

光学学报

基于多层感知机的金属表面光度-结构光测量数据融合方法

杨沅松, 王晰, 任明俊*

上海交通大学机械与动力工程学院, 上海 200240

摘要 提出一种基于多层感知机的金属表面光度-结构光测量数据融合算法。设计了融合面结构光与光度立体视觉两种测量原理的复合传感器, 分别获得同一相机坐标系下的缺失点云及完整法向量多模态数据。为有效融合两种数据, 设计了基于位置编码的多层感知机网络, 以点云为形状约束, 以法向量为纹理约束, 实现金属表面完整的高精度三维重建。通过仿真和实验验证了所提方法的有效性。实验结果表明: 该方法有效避免了法向量积分累计误差的问题, 在获取完整点云的同时, 测量精度相较于结构光测量系统提升了 50.4%。

关键词 测量; 结构光; 光度立体视觉; 视觉测量; 复合传感器; 多层感知机

中图分类号 TP274 **文献标志码** A

DOI: 10.3788/AOS230497

1 引言

在工业制造领域, 机械加工零件表面的高精度重建对加工过程的监控和质量评定起到至关重要的作用, 目前工业上常用的表面三维重建方法可以分为接触式和非接触式两大类。常用的接触式三维重建方法为三坐标测量机(CMM), 其具有高测量精度、强鲁棒性的优点, 但该方法具有测量范围受限和测量效率较低的缺陷。近年来, 非接触式三维重建方法由于具有成本低廉、测量速度快、精度高的优点, 被广泛应用于机械加工表面测量。常用的非接触式三维测量方法包括激光扫描、面结构光投影、光度立体视觉等。

面结构光测量方法是依据三角测量的原理实现物体表面三维重建的^[1], 该方法将特定的图像经投影仪投射至待测物体表面, 所投射的图像受到物体表面高度信息的调制, 并通过系统标定参数实现物体表面三维坐标的计算。尽管该方法可以有效获取待测物体的面形信息, 但是点云表面易出现高频噪声, 难以实现对纹理丰富构件的有效测量^[2]。此外, 由于结构光系统在测量过程中依赖主动式光源投射, 受金属表面复杂高光反射现象影响, 反光区域的调制度降低, 点云精度及完整度急剧下降。研究人员针对高亮金属表面开展了大量的研究^[3], 提出多重曝光法^[4]、自适应调制法^[5]等方法, 但这些方法均在一定程度上存在高亮区域点云恢复有限、测量效率低的问题。

光度立体视觉系统通常由一台相机和多个不同位置的光源组成, 根据物体在不同光照方向下的光反射情况, 实现表面法向量估计^[6], 并通过法向量积分重建物体表面深度信息^[7]。光度立体视觉方法对于物体表面深度变化的分辨率较高, 可以实现精细纹理重构^[8], 但在法向量积分的过程中容易引起误差累积, 导致重建形状扭曲。该方法不具备形状恢复能力, 大多用于物体表面缺陷检测, 例如钢板缺陷检测^[9]、织物瑕疵检测^[10]、子弹缺陷检测^[11]等。

针对以上问题, 本文提出一种基于多层感知机的金属表面光度-结构光测量数据融合方法。利用光度-面结构光复合传感器同时获取待测物体表面的点云与法向量信息。点云中包含较为准确的形状信息, 但测量噪声大且容易造成数据缺失; 表面法向量包含的信息相对完整、精度较高, 但缺少尺度与位置信息, 且法向量积分易造成形状扭曲。为有效融合点云与法向量信息, 本文提出一种基于位置编码的多层感知机网络, 在网络训练过程中引入点云形状约束与法向量纹理约束, 补充缺失的点云信息, 减小高频噪声, 从而实现金属表面完整高精度三维重建。

2 光度-结构光复合传感器测量原理

光度-结构光复合传感器融合光度立体视觉与面结构光两种测量原理, 能够在同一相机坐标系下完成点云与法向量信息的提取。面结构光系统利用投影仪

收稿日期: 2023-02-03; 修回日期: 2023-03-21; 录用日期: 2023-04-20; 网络首发日期: 2023-05-08

基金项目: 国家自然科学基金面上项目(52175477)

通信作者: *renmj@sjtu.edu.cn

向场景中投射条纹,获取物体表面点云信息;光度立体视觉系统利用同一相机在不同光照方向下拍摄的物体图片实现表面法向量估计。两种测量方式共用一台相机,即获得同一相机坐标系下的点云与法向量结果,避免了坐标转换带来的额外误差。

2.1 面结构光系统测量原理

面结构光系统通过投影仪向待测物体表面投射一系列正弦条纹,按照时间顺序解析同一相机像素坐标的灰度值变化来计算条纹相位,并依据相位值恢复物体表面深度。图 1 为条纹投影系统示意图。

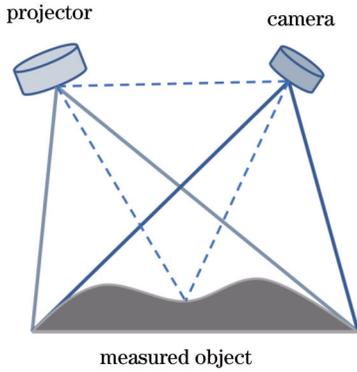


图 1 条纹投影系统示意图
Fig. 1 Schematic of fringe projection system

所提系统采用横向正弦条纹图案,对于投影仪图像的灰度值 $I(x, y)$,其表达式为

$$I(x, y) = \frac{\lambda(x, y) I_{\max}}{2 \left\{ \cos \left[\frac{2\pi N(x - x_{\text{mid}})}{w} + b \right] + 1 \right\}}, \quad (1)$$

式中: $I(x, y)$ 为坐标 (x, y) 处的灰度值; I_{\max} 为条纹图像中的最大灰度值,在单通道 8 位灰度图中通常设置为 255; $\lambda(x, y)$ 为用于像素调制的参数,系统中设置为 1; N 为一幅投影图像上的条纹周期数,根据相移步数,本系统中设置为 8; x_{mid} 为图像中间列所在的索引; w 为图案的宽度; b 为预设的相位偏移量。

对于多步相移方法,相移通常表示为 $\delta_i = 2\pi i/N$, 其中 $i = 1, 2, \dots, N$, 每个像素的相位值可表示为

$$\varphi(u, v) = -\arctan \frac{\sum_{i=1}^N I_i(u, v) \sin(\delta_i)}{\sum_{i=1}^N I_i(u, v) \cos(\delta_i)}, \quad (2)$$

式中: (u, v) 为像素位置点。解相位结果 $\varphi(u, v)$ 的取值范围为 $0 \sim 2\pi$, 表示单周期内的局部相位信息。为实现投影仪-相机像素的精确对应,需要通过相位展开获取全周期下的绝对相位信息,所提系统采用多频外差法^[12]获取绝对相位。在获得绝对相位后,可以根据绝对相位及系统标定的计算结果计算三维点云。

2.2 光度立体视觉测量原理

光度立体视觉方法通过分析不同光照方向下的图像恢复物体法向量,对于非朗伯体表面,任意光照方向下相机成像的灰度值 I 可以表示为

$$I = f(n^T l, n^T v, l^T v), \quad (3)$$

式中: n 为表面法向量; l 为入射光线方向向量; v 为观测方向向量。在光度立体视觉系统中,难点在于未知函数 f 和待测法向量 n 之间存在高度非线性关系,本文采用 Wang 等^[13]提出的基于共位光源的逆向反射模型,表达式为

$$n^T l = g(I, I_c, l^T v), \quad (4)$$

式中: I_c 表示共位光源照射下的光场强度; g 表示逆向反射模型。法向量 n 可以通过已知的光照方向及观测方向的一系列图像解算得到。为方便求解表面法向量,模型可由全连接神经网络构建,并由泛化性训练数据集进行训练^[13]。为适应金属工件表面反射,根据装置中固定光源位置,进一步渲染仅包含金属反射的针对性训练数据集,从而对网络权重进行微调。此外,为了适应近场光度立体视觉的条件,所提网络输入采用 $(I_{i_0}, I_c, I_{i_0}, v)$ 的形式,其中 I_{i_0}, I_c 分别为在光源 i_0 和共位光源 c 照射下图像中 9×9 局部邻域信息。将不同光照方向下的输入数据映射为 512 维向量,经最大池化获取融合特征,从而实现表面法向量的端对端估计。图 2 为近场光度立体视觉(NFPS)系统的网络示意图。

3 多模态数据融合技术

在结构光三维测量过程中,相位解析的精度依赖于光照、成像等条件,所获取的点云虽然具有较高的低频形状精度,但金属表面易出现反光现象,反光区域的投射图像调制度下降,造成点云的测量精度及完整度锐减。所设计的基于点云-法向量多模态协同数据融合的三维重建方案如图 3 所示,其中, $\gamma(x)$ 为隐含变量, x 为图像的像素位置坐标信息。在利用已有光度立体视觉方法获得稠密法向量后,搭建神经网络,通过多层感知机 M_θ 代替传统方法实现法向量积分,设计损失函数,在积分过程中引入法向量-点云的多模态监督信息,其中法向量作为纹理约束,点云作为形状约束,最终实现点云-法向量的多模态信息融合,获取高完整度、高测量精度的点云。

3.1 法向量积分算法

稠密法向量描述物体的纹理信息,该信息可用来估计深度梯度,从而完整还原物体表面深度图。考虑图 4 所示的像素邻域,对于像素位置点 (u, v) ,可以利用 u, v 方向的深度差分来近似该点两个方向的梯度信息 p, q , 又因积分过程中,深度信息 $Z(u, v) > 0$, 那么可以转换为 $\tilde{z} = \ln(Z)$ ^[7], 则对应的 \tilde{p}, \tilde{q} 可表示为

$$\nabla \tilde{z} = (\tilde{p}, \tilde{q})^T. \quad (5)$$

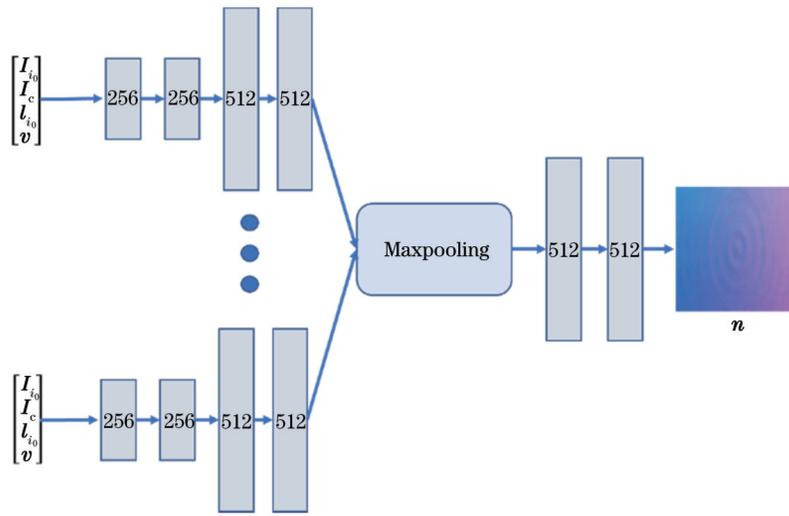


图 2 光度立体视觉网络示意图

Fig. 2 Schematic of photometric stereo vision network

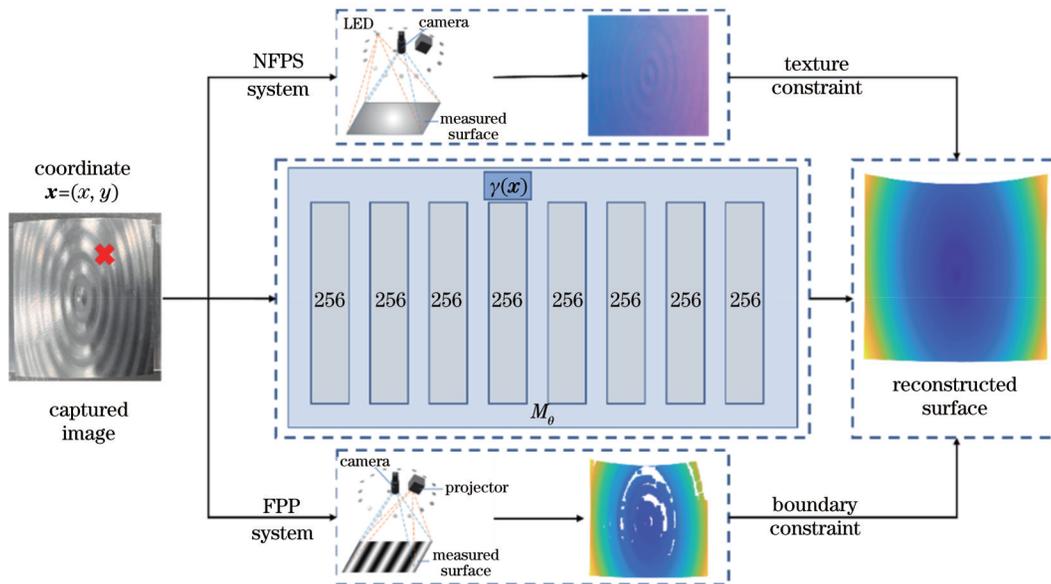


图 3 多模态数据融合流程

Fig. 3 Flowchart of multimodal data fusion

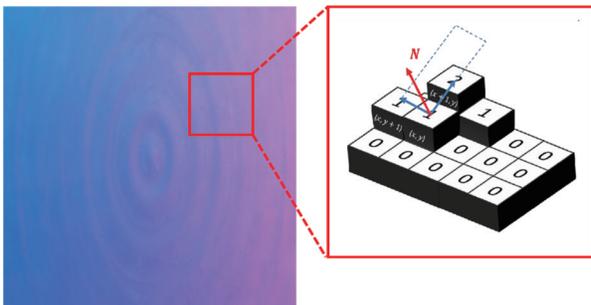


图 4 深度梯度示意图

Fig. 4 Schematic of depth gradient

从式(5)可以看出,法向量积分问题即为求解合适的 \tilde{z} ,使其满足式(5)。由定义可知, u, v 方向的深度梯度变化可以表示为

$$\begin{cases} \tilde{z}(u, v) - \tilde{z}(u - 1, v) = \tilde{p} \\ \tilde{z}(u, v) - \tilde{z}(u, v - 1) = \tilde{q} \\ \tilde{z}(u + 1, v) - \tilde{z}(u, v) = \tilde{p} \\ \tilde{z}(u, v + 1) - \tilde{z}(u, v) = \tilde{q} \end{cases} \quad (6)$$

若对于图像上的待积分区域,各点组合而成的深度向量可以表示为 $\tilde{z} = (\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_m)^T$,其中 m 为像素点的数量。由式(6)可知,若要整体重建物体表面深度,需要求解式(7)。

$$A\tilde{z} = b, \quad (7)$$

式中: A 为大型的稀疏矩阵,其内部系数仅与像素坐标有关; \tilde{z} 为待求解的深度向量; b 为由法向量及相机内参解算而得的梯度向量,可由最小二乘法求解。然而,由于法向量数据本身存在噪声,积分过程中会引入累

计误差,因此需要在法向量积分过程中引入点云形状约束。

3.2 基于多层感知机的点云-法向量融合算法

由 3.1 节的推导可知,法向量图具备一定的物体形貌恢复能力,然而难以保证形状精度。基于点云-法向量的融合算法^[14-15]需要构建复杂的优化方程,该方程对初值敏感且容易陷入局部最优解,因此求解过程较为繁琐。由式(7)可知,法向量积分过程即通过法向量求解像素坐标与深度信息的映射关系,以梯度信息为监督可实现深度求解,所提方法将该过程隐式表达在神经网络的推理过程中,以自监督的形式实现多模态数据融合。

3.2.1 网络结构

所提算法使用的网络结构与神经辐射场^[16]类似,如图 5 所示,网络主体由多层感知机 M_θ 构成,其中 θ 为待优化的权重参数,每层全连接块由线性连接层、BN 层和激活函数 ReLU 构成。所提网络结构和神经辐射场的区别在于,所提网络的输入输出是根据法向量积分的基本原理确定的,以图像的像素位置信息坐标 $x=(x, y)^T \in \mathbb{R}^2$ 为输入,输出为该点所对应的三维坐

标 $X=(X, Y, Z)^T \in \mathbb{R}^3$ 。在无额外约束的情况下,直接由像素坐标映射至三维坐标的难度较大。由 3.1 节可知,法向量重建是一个积分过程,需要在积分前给定法向量的参考深度,因此可以将所提网络表示为残差结构,即

$$X = M_\theta(x) + d_{MD}, \quad (8)$$

式中: d_{MD} 为点云的平均深度,即法向量积分过程中的基准平面高度,在所提算法中由结构光传感器测量得到的点云取平均值获得。引入残差结构后,网络的学习任务从预测重构表面的三维坐标转换为预测平均深度基准下深度的变化情况,即 ΔZ 。简化网络的拟合过程与神经辐射场的训练过程类似,考虑到多层感知机的拟合特性,仅利用像素坐标作为输入会导致输出点云的细节模糊,所提网络难以学习法向量的细节信息。神经辐射场的训练结果表明,在输入信息传递至网络前,使用高频函数将输入信息映射到高维空间,能够更好地拟合高频变化的数据^[16]。所提网络采用位置编码算法,将图像的像素坐标映射到高维空间并作为隐含变量输入网络,即将一维空间映射至频域空间,该映射过程可表示为

$$\gamma(x) = [\sin(2^0 \pi x), \cos(2^0 \pi x), \sin(2^1 \pi x), \cos(2^1 \pi x), \dots, \sin(2^{L-1} \pi x), \cos(2^{L-1} \pi x)], \quad (9)$$

式中: L 为高频空间的维度。为了加强网络对位置编码的感知,在网络输入阶段及第四层全连接部分堆叠输入位置编码向量。

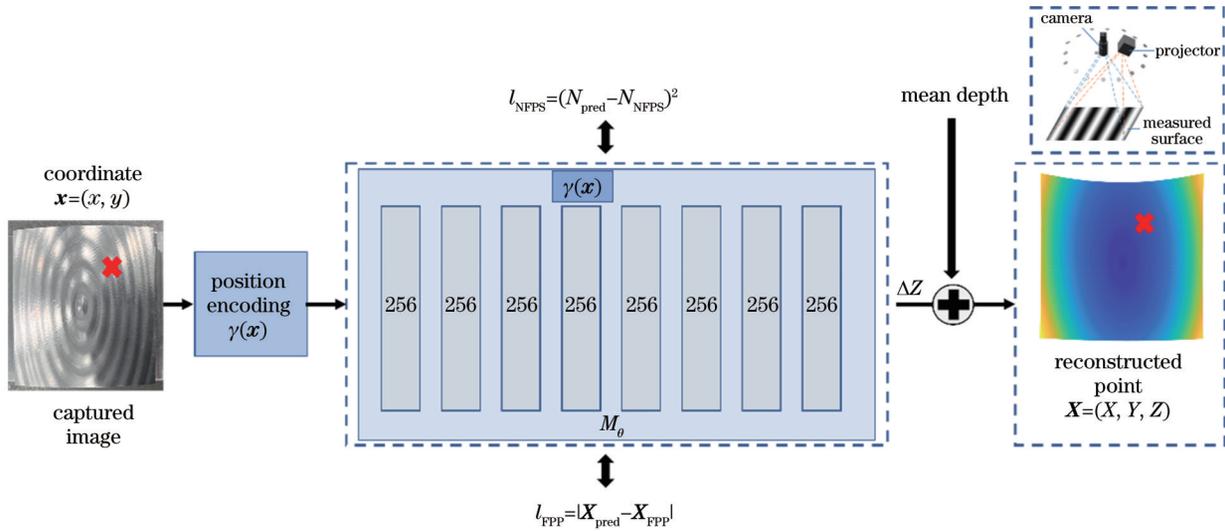


图 5 所提网络结构示意图

Fig. 5 Schematic of proposed network structure

3.2.2 损失函数设计

在网络训练中,主要采取自监督的方式实现法向量积分及积分过程中结构光传感器点云的约束,这意味着网络输出既要保留法向量的细节信息,又要依靠结构光点云约束积分尺度。

所提算法利用多损失函数的策略,将光度立体视觉传感器的法向量与结构光点云作为复合监督源来计算损失函数,即

$$l_{oss} = w_1 l_{NFPS} + w_2 l_{FPP}, \quad (10)$$

式中: l_{NFPS} 为光度立体视觉下的法向量损失,以法向量误差为重建的纹理约束; l_{FPP} 为结构光下的点云损失,以坐标误差为重建的形状约束; w_1 、 w_2 分别为考虑点云和法向量误差尺度时的损失权重系数(实际训练时分别使用 1.0、0.01)。由式(10)可知,光度立体视觉传感器的整体损失由法向量损失和点云损失两部分构成。

法向量及点云损失可表示为

$$\begin{cases} l_{\text{NFPS}} = (N_{\text{pred}} - N_{\text{NFPS}})^2 \\ l_{\text{FPP}} = |X_{\text{pred}} - X_{\text{FPP}}| \end{cases}, \quad (11)$$

式中: X_{pred} 为给定输入像素坐标下网络输出的三维坐标; X_{FPP} 为结构光传感器测量得到的该点三维坐标, 考虑到结构光传感器测得的点云具有较大的噪声, 为避免离群值对损失函数计算的影响; N_{NFPS} 为光度立体视觉传感器测得的法向量值; N_{pred} 为根据网络输出的三维坐标通过局部邻域点云估计得到的法向量值。由于 N_{NFPS} 是相机坐标系下的法向量, 由 3.1 节中法向量积分在图像坐标系下的推导得出, 像素坐标 (u, v) 在相机坐标系下的深度梯度分量 p_1, q_1 可以表示为

$$\begin{cases} p_1 = \frac{Z_c(u+2, v) - Z_c(u-2, v)}{X_c(u+2, v) - X_c(u-2, v)} \\ q_1 = \frac{Z_c(u, v+2) - Z_c(u, v-2)}{Y_c(u, v+2) - Y_c(u, v-2)} \end{cases}, \quad (12)$$

式中: X_c, Y_c, Z_c 为相机坐标系下的三维坐标。若定义法向量为 (n_1, n_2, n_3) , 则相机坐标系下的 p_1, q_1 ^[7] 可以表示为

$$\begin{cases} p_1 = -n_1/n_3 \\ q_1 = -n_2/n_3 \end{cases}, \quad (13)$$

故 N_{pred} 可以表示为

$$N_{\text{pred}} = (p_1, q_1, -1). \quad (14)$$

3.2.3 算法实现

由 3.2.1 节可知, 使用位置编码策略在像素坐标输入到网络前对其进行编码, 网络 M_θ 使用 8 个完全连接的 ReLU 层, 每层均有 256 个通道。由式 (9) 可

知, 所提方法的高频空间维度 $L = 10$, 在第四层重复引入位置编码以加强网络对位置编码的感知, 最终从线性层输出预测深度。所提算法不需要输入额外的训练数据, 仅利用复合传感器下的光度立体视觉及结构光模块即可实现法向量及点云的自监督约束。为提升训练效率, 将像素坐标提前进行随机采样处理, 随机采样策略如图 6 所示。设置采样窗口为 $256 \text{ pixel} \times 256 \text{ pixel}$, 对于每一个 epoch, 先随机生成采样窗口的位置, 获取采样窗口下复合传感器的测量数据 (点云及法向量), 再将像素坐标重新排布成 65536×2 维向量并作为训练数据, 经位置编码后输入网络进行训练。在新一轮 epoch 中刷新采样窗口的位置, 每一轮窗口的位置刷新相互独立。基于深度学习框架 PyTorch 构建网络, 初始学习率为 0.0005, 衰减系数为 0.01, 训练过程中损失函数在 epoch 为 500 时显著降低, 在 2500 时趋于收敛, 因此训练 epoch 设置为 2500, 对应的训练时长为 20 s, 运行计算机的 CPU 为 AMD Ryzen 7 5800H, GPU 为 GeForce RTX 3060。

4 实验结果与分析

为了验证所提出的基于复合传感器的三维重建算法效果, 搭建了基于复合传感器的金属工件测量场景, 如图 7 所示, 并通过仿真和实验来验证所提算法的有效性。复合传感器由一台投影仪、一台工业相机和 29 个 LED 构成。投影仪向物体表面投射正弦条纹实现结构光测量, 29 个 LED 作为点光源为光度立体视觉系统提供多方向的光照, 在这一过程中, 系统始终使用同一台工业相机获取图像, 共采集 24 幅条纹图和 29 幅不同光照方向下的图像, 测量时长为 20 s。

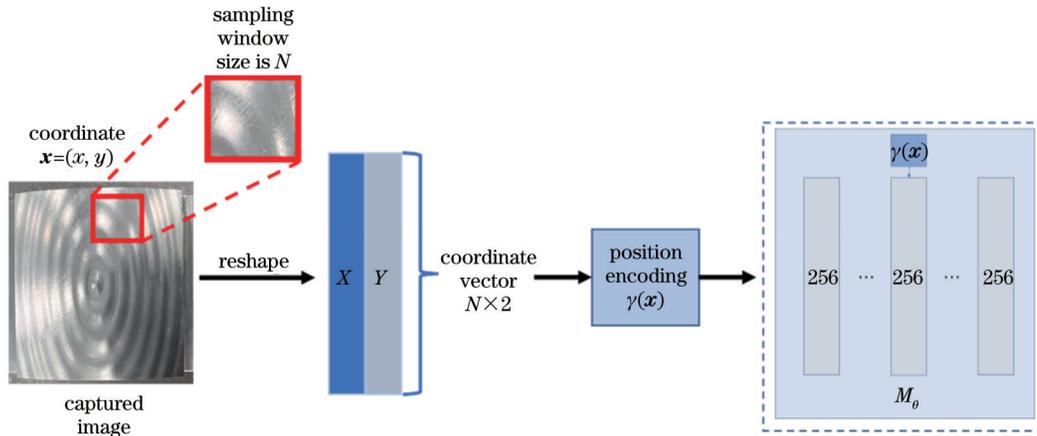


图 6 网络采样策略示意图

Fig. 6 Schematic of network sampling policy

系统硬件包括投影仪 (德州仪器, 型号为 DLP4710, 分辨率为 $1920 \text{ pixel} \times 1080 \text{ pixel}$)、工业相机 (大恒相机, 型号为 MER-503-36U3C, 分辨率为 $2448 \text{ pixel} \times 2048 \text{ pixel}$) 和 29 个 LED。其中, 相机镜头 (Fujinon, 型号为 CF16ZA-1S) 的焦距为 16 mm, 考虑

到待测物体的尺寸, 系统的工作距离为 280~340 mm, 工作范围为 $180 \text{ mm} \times 150 \text{ mm}$ 。

本实验的测量结果与设计的三维模型之间的偏差可以表示为轮廓误差, 由于被测表面存在加工误差, 因此轮廓误差不能直接作为测量误差, 考虑到三坐标测

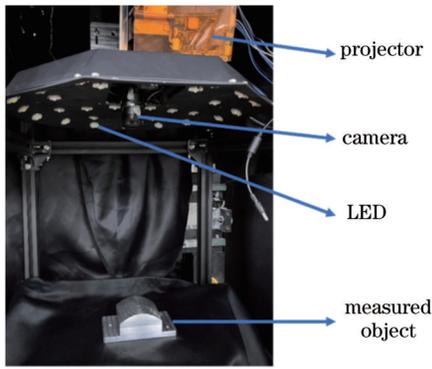


图 7 实验装置整体图

Fig. 7 Overall diagram of experiment equipment

量机的测量精度远超光学设备,本实验将利用三坐标测量机得到的测量值作为被测工件的真实值,测量系统与三坐标测量机得到的测量值之间的相对轮廓误差被认为是系统的测量误差。本实验使用三坐标测量机

(LEITZ PMM XI, Hexagon)测量工件表面,探针的尖端直径为 3 mm,表面测量精度约为 0.6 μm ,采用线扫描的测量模式,利用 PolyWorks 商业软件实现基于迭代最近点的点云与三维模型间的配准,通过计算得到相应的误差图。

4.1 仿真精度分析

设计了具有周期性纹理的平面进行算法精度的仿真实验,综合考虑工件加工过程中可能出现的纹理分布,在理论平面的基础上叠加波纹来模拟加工刀纹所带来的纹理结构,在工件表面采样获得的理论三维坐标可作为仿真点云真值。为获取工件的表面法向量真值,在三维图形图像软件 Blender 内架设相机,参考真实的测量场景,相机的视场角设置为 8° ,渲染图像分辨率设置为 512 pixel \times 512 pixel,根据仿真工件及 Blender 内设场景计算法向量,仿真工件的点云及法向量结果如图 8(a)、(b)所示。

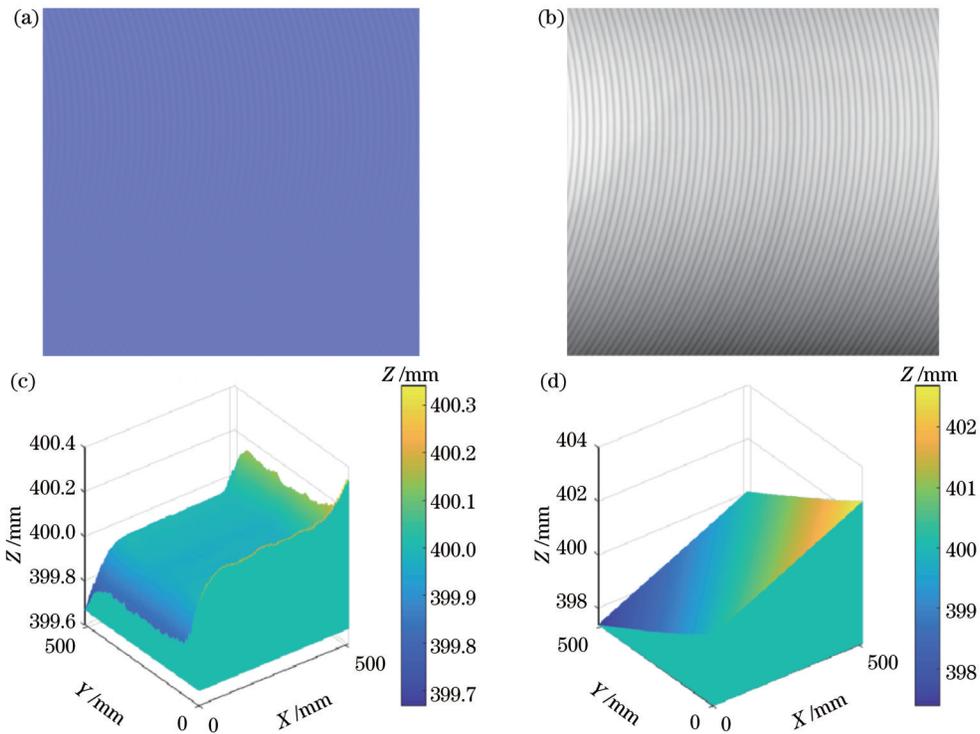


图 8 仿真结果。(a)仿真法向量图;(b)仿真深度图;(c)融合算法结果;(d)法向量积分结果

Fig. 8 Simulation results. (a) Simulation normal image; (b) simulation depth image; (c) result of fusion algorithm; (d) result of normal integrity algorithm

为模拟复合传感器测量过程中的噪声,分别为点云和法向量真值数据添加高斯噪声,生成仿真测量数据,其中参考所搭建的真实结构光传感器的测量精度约为 0.04 mm,向点云数据加入标准差 σ 为 0.04 的高斯噪声,以模拟结构光测量过程中的高频点云噪点。为模拟所提算法在不同法向量测量精度时的融合性能波动变化,分别向法向量数据添加均值为 0.01、0.03、0.05 和 0.10,标准差 σ 均为 0.025 的高斯噪声,用以模拟光度立体视觉测量中的系统误差,其中,噪声水平 μ

与法向量角度误差之间的关系如表 1 所示。

为验证所提点云-法向量融合算法的优越性,将其与传统基于法向量积分的重建算法进行比较。以法向量噪声水平 $\mu = 0.05$ 为例,所提融合算法及法向量积

表 1 μ 与法向量角度误差之间的关系

Table 1 Relationship between μ and measurement error of normal vector angle

μ	0.01	0.03	0.05	0.10
Normal noise /($^\circ$)	2	3	4	7.5

分算法的重建结果如图 8(c)、(d)所示。可以看到,由仿真平面生成的噪声法向量积分所得到的结果存在累计误差,在 Z 方向上出现明显的翘起现象,而所提点云-法向量融合算法由于在法向量重建过程中加入点云约束,整体依然维持平面形状。

为定量比较不同算法的性能,以仿真点云的 Z 轴深度为真值,比较融合算法与法向量积分算法在 Z 轴方向上的偏差,以法向量噪声水平 $\mu = 0.05$ 为例,误

差结果如图 9 所示。由误差分布结果可知:由于传统法向量积分算法在积分过程中会存在累计误差,当在法向量中加入均值为 0.05 的噪声来模拟系统误差时,重建过程中 Z 轴方向的精度从中心至对角线迅速劣化,导致重建结果出现极大的偏差;所提算法相比于传统法向量积分算法存在位置约束,因此重建平面 Z 轴方向的测量精度有极大提升,Z 轴的极值误差从 2.62 mm 优化至 0.34 mm,提升了 87%。

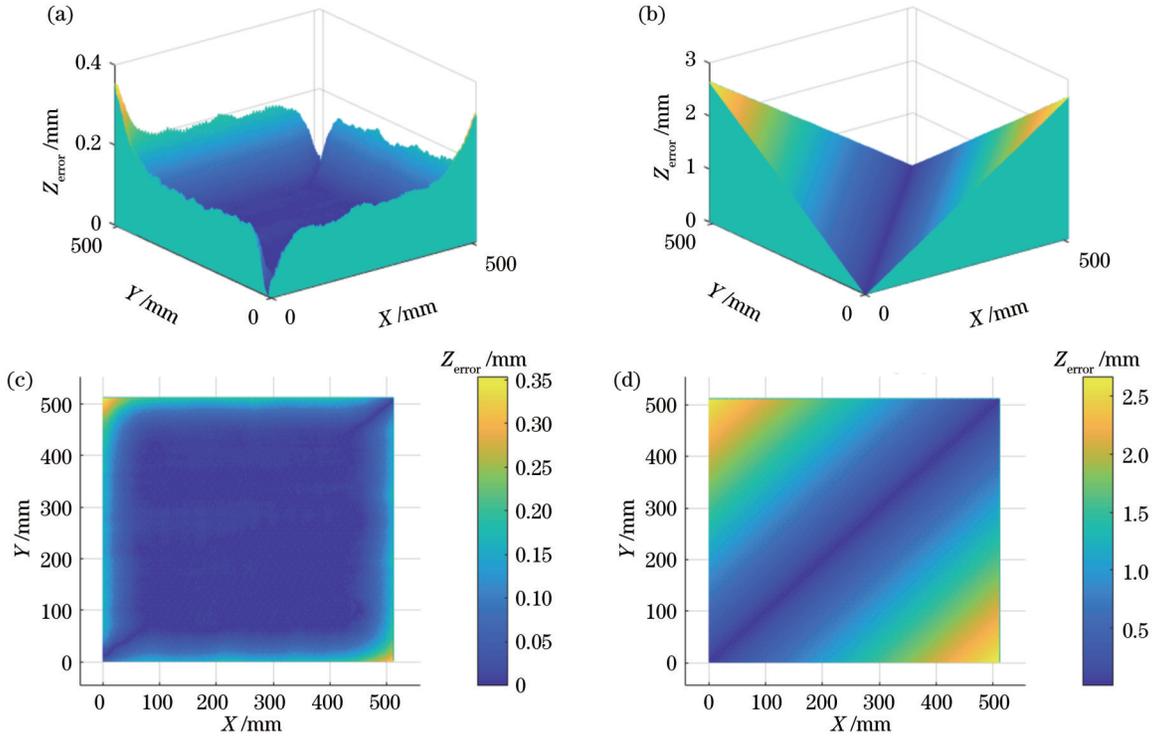


图 9 不同算法沿 Z 方向的平均误差结果。(a)(b)三维分布;(c)(d)二维分布

Fig. 9 Mean error results along Z-axis of different algorithms. (a) (b) 3D distribution; (c) (d) 2D distribution

图 10 给出了所有噪声水平下两种算法重建点云的 Z 轴误差三维分布图。由图 10 可知:当法向量仅存在均值为 0 的高斯噪声时,由于法向量精度接近真值,两种算法都能获得良好的重建效果;随着噪声水平提高,传统法向量积分算法的重建结果开始出现明显的深度偏差,这是因为其网络采用滑窗采样的方式获取数据,相比于中心像素区域,像素边缘区域受窗体大小的限制,采样次数偏少,因此整体误差分布呈现四周大于中心的情况。计算 Z 轴平均误差,结果如图 11 及表 2 所示。由实验结果可知,随着法向量噪声增大,所提算法依然能够维持良好的重建精度,而法向量积分算法存在明显的劣化情况。比较平均深度误差发现,4 种噪声水平下所提算法的精度相比于传统积分算法分

别提升了 96.6%、96.4%、96.2% 和 92.4%。

4.2 实验精度分析

本实验选用图 12 所示的工件进行实验精度分析,对系统的结构光模块及光度立体视觉模块进行标定,包括相机标定^[17]、投影仪-相机系统标定^[18]和 LED-相机标定^[19],在测量过程中,结构光系统投射条纹图案至待测物体表面,解算相位获取表面点云;光度立体视觉系统采集 29 幅不同光照方向下的图像,分别利用所提方法及基于最小二乘的传统光度立体法^[17]获取表面法向量,法向量恢复结果如图 13 所示;将所得法向量分别进行积分重建及融合重建,实验结果如图 14 所示。

由图 14(a)、(b)可知,由于金属表面具有较为复

表 2 不同噪声水平 Z 轴平均误差对比

Table 2 Comparison of mean error along Z-axis under different noise levels

unit: mm

μ	0	0.01	0.03	0.05	0.10
Fusion algorithm	0.0025	0.0062	0.0193	0.0336	0.1288
Integrity algorithm	0.0023	0.1851	0.5445	0.8884	1.6980

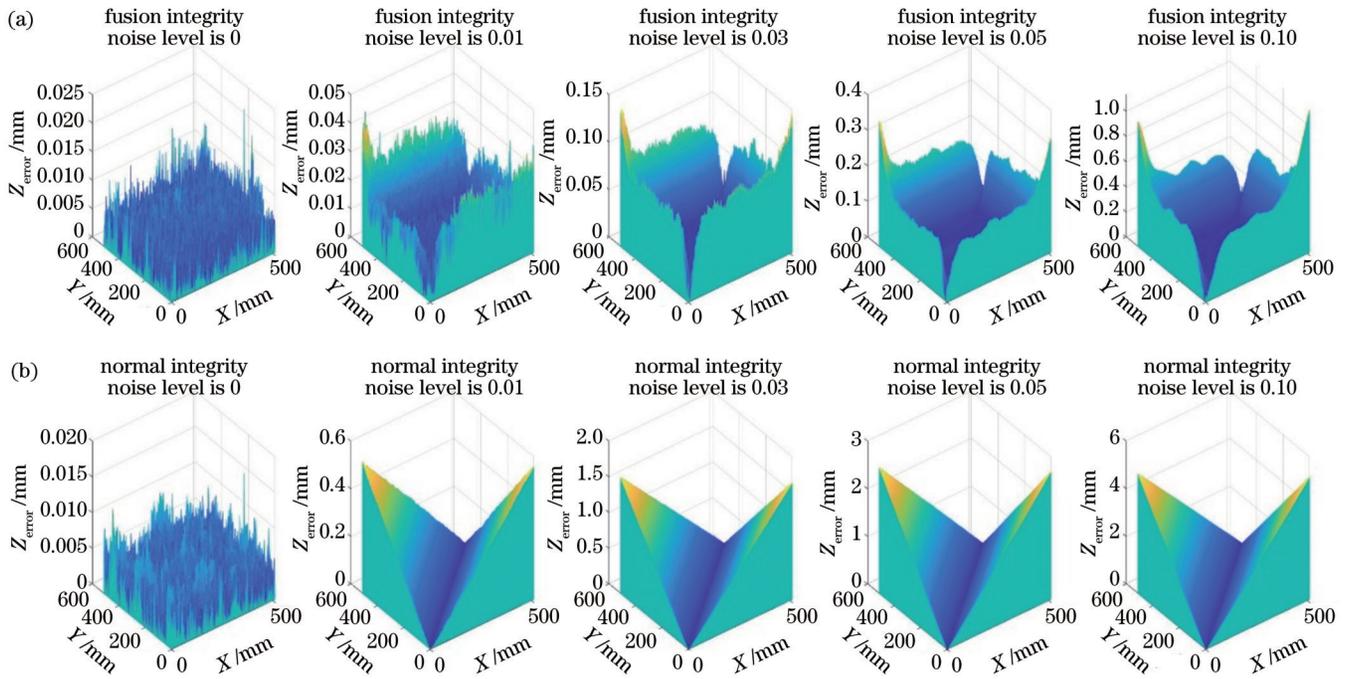


图 10 不同噪声水平下不同算法沿 Z 方向的平均误差结果。(a)融合算法;(b)法向量积分算法

Fig. 10 Mean error results along Z-axis of different algorithms under different noise levels. (a) Fusion algorithm; (b) normal integrity algorithm

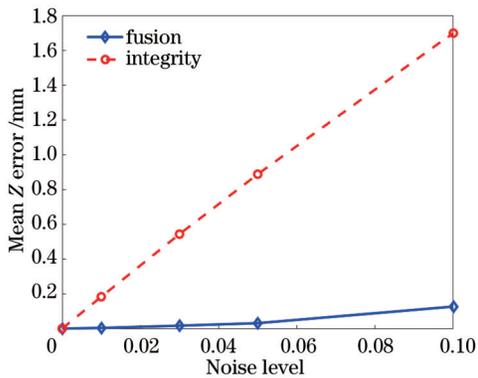


图 11 不同噪声水平 Z 轴平均误差对比

Fig. 11 Comparison of mean error along Z-axis by different noise level

杂的光反射特性, 结构光测量方法所得结果的部分区域的调制度较低, 表面点云不仅会存在高频噪声, 甚

至会出现较大范围的点云缺失; 由图 14(c)、(d)可知, 利用最小二乘法求解得到的法向量本身存在很大的系统误差, 无法为点云提供高精度约束; 由图 14(e)、(f)可知, 所提方法获取的法向量在积分过程中会逐渐累积误差, 在边缘处存在明显翘曲; 由图 14(g)、(h)可知, 所提融合算法由于在重建过程中引入点云约束, 其点云的整体形状相比于积分算法的点云没有出现明显的累积误差, 即使是在出现高频噪声及点云缺失的区域, 而引入光度立体视觉的高精度法向量约束后, 融合点云也没有出现表面缺失, 并且高频噪声大幅减少。

为验证所提的点云-法向量融合算法的测量重复性, 对两种工件进行 10 次重复实验, 表 3 所示为融合点云方法的重复测量精度, 该精度由均方根(RMS)误差表示, 是重复测量点云与第一组测量点云的匹配差异。

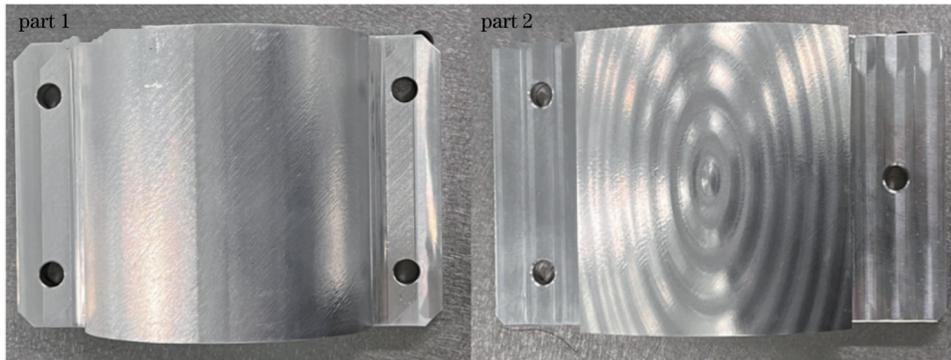


图 12 实验工件

Fig. 12 Experiment parts

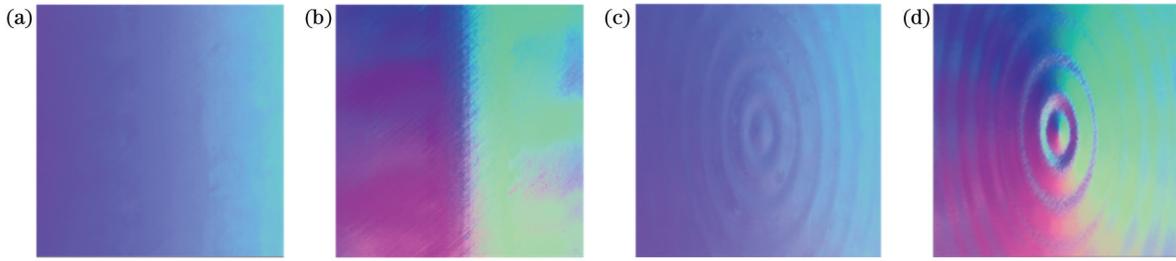


图 13 法向量结果。(a)工件 1 的深度学习结果；(b)工件 1 的最小二乘法结果；(c)工件 2 的深度学习结果；(d)工件 2 的最小二乘法结果

Fig. 13 Normal vector results. (a) Result of part 1 by deep learning method; (b) result of part 1 by least square method; (c) result of part 2 by deep learning method; (d) result of part 2 by least square method

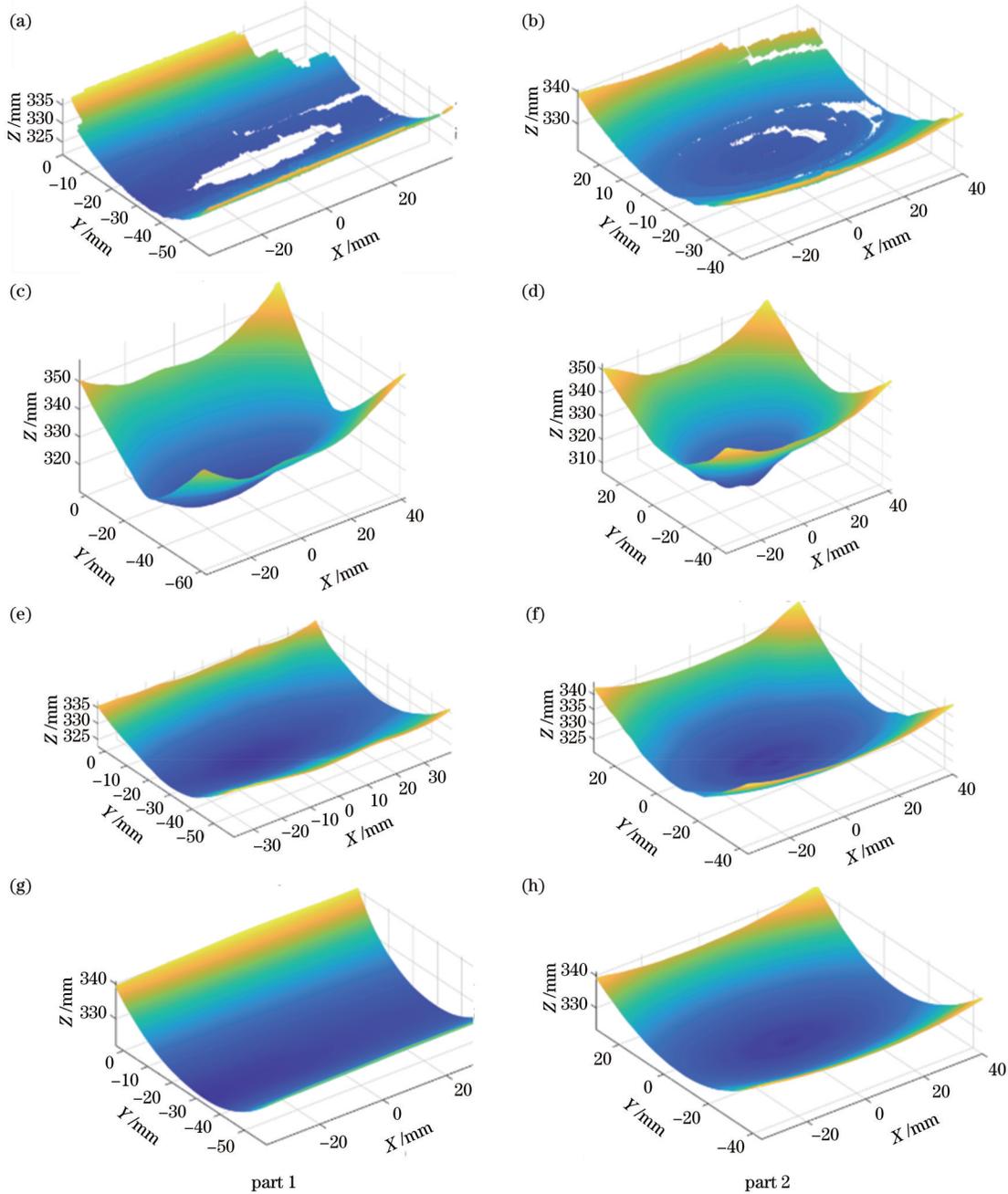


图 14 工件 1 和工件 2 的实验结果。(a)(b)结构光重建结果；(c)(d)最小二乘法向量融合重建结果；(e)(f)积分算法重建结果；(g)(h)所提方法重建结果

Fig. 14 Experimental results of part 1 and part 2. (a)(b) Structured light reconstruction results; (c)(d) fusion reconstruction results by least square normal vector; (e)(f) integration algorithm reconstruction results; (g)(h) reconstruction results by proposed method

由表 3 可知,两个工件的重复测量精度平均值分别为 2.5、1.0 μm ,表明系统的重复测量精度较好。

表 3 融合点云重复测量精度(RMS)

No.	1	2	3	4	5	6	7	8	9
Cylinder	2.0	2.3	1.2	2.1	2.4	3.7	2.9	2.9	3.4
Quadric	1.4	1.4	0.6	1.0	0.7	0.7	1.4	1.0	0.9

为验证所提出的点云-法向量融合算法的精度,利用 PolyWorks 商业软件计算测量点云与真实工件的误差,利用三坐标测量机对工件表面进行测量,所得结果作为工件真值。两个实验工件的误差分布如图 15 所示,表 4 所示为两种测量方法的 RMS 误差。相较于结构光测量方法,所提算法在保证点云完整度的基础上,对于圆柱面工件的 RMS 误差由 156 μm 减少至

113 μm ,而对于二次曲面工件的 RMS 误差由 188 μm 减少至 131 μm 。以三坐标测量机测量的加工工件真值为基准,所提算法对于圆柱面工件的 RMS 误差从 121 μm 减少至 78 μm ,精度提升了 35.5%;对于二次曲面工件的 RMS 误差从 113 μm 减少至 56 μm ,精度提升了 50.4%。可见,所提点云-法向量融合算法显著提升了工件测量的完整性和精度。

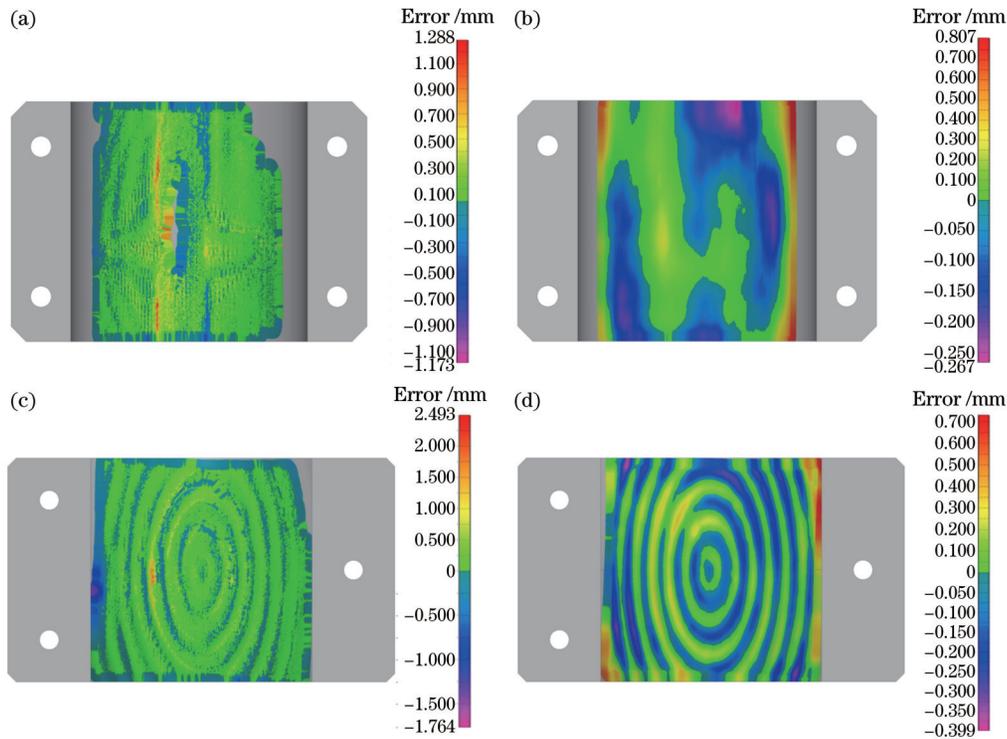


图 15 误差分布结果。(a)工件 1 的结构光方法结果;(b)工件 1 的融合算法结果;(c)工件 2 的结构光方法结果;(d)工件 2 的融合算法结果

Fig. 15 Error distribution results. (a) Structured light method result of part 1; (b) fusion algorithm result of part 1; (c) structured light method result of part 2; (d) fusion algorithm result of part 2

表 4 不同方法测量 RMS 误差

Surface	CMM	Structured light method	Fusion algorithm
Cylinder	35	156	113
Quadric	75	188	131

5 结 论

提出一种基于多层感知机的金属表面光度-结构光测量数据融合方法。在金属构件测量过程中,通过复合传感器获取点云-法向量数据,搭建基于位置编码

的多层感知神经网络。为实现点云与法向量的有效融合,在网络自监督训练过程中,以点云为形状约束,以表面法向量为纹理约束。实验结果表明:相比于传统结构光重建方法,所提算法能够在保证点云完整度的同时滤除高频噪声;以三坐标测量机的测量结果为基准,所提算法的测量精度较结构光测量方法提升了 50.4%。未来的研究工作将进一步分析复合传感器系统的不确定度。

参 考 文 献

[1] 熊义可. 基于自适应结构光投影的三维测量系统研究[D]. 杭

- 州: 浙江大学, 2016: 1-7.
- Xiong Y K. Research on 3D measurement system based on adaptive structured light projection[D]. Hangzhou: Zhejiang University, 216: 1-7.
- [2] 林辉. 高动态范围光亮表面的结构光三维形貌测量方法研究与实现[D]. 广州: 广东工业大学, 2017: 4-15.
- Lin H. Research and implementation of structured light three-dimensional topography measurement method for bright surface with high dynamic range[D]. Guangzhou: Guangdong University of Technology, 2017: 4-15.
- [3] 张宗华, 于瑾, 高楠, 等. 高反光表面三维形貌测量技术[J]. 红外与激光工程, 2020, 49(3): 0303006.
- Zhang Z H, Yu J, Gao N, et al. Three-dimensional shape measurement techniques of shiny surfaces[J]. Infrared and Laser Engineering, 2020, 49(3): 0303006.
- [4] Zhang S, Yau S T. High dynamic range scanning technique[J]. Optical Engineering, 2009, 48(3): 033604.
- [5] 詹瑜, 胡丹, 刘凯. 面向高反光区域的自适应结构光[J]. 光学学报, 2022, 42(16): 1612001.
- Zhan Y, Hu D, Liu K. Adaptive structured light for highly reflective areas[J]. Acta Optica Sinica, 2022, 42(16): 1612001.
- [6] Woodham R J. Photometric method for determining surface orientation from multiple images[J]. Optical Engineering, 1980, 19(1): 139-144.
- [7] Quéau Y, Durou J D, Aujol J F. Normal integration: a survey [J]. Journal of Mathematical Imaging and Vision, 2018, 60(4): 576-593.
- [8] 简振雄, 王晰, 任杰骥, 等. 基于近场光度立体视觉的金属表面纹理重构[J]. 光学学报, 2021, 41(11): 1112002.
- Jian Z X, Wang X, Ren J J, et al. Metal surface texture reconstruction based on near-field photometric stereo[J]. Acta Optica Sinica, 2021, 41(11): 1112002.
- [9] 徐科, 周鹏, 杨朝霖. 基于光度立体学的金属板带表面微小缺陷在线检测方法[J]. 机械工程学报, 2013, 49(4): 25-29.
- Xu K, Zhou P, Yang C L. On-line detection technique of tiny surface defects for metal plates and strips based on photometric stereo[J]. Journal of Mechanical Engineering, 2013, 49(4): 25-29.
- [10] 刘瑞鑫, 李立轻, 汪军, 等. 基于光度立体视觉的织物褶裥等级评定[J]. 东华大学学报(自然科学版), 2013, 39(1): 48-52, 59.
- Liu R X, Li L Q, Wang J, et al. The evaluation of fabric pleat grade based on photometric stereo[J]. Journal of Donghua University (Natural Science), 2013, 39(1): 48-52, 59.
- [11] 付琳, 洪海波, 王晰, 等. 基于逆向反射模型的非朗伯光度立体视觉[J]. 光学学报, 2020, 40(5): 0520001.
- Fu L, Hong H B, Wang X, et al. Non-lambertian photometric stereo vision based on inverse reflectance model[J]. Acta Optica Sinica, 2020, 40(5): 0520001.
- [12] Towers C E, Towers D P, Jones J D C. Absolute fringe order calculation using optimised multi-frequency selection in full-field profilometry[J]. Optics and Lasers in Engineering, 2005, 43(7): 788-800.
- [13] Wang X, Jian Z X, Ren M J. Non-Lambertian photometric stereo network based on inverse reflectance model with collocated light[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2020, 29: 6032-6042.
- [14] Pei X H, Ren M J, Wang X, et al. Profile measurement of non-Lambertian surfaces by integrating fringe projection profilometry with near-field photometric stereo[J]. Measurement, 2022, 187: 110277.
- [15] Zhang Q, Ye M, Yang R G, et al. Edge-preserving photometric stereo via depth fusion[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 2472-2479.
- [16] Mildenhall B, Srinivasan P P, Tancik M, et al. NeRF: representing scenes as neural radiance fields for view synthesis [J]. Communications of the ACM, 2022, 65(1): 99-106.
- [17] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [18] Zhang Z H, Huang S J, Meng S S, et al. A simple, flexible and automatic 3D calibration method for a phase calculation-based fringe projection imaging system[J]. Optics Express, 2013, 21(10): 12218-12227.
- [19] Nie Y, Song Z, Ji M, et al. A novel calibration method for the photometric stereo system with non-isotropic LED lamps[C]//2016 IEEE International Conference on Real-time Computing and Robotics (RCAR), June 6-10, 2016, Angkor Wat, Cambodia. New York: IEEE Press, 2016: 289-294.

Multilayer Perceptron-Based Fusion Method for Metal Surface Measurement Data by Multi-Sensors Incorporating Photometric Stereo and Structured Light

Yang Yuansong, Wang Xi, Ren Mingjun*

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

Objective Accurate reconstruction of the workpiece surface is crucial to evaluate the machining process and control product quality. Non-contact three-dimensional (3D) reconstruction methods are widely employed due to their low cost, high measurement speed, and high measurement accuracy. Common non-contact reconstruction methods, such as structured light technology, photogrammetry, and laser scanning, generally assume a diffuse surface. However, the machined metal surfaces obviously deviate from this assumption and exhibit complicated high light reflectance, which incurs heavy high-frequency noises and numerous invalid measurement data. Although high dynamic structured light research works focus on this challenge, all these methods fail to break the assumption of diffuse surface, and thus the

measurement efficiency is low, and the recovery of missing measurement data is limited. Compared with these methods, photometric stereo can estimate the complete surface normal of the metal surface and overcome the influence of non-diffuse reflectance by inversely modeling the reflectance. However, the error accumulation happens during the normal integration, resulting in distorted shapes. In order to solve these problems, a multilayer perceptron-based fusion method is proposed for metal surface measurement by multi-sensors incorporating photometric stereo and structured light. The point cloud by structured light offers the geometry constraint, and the surface normal by photometric stereo provides the texture constraint. As a result, the accurate 3D reconstruction of the metal surface is achieved by fusing the point cloud and the surface normal.

Methods In this paper, one multi-sensor system including several light-emitting diodes (LEDs), one projector, and one camera is designed, and the structured light measurement is achieved by the projector and the camera. The photometric stereo system consists of the camera and the LEDs. These two measurement sensors share the same camera, and thus the error of coordinate system matching is avoided. The structured light gives the imperfect noisy point cloud, and the complete surface normal with small high-frequency noises is estimated by the photometric stereo. In order to fuse these two different measurement data, a self-supervised multilayer perceptron network based on position encoding is designed according to the principle of normal integration, which achieves the mapping from the pixel coordinate to the depth under the camera coordinate. In the training process, the point cloud provides shape supervision, and the accurate surface normal gives the texture supervision to complete the point cloud. Thus, the complete and highly accurate 3D reconstructions of metal surfaces are output. Both synthetic experiments and real experiments verify the effectiveness of the proposed fusion method.

Results and Discussions In this paper, synthetic experiments are designed to test the influence of different surface normal estimation errors on the proposed fusion method and verify the superiority of the proposed method to normal integration, especially under the condition of the noisy surface normal. The estimation error of structured light is synthesized by adding the Gaussian noise with a mean value of 0 and a standard deviation of 0.040 mm. In order to simulate different levels of surface normal estimation errors, the Gaussian noise with mean values of 0.01, 0.03, 0.05, and 0.10 and standard deviation of 0.025 is imposed respectively. With the increase in the surface normal estimation error, the proposed method can maintain excellent reconstruction accuracy, but the conventional normal integration algorithm exhibits dramatic degradation (Fig. 11). For different noise levels, the accuracy of the proposed fusion method enhances by 96.6%, 96.4%, 96.2%, and 92.4% compared with that of normal integration. In the real experiments, the point cloud and normal vector data are measured by the multi-sensors system incorporating the photometric stereo and structured light. As exhibited in Fig. 14, the reconstruction results of normal integration obviously distort due to the cumulative error, the point cloud by the structured light has missing data and high-frequency noises. The proposed fusion method effectively avoids the cumulative error of normal integration, completes the invalid data, and reduces the measurement error. According to the measurement results of the coordinate measurement machine (CMM), the accuracy of the proposed method enhances by about 50.4% compared with that of the structured light measurement (Fig. 15).

Conclusions In this paper, a multilayer perceptron-based fusion method is proposed for the measurement data of the multi-sensor system incorporating photometric stereo and structured light. In the metal surface measurement, the point cloud and surface normal are obtained by the multi-sensor, and a multilayer perceptron network based on position encoding is designed to achieve the final measurement. In order to effectively fuse the point cloud and the surface normal, the point cloud is employed as the shape constraint, and the surface normal vector is adopted as the texture constraint to supervise the proposed network. The synthetic experiments prove that the accuracy of the proposed fusion method improves by over 90% than that of the normal integration under the condition of noisy normal. The real experiments indicate that the proposed method can simultaneously filter out the high-frequency noise and complete the invalid measurement data. Besides, compared with the structured light reconstruction results, the accuracy of the proposed methods improves by about 50.4% based on the measurement results of the CMM. Future research work can further analyze the uncertainty of the proposed multi-sensor systems.

Key words measurement; structured light; photometric stereo; vision measurement; multi-sensor; multilayer perceptron