ 为常量。然而,在实际迭代开始时,优化问题为一个大的残差问题,需要选择较大的权重系数,以保持一定的点-切面距离比,从而保证初始收敛的稳定性;经过一定次数的迭代后,残差逐渐减小,需要选择一个相对较小的权重系数来加快收敛速度,权重系数 μ 应能够随迭代次数变化动态调整。因此,本文提出将惯性权重的概念引入该算法,便可以针对不同程度的残差自动选择适当的权重系数计算切线距离,以反应曲率特性。其中,常用的惯性权重计算公式如下：

$$\mu(I)=\mu_{start}-(\mu_{star}-\mu_{end})\times I/I_{\max} \tag{4}$$

$$\mu(I)=\mu_{start}-(\mu_{star}-\mu_{end})(I/I_{\max})^2 \tag{5}$$

$$\mu(I)=\mu_{start}+(\mu_{star}-\mu_{end})[2I/I_{\max}-(I/I_{\max})^2] \tag{6}$$

式中：$I_{\max}$ 和 I 分别表示最大迭代次数和当前迭代。$\mu_{start}=1$ 为权重的初始值,$\mu_{end}=0$ 为迭代结束时的权重值。为了保证数据匹配的稳定性和收敛速度,本文结合现有的优化算法提出了一种新的非线性递减权重方法,表达式为：

$$\mu(I)=(\mu_{star}-\mu_{end})\exp(-2\pi I/I_{\max}) \tag{7}$$

式(7)表明了开始迭代时,权重系数较大且接近 1,目标函数近似方差最小化函数,此时残差距离 $d_i{}'$ 最小,

目标函数中各点贡献相同,配准不会向高密度点区域倾斜,一方面克服了测点密度不均对配准精度的影响,另一方面减少了噪声对配准精度的影响,保持稳定收敛;随着迭代次数增加,修正系数非线性减小并逐渐接近 0,具有 TDM 收敛快的特性,加快了收敛速度。

2.3 求解方法

求解基于自适应距离的目标函数,可以通过求解线性方程组来计算变换 $g = (\boldsymbol{R},\boldsymbol{t})$,类似于文献[8]。定义 $\boldsymbol{\omega} = [\delta_x,\delta_y,\delta_z]^{\mathrm{T}}$ 为旋转矩阵, $\boldsymbol{v} = [v_x,v_y,v_z]^{\mathrm{T}}$ 为平移矩阵,则 $p_{i+} = p_i + \boldsymbol{\omega}\times p_i + \boldsymbol{v}$,因此

$$d_i = \boldsymbol{n}_j^{\mathrm{T}}(p_{i+} - q_j) = d_{i0} + \boldsymbol{n}_j^{\mathrm{T}}(\boldsymbol{\omega}\times p_i + \boldsymbol{v}) = d_{i0} + \boldsymbol{A}_i\boldsymbol{\varepsilon} \tag{8}$$

其中, $\boldsymbol{\varepsilon} = [\boldsymbol{v},\boldsymbol{\omega}]^{\mathrm{T}}$, $d_{i0} = \boldsymbol{n}_j^{\mathrm{T}}(p_i - q_j)$, $\boldsymbol{A}_i$ 是 1×6 的矩阵, $\boldsymbol{\varepsilon} = [\boldsymbol{v},\boldsymbol{\omega}]^{\mathrm{T}}$ 是 6×1 的矩阵。将 $d_i = d_{i0} + \boldsymbol{A}_i\boldsymbol{\varepsilon}$ 代入式(2),目标函数可表示为:

$$\min F(\boldsymbol{R},\boldsymbol{t}) = \sum_{i=1}^{n}(d_{i0} + \boldsymbol{A}_i\boldsymbol{\varepsilon})^2 - \frac{1}{n}\left[\sum_{i=1}^{n}(d_{i0} + \boldsymbol{A}_i\boldsymbol{\varepsilon})\right]^2 = (D_1 + 2\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{\varepsilon}) + (D_2 + 2\boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{\varepsilon}) = D + 2\boldsymbol{C}^{\mathrm{T}}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^{\mathrm{T}}A\varepsilon \tag{9}$$

其中, $A_1 = \sum_{i=1}^{n}\boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{A}_i$, $A_2 = (\sum_{i=1}^{n}\boldsymbol{A}_i)^{\mathrm{T}}(\sum_{i=1}^{n}\boldsymbol{A}_i)$, $\boldsymbol{A} = A_1 + \boldsymbol{A}_2$ 是 6 × 6 的对称矩阵, $\boldsymbol{C}_1 = \sum_{i=1}^{n} d_i\boldsymbol{A}_i^{\mathrm{T}}$, $\boldsymbol{C}_2 = \sum_{i=1}^{n} d_{i0}\sum_{i=1}^{n}\boldsymbol{A}_i$, $\boldsymbol{C} = \boldsymbol{C}_1 + \boldsymbol{C}_2$ 是 6×1 的列向量, $D = D_1 + D_2$ 为标量。则参数 $\boldsymbol{\varepsilon} = [\boldsymbol{v},\boldsymbol{\omega}]^{\mathrm{T}}$ 可由下列线性方程求解:

$$\boldsymbol{A}\boldsymbol{\varepsilon} + \boldsymbol{C} = 0 \tag{10}$$

进而,变换矩阵 $\boldsymbol{g} = (\boldsymbol{R},\boldsymbol{t})$ 由下式计算:

$$R = e^{\boldsymbol{\omega}_0}, t = \boldsymbol{v} \tag{11}$$

式中: $\boldsymbol{\omega}_0$ 是旋转矩阵 $\boldsymbol{\omega}$ 的反对称矩阵。根据上述推导过程,可以推出该配准方法是采用 LM(Levenberg-Marquardt)算法求解非线性最小二乘问题,具有良好的收敛稳定性和二次收敛特性[21]。

基于自适应距离的点云配准方法流程如图 3 所示。

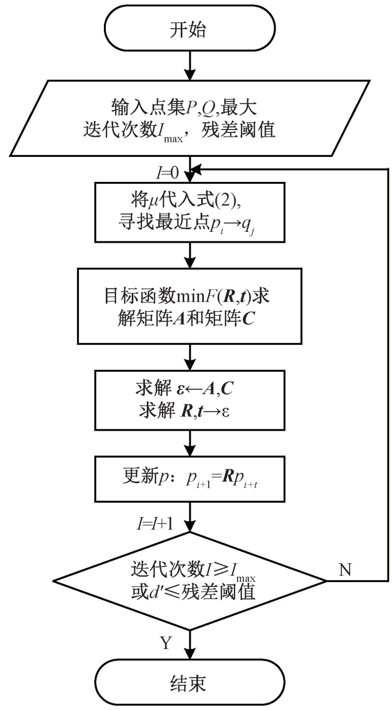


图 3 自适应距离的点云配准流程

Fig. 3 Point cloud registration flow chart with adaptive distance

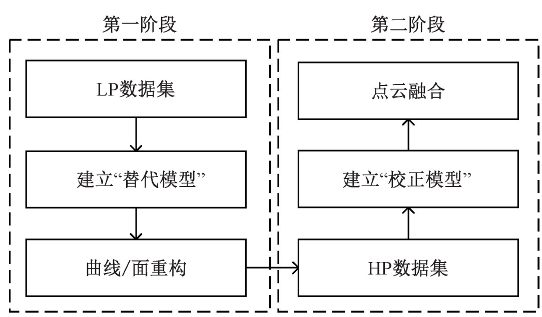


图 4 点云融合框图

Fig. 4 Block diagram of point cloud fusion

3 基于高斯过程模型的点云融合

点云融合是将配准到同一坐标系的不同数据集进一步融合校正,剔除异常和冗余并输出。点云融合需要两个阶段:1) 建立 LP 数据集的“替代模型”;2) 建立“校正模型”融合 LP 和 HP 数据集。具体如图 4 所示。

3.1 数据集的表示形式

假设用离散函数 $z(x_i,y_i)$ 表示点集,即第 i 点的 z 坐标表示其在 (x,y) 平面上的位置函数,这种-形式有利于 LP 数据集的表示。一方面,由于结构光扫描仪获取的范围图像数据集,在任何给定的 (x,y) 位置只能存在一个 z 值;另一方面,可以适应大多数 CMM 点集(HP 数据集),即能够在同一 (x,y) 平面上找到 HP 数据集的映射。因此,LP 数据集可表示为: $z_{LP}(x_i,y_i)$,其中, $(x_i,y_i) \in v_{LP} \subset \mathrm{R}^2, i = 1,2,\cdots,n_{LP}$;HP 数据集可表示为: $z_{HP}(x_i,y_i)$,其中 $(x_i,y_i) \in v_{HP} \subset \mathrm{R}^2, i = 1,2,\cdots,n_{HP}$ 。如前所述,LP 数据集大于 HP 数据集,即 $n_{LP} > n_{HP}$ 。通常,LP 数据集和 HP 数据集在平面 (x,y) 上的点的坐标位置不完全相同,即 $v_{LP} \cap v_{HP} = \varnothing$,因为光扫描成像设备所获取的某个点,CMM 触控探针很难在同一位置获取该点。

3.2 建立“替代模型”

由 LP 数据集可建立如下模型:

$$z_{LP}(x_i,y_i) = f_{LP}(x_i,y_i) + \varepsilon_{LP} \tag{12}$$

式中: $f_{LP}(x_i,y_i)$ 表示在完美拟合的理想情况下被测表面

的理想值，ε_{LP} 表示测量误差，假设误差服从 $\varepsilon_{LP} \sim N(0, \sigma^2_{\varepsilon_{LP}})$ 的正态分布。本文首先考虑使用高斯过程（GP）模型来表示 $f_{LP}(x_i, y_i)$：

$$f_{LP}(x_i, y_i) \sim N(m_{z_{LP}}(v_i), k_{z_{LP}}(v_i, v_j)) \tag{13}$$

其中，$v_i = (x_i, y_i)$，$v_j = (x_j, y_j)$；$m_{z_{LP}}(v_i) = E[z_{LP}(v_i)]$ 是均值函数，表示在点 v_i 处预测的 z 值；$k_{z_{LP}}(v_i, v_j) = E[(z_{LP}(v_i) - m_{z_{LP}}(v_i))(z_{LP}(v_j) - m_{z_{LP}}(v_j))]$ 是协方差函数，用于描述在点 $v_i = v_j$ 处 z 值的方差，以及在点 $v_i \neq v_j$ 处 z 值的协方差。在本文中，使用线性方程来表示 GP 模型的均值函数：

$$m_{z_{LP}}(v_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i \tag{14}$$

另外，本文选择 Matern 核函数来表示 GP 模型的协方差函数，即：

$$k_{z_{LP}}(v_i, v_j) = \sigma^2_{z_{LP}} \frac{2^{1-\lambda}}{\Gamma(\lambda)} \left(\frac{\sqrt{2\lambda}\, \| v_i - v_j \|}{l_{LP}} \right)^{\lambda} K\left(\frac{\sqrt{2\lambda}\, \| v_i - v_j \|}{l_{LP}} \right) \tag{15}$$

式中：$\| v_i - v_j \|$ 表示在 (x, y) 平面上点 v_i 和 v_j 之间的欧氏距离；$\sigma^2_{z_{LP}}$ 表示 GP 模型的常数方差；λ 是平滑系数，当 $\lambda = 0.5$ 时，Matern 核为指数核，当 $\lambda \to \infty$ 时，为高斯核；l_{LP} 是核函数的长度参数，即两个 z 值之间没有显著关系的长度；Γ 为 Gamma 函数，K 为第二类修正 Bassel 函数。Matern 核函数是 GP 模型最常用的选择之一，它产生的正定矩阵能够实现贝叶斯估计算法的适当收敛。

式（12）～（15）中，参数 $\{\sigma^2_{\varepsilon_{LP}}, \beta_0, \beta_1, \beta_2, \sigma^2_{z_{LP}}, l_{LP}\}$ 是未知的，需要通过测量值 $z_{LP}(x_i, y_i)$ 对参数进行估计。通常，超参数由最大似然法获得，即让已知数据评估的条件概率 $p(\theta \mid \mathbf{Z}, v_i)$ 达到最大，其中，θ 为超参数的统称，$\mathbf{Z}$ 为一组表示测量值 $z_{LP}(x_i, y_i)$ 的向量，$v_i = (x_i, y_i)$，$i = 1, 2, \cdots, n_{LP}$。这个过程相当于构造已知数据集的条件概率负对数 $L(\theta) = -\log p(z \mid v_i, \theta)$ 为目标函数，表达式如下：

$$L(\theta) = \frac{1}{2}\mathbf{Z}^{\mathrm{T}} k^{-1}_{z_{LP}} \mathbf{Z} + \frac{1}{2}\log | k_{z_{LP}} | + \frac{n_{LP}}{2}\log 2\pi \tag{16}$$

通过共轭梯度法对式（16）中的超参数求偏导并极大化，获得超参数的最优解。在参数估计完成后便可根据式（17）联合概率分布对任意位置进行预测。

$$\begin{bmatrix} z_{LP} \\ z'_{LP} \end{bmatrix} \sim N\left(\begin{bmatrix} m(v_r) \\ m(v_t) \end{bmatrix}, \begin{bmatrix} k(v_r, v_r) + \sigma^2_{\varepsilon_{LP}} I & k(v_r, v_t) \\ k(v_t, v_r) & k(v_t, v_t) \end{bmatrix} \right) \tag{17}$$

因此，预测值 z'_{LP} 的后验分布可以被描述为 $z'_{LP} \mid v_r, z_{LP}, v_t \sim N(\mu(v_t), c(v_t))$，其中，均值 $\mu(v_t)$ 和方差 $c(v_t)$ 表示为：

$$\mu(v_t) = k(v_t, v_r)[k(v_r, v_r) + \sigma^2_{\varepsilon_{LP}} I]^{-1} z_{LP} \tag{18}$$

$$c(v_t) = k(v_t, v_t) - k(v_t, v_r)[k(v_r, v_r) + \sigma^2_{\varepsilon_{LP}} I]^{-1} k(v_r, v_t) \tag{19}$$

式中：$\mu(v_t)$ 是最可能的预测，方差 $c(v_t)$ 表示置信度，即方差越大，GP 预测的不确定性越高。v_r 表示已测量的位置 (x_i, y_i)，v_t 表示需要预测的位置。

3.3　建立“校正模型”

如上所述，在 LP 数据集上构建的 GP 模型为在任意位置可以查询的连续函数。然而，该模型无法校正原始 LP 数据集中可能存在的局部偏差。另一组高精度、低密度点集 HP，在测量范围内具有高精度（即测量误差的方差较小）和无偏性（即平均测量误差为 0）。无偏性意味着，即使 LP 点集偏离被测表面的位置，HP 数据集仍能提供可靠的参考性。因此，本文提出一种“校正模型”，使 HP 高精度数据集所包含的信息能够校正 LP 数据集的局部偏差。

首先引入“校正模型”，即一个用来描述由 LP 数据集预测的 $z'_{LP}(x_i, y_i)$ 值与 HP 数据集的测量值 $z_{HP}(x_i, y_i)$ 之间差异的模型[22]，其表达式如下：

$$z_{HP}(x_i, y_i) = \rho(x_i, y_i) z'_{LP}(x_i, y_i) + \delta(x_i, y_i) + \varepsilon_{HP} \tag{20}$$

式中：$(x_i, y_i) \in \mathrm{v}_{HP} \subset \mathrm{R}^2$，$i = 1, 2, \cdots, n_{HP}$，$\rho(x_i, y_i)$ 和 $\delta(x_i, y_i)$ 分别表示缩放因子和移动因子，ε_{HL} 表示误差。由于“校正模型”是在统计基础上建立的，所以误差 ε_{HL} 表示 HP 数据集 $z_{HP}(x_i, y_i)$ 与 $\rho(x_i, y_i) z'_{LP}(x_i, y_i) + \delta(x_i, y_i)$ 之间的残差。假设残差 ε_{HL} 也服从正态分布：$\varepsilon_{HL} \sim N(0, \sigma^2_{\varepsilon_{HL}})$。

缩放因子 ρ 用线性模型表示：

$$\rho(x_i, y_i) = \rho_0 + \rho_1 x_i + \rho_2 y_i \tag{21}$$

移动因子 δ 由常数均值 δ_0 和协方差函数 $k_\delta(v_i, v_j)$ 组成：

$$\delta(v_i) \sim GP(\delta_0, k_\delta(v_i, v_j)) \tag{22}$$

其中，协方差 函数同样由 Matern 核函数表示：

$$k_\delta(v_i, v_j) = \sigma^2_\delta \frac{2^{1-\lambda}}{\Gamma(\lambda)} \left(\frac{\sqrt{2\lambda}\, \| v_i - v_j \|}{l_\delta} \right)^{\lambda} K\left(\frac{\sqrt{2\lambda}\, \| v_i - v_j \|}{l_\delta} \right) \tag{23}$$

式中：参数 $\{\sigma^2_{\varepsilon_{HL}}, \rho_0, \rho_1, \rho_2, \sigma^2_\delta, l_\delta\}$ 的计算过程同 3.2 节，需要由 CMM 测量值 $z_{HP}(x_i, y_i)$ 进行估计。所有未知参数成功估计后即可完成“校正模型”的建模，使原始 GP 模型（有偏差）倾向 HP 数据集（无偏差）。因此，在 (x, y) 平面上的任意位置 $v_i = (x_i, y_i)$ 的预测结果 $z'_{HL}(x_i, y_i)$ 可以表示为：

$$z'_{HP}(x_i, y_i) = \rho(x_i, y_i) z'_{LP}(x_i, y_i) + \delta(x_i, y_i) \tag{24}$$

3.4　高斯过程建模的具体步骤

基于高斯过程模型的点云融合方法具体步骤如下：

步骤 1）建立“替代模型”式（12），并选择合适的均值函数和协方差函数，根据核函数设置初始超参数；

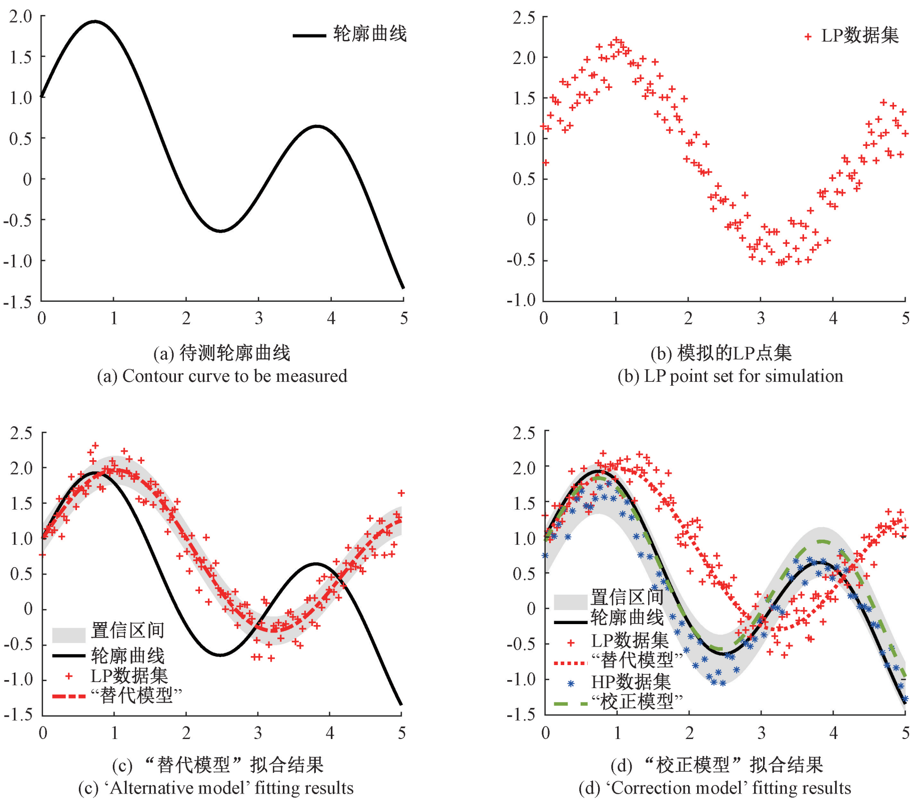
步骤 2)输入 LP 数据集,根据式(16),采用梯度法得到最优超参数,确定后验模型;

步骤 3)建立“校正模型”式(20),并选择合适的缩放因子和移动因子,设置初始超参数;

步骤 4)输入 HP 数据集,计算“校正模型”中的超参数。

步骤 5)输入需要预测的位置,通过式(24)获取最终预测值 $z'_{HP}(x_i,y_i)$,其中, $z'_{LP}(x_i,y_i)$ 是由步骤 2)中后验模型在相同位置获取的预测值。

为了进一步说明高斯过程在点云融合中是如何工作的,本文给出了一个简单的模拟示例,在通用性的情况下,以轮廓曲线为参考,待测量的轮廓曲线如图 5(a)所示。图 5(b)模拟光学测量仪器获取的低精度、高密度的 LP 点集,整个截面轮廓的精度较低,尤其是在边界处,光学像差等效应会引入较大偏差。图 5(c)为“替代模型”的拟合结果,虚线表示预测值 $z'_{LP}(x_i,y_i)$,灰色带表示 95%置信水平的预测区间。图 5(d)为“校正模型”的预估结果,其中,实线表示被测量的轮廓,由光成像原理测量的原始低精度、高密度的 LP 点集为十字形点,预测值 $z'_{LP}(x_i,y_i)$ 由点虚线表示,星形点表示原始高精度、低密度的 HP 点集,最终融合模型由长虚线(校正后的预测值)和相关预测区间(灰色带)表示。通过将图 5(d)与图 5(c)的结果进行比较,预测偏差有明显减小,修正后的曲线(长虚线)在真实曲线(实线)周围上下波动。



(a) 待测轮廓曲线
(a) Contour curve to be measured

(b) 模拟的LP点集
(b) LP point set for simulation

(c) “替代模型”拟合结果
(c) ‘Alternative model’ fitting results

(d) “校正模型”拟合结果
(d) ‘Correction model’ fitting results

图 5 基于 GP 的拟合结果

Fig. 5 The fitting results based on GP

4 实验验证与分析

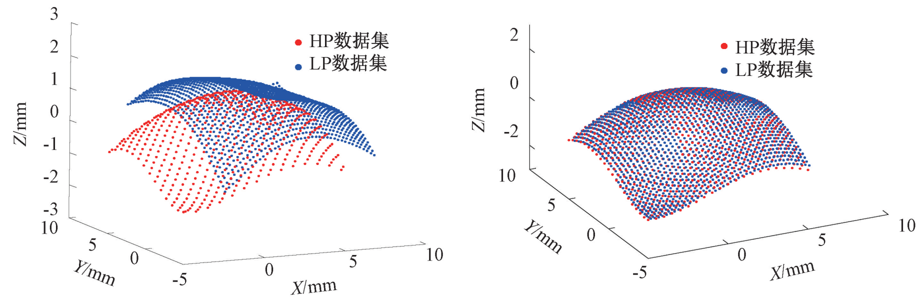
4.1 仿真验证

在仿真研究中,多源数据来自同一数据集,通过均匀采样两组数据模拟设计的自由曲面,分别由低精度(LP)和高精度(HP)表示。为了分析本文提出的 AD-GP 融合算法的性能,定义的自由曲面形式如下:

$$z = \sin(0.5x) + \cos(0.5y) \tag{25}$$

其中, $x,y \in [-2\pi,2\pi]$ mm,在此曲面上,以 0.4 mm 为间隔均匀采样一组点集,定义为 HP 数据集。另一组点集以 0.15 mm 为间隔,定义为 LP 数据集。在 HP 数据集中添加 5 μm 的高斯噪声,LP 数据集中添加 10 μm 的高斯噪声,以表示测量误差。将 LP 数据集按照给定的变换矩阵 $\boldsymbol{g}=(\boldsymbol{R},\boldsymbol{t})$ 移动到特定位置,以便两组点云处于不同的坐标系中。其中旋转参数 $\theta_z=30°$,即绕 z 轴旋转的角度,平移参数 $t_x=5$ mm, $t_y=5$ mm, $t_z=10$ mm。生成的曲面点云如图 6(a)所示。此时,两数据集具有不同的分辨率,不同的相关不确定性,以及处于不同的坐标系中。

首先,使用 AD 的点云配准算法对 LP 数据集和 HP 数据集进行配准,将二者转换到同一坐标系,缩小两片点云的旋转与平移错位。图 6(b)显示了配准后的结果。为了评估所提出算法的准确性和效率,与 TDM 方法、方差最小化方法、RICP 方法进行对比,变换参数误差和计算时间如表 1 所示。图 7 显示了 4 种方法在配准过程中对应点间的均方根误差(RMSE)随迭代次数的变化。该实验使用 MATLABR2016a 实现,设置最大迭代次数为 50,误差阈值为 0.1 μm。



(a) 模拟数据集
(a) Simulated data set

(b) 点云配准结果
(b) Point cloud registration result

图 6　模拟曲面的点云配准

Fig. 6　Point cloud registration of simulated surfaces

从表 1 配准结果对比可以看出,本文使用的自适应距离比 TDM、方差最小化、RICP 计算出的变换参数误差都小,表明自适应距离具有更高的配准精度。此外,该方法的计算成本与 RICP 接近,小于另两种算法。图 7 可以看出,自适应距离函数的收敛速度是 4 个算法中最快的,因为该点云配准是基于 Levenberg-Marquardt 方法实现的,在计算过程中用雅克比矩阵代替了二阶导数信息,使得优化效率有了很大提升。因此,基于自适应距离的点云配准方法对 HP 数据集和 LP 数据集的配准具有更高的精度和更快的收敛速度,并且受测量误差的影响较小,体现了该配准算法的鲁棒性。

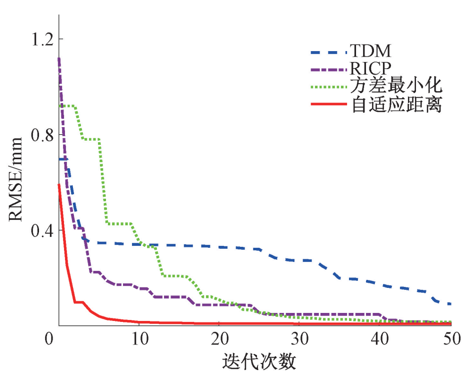


图 7　RMSE 随迭代次数的变化

Fig. 7　RMSE with iteration number

然后,根据第 3 节,构建两数据集重叠区域的点云融合模型,融合数据曲面如图 8(a)所示,图 8(b)显示了融合数据与设计曲面对比的形状误差。为了评估基于高斯过程融合方法的准确性,将其与基于加权最小二乘法(WLS)的融合方法进行比较测试,计算重建模型与设计曲面的平均绝对误差(MAE)和均方根误差(RMSE)。表 2 中 HP 数据集和 LP 数据集的误差值表明了与数据集中添加的高斯噪声是一致的。采用本文提出的 AD-GP 融合方法,融合后的数据平均绝对误差(MAE)降低到 5.7 μm,显著低于 HP 数据集和 LP 数据集的平均绝对误差,均方根误差(RMSE)降低到 4.6 μm,表明该方法能够提高测量数据集的准确性;表 2 中 AD-GP 与 AD-WLS 的数据表明,执行相同的点云配准方法后,基于 GP 的融合方法比基于 WLS 的融合方法精度更高,其中,MAE 和 RMSE 分别降低了 24% 和 29%;此外,AD-GP 分别与 RICP-GP、TDM-WLS 对比,也表明了 AD-GP 在点云配准与点云融合上都提供了更高的精度。一方面,由于 WLS 方法依赖于线性几何结构的数据集,当数据集发生急剧的几何变化时,不可避免的导致融合精度降低;另一方面,GP 的统计特性使得基于 GP 的融合方法在学习过程中能够有效避免噪声干扰。从表 2 的结果还可以看出,AD-GP 具有较快的计算效率仅次于 RICP-GP。

表 1　点云配准结果对比

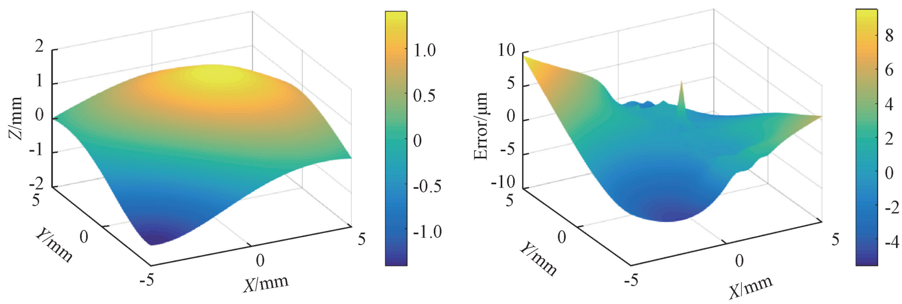
Table 1　Comparison of point cloud registration results

配准方法	变换参数误差				耗时/s
	θ_z	t_x	t_y	t_z	
TDM	0.152°	0.102 5	0.061 2	0.012 5	19.719 2
方差最小化	0.043°	0.060 7	0.030 3	0.004 6	18.156 3
RICP	0.113°	0.043 7	0.017 2	0.001 1	**16.415 6**
自适应距离	**0.009°**	**0.010 5**	**0.004 7**	**0.000 5**	**16.448 7**

表 2 数据融合结果对比

Table 2 Comparison of data fusion results

	HP 数据集	LP 数据集	TDM-WLS	AD-WLS	RICP-GP	AD-GP
MAE/μm	6.2	11.4	10.6	7.5	6.3	**5.7**
RMSE/μm	5.4	10.5	8.1	6.5	5.7	**4.6**
Time/s	—	—	10.5	9.2	**7.1**	7.3



(a) 融合结果
(a) The fusion results

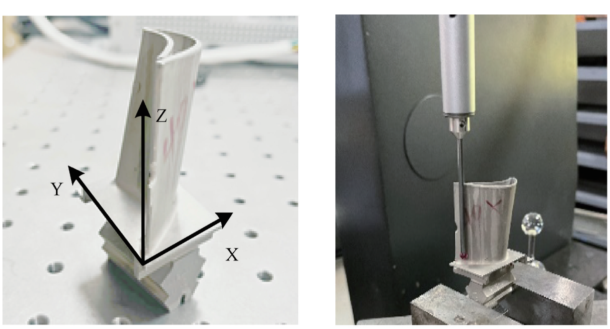
(b) 融合误差
(b) The fusion error

图 8 基于 GP 的数据融合结果

Fig. 8 Results of data fusion based on GP

4.2 实际测量

叶片是航空发动机的关键零部件，具有种类多、曲面形状多样、加工工艺复杂，加工难度大等特点，因此，对航空发动机叶片的测量精度要求较高[23]。以某航空发动机叶片为例，检测的主要目标为叶片型面上的参数，需要在截面线上采点，并进行曲线拟合。如图 9 所示，定义了工件坐标系。



(a) 工件坐标系
(a) Workpiece coordinate system

(b) 工件的CMM采样
(b) CMM sampling

图 9 叶型曲面测量过程

Fig. 9 The measurement process of blade surface

叶片分别通过两个计量装置进行采点，首先使用光扫描设备对该叶片进行扫描，并提取某一截面线的点云数据，作为 LP 数据集。另一组点由三坐标测量机(CMM)在同一截面线复制测量，作为 HP 数据集。其中，HP 数据集由 55 个点组成，LP 数据集由 1 303 个点组成。为了评估本文方法的融合性能，还有一组参考数据集作为实际值，用于验证。

为了满足叶片截面参数的测量需求，将本文提出的 AD-GP 方法应用于两数据集的融合。首先，本文提出的 AD 点云配准算法将两个数据集转换到公共坐标系，然后采用基于 GP 的融合方法生成融合数据集并输出，融合后的拟合曲线如图 10 所示。点线表示 LP 数据集的预测值，十字线表示通过 HP 数据集校正后的预测值。

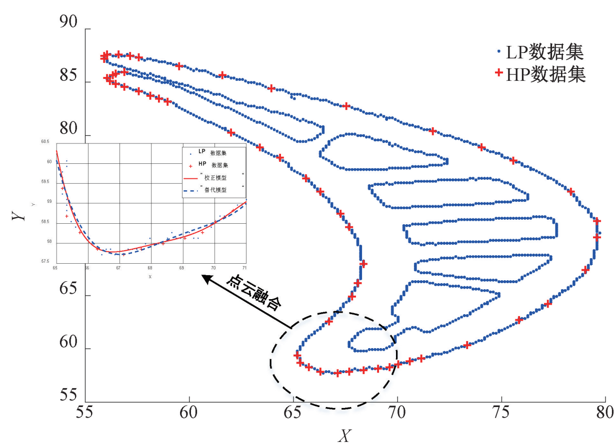


图 10 多传感器数据融合

Fig. 10 Multi-sensor data fusion

本文提出的 AD-GP 融合方法的评估结果如表 3 所示，包括 CMM 数据集，光扫描数据集，参考数据集和融合数据集。测量精度分别由平均绝对误差(MAE)和均方根误差(RMSE)表征。可以看出，本文方法生成的融合数据集与光扫描数据集相比，融合数据集的平均绝对误

差减小了 38.8%,均方根误差减小了 50.7%。同时,从表 3 还可以看出,融合数据集的误差值与参考数据集接近,表明融合数据集的数据质量已近似于参考数据集。

表 3　数据融合评估结果

Table 3　Data fusion evaluation results

	CMM	光扫描仪	参考集	AD-GP
MAE/μm	6.8	22.4	8.1	13.7
RMSE/μm	5.3	20.9	7.6	10.3

本节将对第 2 节中 4 种不同权重系数下的算法性能进行比较,验证该叶型点云数据融合的准确性。如图 11 所示,图中曲线分别表示式(4)~(7) 权重系数 μ 随迭代次数的变化过程,实验中其他因素保持不变,仅改变权重系数 μ 的计算方法,对比结果如表 4 所示。从表中可以看出,本文提出的权重系数计算方法在精度上优于常用方法。

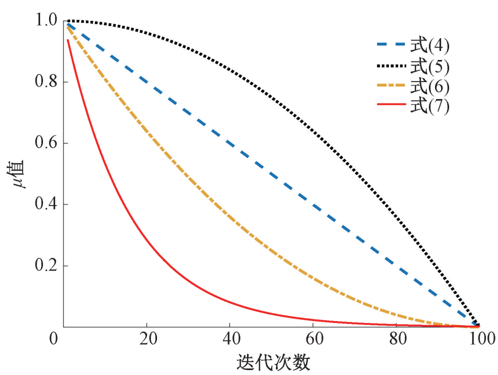


图 11　μ 随迭代次数的变化曲线

Fig. 11　The curve of μ with the number of iterations

表 4　不同权重系数下的算法性能比较

Table 4　Performance comparison of algorithms with different weight coefficients

方法	MAE/μm	RMSE/μm
式(4)	15.7	14.8
式(5)	17.1	15.2
式(6)	15.5	13.1
本文	**14.8**	**12.3**

另外,第 3 节中选择合适的核函数,对高斯过程模型的性能起着重要的作用。线性核、幂指数核以及本文使用的 Matern 核都是常用核函数,本节分别对基于以上 3 种内核的高斯过程模型的性能进行对比,验证基于 Matern 核函数的高斯过程模型对于该叶型点云数据融合的准确性更高。3 种核函数对模型性能评价结果如表 5 所示,本文使用的 Matern 核函数预测性能优于其他两种内核,MAE、RMSE 分别为 14.1 和 11.7,表现出了较高的准确性。图 12 为基于 3 种核函数建立的不同模型的预测结果与真实叶片截面数据的误差曲线,图 12 中可以看出本文使用的基于 Matern 核函数的误差曲线波动较小,且更接近真实值,进一步验证了 Matern 内核的稳定性。

表 5　不同核函数的算法性能对比

Table 5　Performance comparison of algorithms with different kernel functions

核函数	MAE/μm	RMSE/μm
线性核	133.2	117.3
幂指数核	41.7	36.4
Matern 核	**14.1**	**11.4**

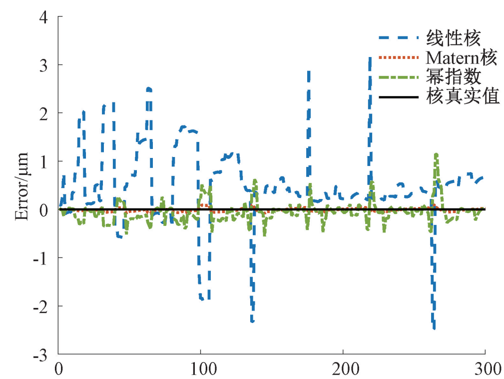


图 12　不同核函数的误差曲线对比

Fig. 12　Comparison of error of different kernel functions

5　结　论

本文提出了一种多传感器组合测量的数据融合方法,旨在将高精度、低密度点云(如 CMM 测量数据)与低精度、高密度点云(如光扫描仪测量数据)相结合,以获取更高精度的表面数据。为了解决多传感器融合的关键问题,首先采用自适应距离的配准方法来统一多数据集的坐标系,该方法以点-切面距离为基础,可线性求解刚体变换参数,受噪声影响较小;然后,引入基于 Matern 核函数的高斯过程模型获取高精度的融合数据集;最后,通过实验验证该方法的有效性。实验结果表明,本文提出的数据融合方法能够处理具有复杂几何结构的复杂曲面,实现了测量效率和融合性能的良好结合。

参考文献

[1] 张帅, 缪东晶, 李建双, 等. 多边法位姿测量系统中跟踪方式对测量精度的影响[J]. 计量学报, 2020, 41(9): 1055-1061.
ZHANG SH, MIAO D J, LI J SH, et al. Influence of tracking mode on measurement accuracy in multi-purpose pose measurement system [J]. Acta Meteorologica Sinica, 2020, 41(9):1055-1061.

[2] 何雪明, 孔丽娟, 何俊飞, 等. 基于三坐标测量机自适应测量的自由曲面逆向[J]. 机械工程学报, 2014, 50(15):155-159.

HE X M, KONG L J, HE J F, et al. Free-form surface reverse based on CMM self-adapting measurement[J]. Journal of Mechanical Engineering, 2014, 50(15): 155-159.

[3] SŁADEK J, BŁASZCZYK P M, KUPIEC M, et al. The hybrid contact-optical coordinate measuring system[J]. Measurement, 2011, 44(3):503-510.

[4] 罗俊海, 杨阳. 基于数据融合的目标检测方法综述[J]. 控制与决策, 2020(1):1-15.

LUO J H, YANG Y. An overview of target detection methods based on data fusion[J]. Control and Decision, 2020(1):1-15.

[5] 张品, 董为浩, 高大冬. 一种优化的贝叶斯估计多传感器数据融合方法[J]. 传感技术学报, 2014, 27(5):643-648.

ZHANG P, DONG W H, GAO D D. An optimal method of data fusion for multi-sensors based on bayesian estimation[J]. Journal of Sensors and Actuators, 2014, 27(5):643-648.

[6] BESL P J, MCKAY H D. A method for registration of 3-D shapes[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1992, 14(2):239-256.

[7] POTTMANN H, HUANG Q X, YANG Y L, et al. Geometry and convergence analysis of algorithms for registration of 3D shapes[J]. 2006, 67:277-296.

[8] POTTMANN H, LEOPOLDSEDER S, HOFER M. Registration without ICP[J]. Computer Vision & Image Understanding, 2004, 95(1):54-71.

[9] 杨晗, 谭川东, 盛晋银, 等. 基于工业 CT 的零件内外曲面形位误差分析[J]. 仪器仪表学报, 2021, 42(11):230-238.

YANG H, TAN CH D, SHENG J Y, et al. Form and position error analysis of inner and outer surface of parts based on the industrial CT[J]. Chinese Journal of Scientific Instrument, 2021, 42(11):230-238.

[10] RUSINKIEWICZ S. A symmetric objective function for ICP[J]. ACM Transactions on Graphics, 2019, 38(4): 1-7.

[11] ZHANG J, YAO Y, DENG B. Fast and robust iterative closest point[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7):3450-3466.

[12] DING J, LIU Q, BAI M X, et al. A multisensor data fusion method based on gaussian process model for precision measurement of complex surfaces[J]. Sensors, 2020, 20(1):278.

[13] JAMSHIDI J, OWEN G W, MILEHAM A R. A new data fusion method for scanned models[J]. Journal of Computing & Information Science in Engineering, 2006, 6(4):340-348.

[14] REN M J, LIU M Y, CHEUNG C F, et al. Multi-sensor data fusion for measurement of complex freeform surfaces[C]. International Symposium on Precision Mechanical Measurements. International Society for Optics and Photonics, 2016.

[15] SUN S, LIN H, MA J, et al. Multi-sensor distributed fusion estimation with applications in networked systems[J]. Information Fusion, 2017, 38(11):122-134.

[16] XIA H F, DING Y, WANG J. Gaussian process method for form error assessment using coordinate measurements[J]. IIE Transactions, 2008, 40(10):931-946.

[17] SU G, JIANG J, YU B, et al. A Gaussian process-based response surface method for structural reliability analysis[J]. Structural Engineering & Mechanics, 2016, 56(4): 549-567.

[18] 赖朝安, 龙漂. 基于高斯过程回归和 WiFi 指纹的室内定位方法[J]. 电子测量与仪器学报, 2021, 35(3):158-146.

LAI CH AN, LONG P. Indoor localization method based on gaussian process regression and WIFI fingerprint[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(3):158-146.

[19] COLOSIMO B M, PACELLA M, SENIN N. Multisensor data fusion via Gaussian process models for dimensional and geometric verification[J]. Precision Engineering, 2015, 40:199-213.

[20] CASTILLO E D, COLOSIMO B M, TAJBAKHSH S D. Geodesic Gaussian processes for the parametric reconstruction of a free-form surface[J]. Technometrics, 2015, 57:87-99.

[21] FAN J, PAN J. On the convergence rate of the inexact Levenberg-Marquardt method[J]. Journal of Industrial & Management Optimization, 2017, 7(1):199-210.

[22] XIA H F, DING Y, MALLICK B K. Bayesian hierarchical model for combining misaligned two-resolution metrology data[J]. IIE Transactions, 2011, 43(4):242-258.

[23] 万能,庄其鑫,郭彦亨,等. 拟合精度约束下航发叶片在机测量采样策略[J/OL]. 航空学报:1-13[2023-01-05].

WAN N, ZHUANG Q X, GUO Y H, et al. Sampling strategy for on-machine measurement of aero-engine blade under constraint of fitting accuracy[J/OL]. Acta Aeronautica et Astronautica Sinica:1-13[2023-01-05].

作者简介

师乐，2019 年于西安工程大学获得硕士学位，现为重庆大学博士研究生，主要研究方向为精密测量仪器。

E-mail：lok_si0608@163.com

Shi Le received her M. Sc. degree from Xi′an Polytechnic University in 2019. Now she is a Ph. D. candidate in Chongqing University. Her main research interest includes precision measuring instruments.

罗钧（通信作者），1983 年于重庆大学获得学士学位，1993 年于重庆大学获得硕士学位，现为重庆大学教授，博导，主要研究方向为人工智能、精密机械及测试计量、智能信息处理。

E-mail：luojun@cqu.edu.cn

Luo Jun (Corresponding author) received his B. Sc. degree from Chongqing University in 1983, received his M. Sc. degree in 1993 from Chongqing University. Now he is a professor in Chongqing University. His main research interests include artificial intelligence, precision machinery and measurement, intelligent information processing.

何小妹，2005 年于北京理工大学获得博士学位，现为中国航空工业集团公司北京长城计量测试技术研究所研究员，主要研究方向为复杂曲面、线纹和角度等几何量计量技术研究。

E-mail：hexiaomei2000@163.com

He Xiaomei received her Ph. D. degree from Beijing Institute of Technology in 2005. Now she is a researcher in AVIC Changcheng Institute of Metrology & Measurement. Her main research interests include geometric measurement technology of complex surface, line and angle.