

基于深度相机的物体姿态感知方法

董理, 胡茂海*, 杨志荣

南京理工大学电子工程与光电技术学院, 江苏 南京 210094

摘要 针对机器视觉中目标物体相对姿态的准确测量问题, 提出了一种应用深度相机获取目标物体三维点云数据从而实现其相对姿态感知的方法。该方法根据目标物体空间结构先验知识选择同名点, 应用单个的双目视觉测量系统获取目标物体三维点云数据, 分析和处理该数据实现同名点识别, 进而计算目标物体的三个同名点分别在物体坐标系与相机坐标系上的坐标值, 利用奇异值分解(SVD)解相对姿态。该方法能够有效地解决三维物体姿态感知问题, 提高了姿态感知的效率。实验数据表明, 所提方法测量出的六自由度相对误差在 $\pm 0.5^\circ$ 和 ± 1 mm 以内, 满足了机器视觉姿态感知在工业生产中的应用要求。

关键词 机器视觉; 相对姿态; 深度相机; 同名点; 双目视觉; 六自由度

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP202259.1415028

Object Attitude Perception Method Based on Depth Camera

Dong Li, Hu Maohai*, Yang Zhirong

School of Electronic and Optical Engineering, Nanjing University of Science & Technology,
Nanjing 210094, Jiangsu, China

Abstract To address the problem of accurate measurement of relative pose of the target object in machine vision, this paper proposed a method, which applies a depth camera to acquire 3D point cloud data of target object, to achieve its relative pose perception. This method selected homonymous points based on a priori knowledge of spatial structure of the target object and applied a single binocular vision measurement system to obtain the 3D point cloud data of the target object while analyzing and processing the data to achieve homonymous point recognition. Furthermore, this method calculated the coordinate values of three homonymous points of the target object on the object coordinate system and the camera coordinate system, respectively, and solved the relative pose by using singular value decomposition (SVD). This method effectively solved the problem of 3D object pose perception and improved the efficiency of pose perception. The experimental data demonstrate that the relative error of six degrees of freedom measured by the proposed method is within $\pm 0.5^\circ$ and ± 1 mm, which satisfies the application of machine vision pose perception in industrial production.

Key words machine vision; relative attitude; depth camera; homonymous point; binocular vision; six degrees of freedom

1 引言

三维目标姿态感知是机器视觉^[1-4]和未来智能科技的核心内容, 已应用于智能机器人运作^[5-6]、视频监控、医疗教育、体育竞技、国防安全等领域。传统方法估测三维姿态时存在许多弊端: GPS 三维姿态测量^[7]虽成本低, 测量精度高, 受气候影响很小, 但很难应用于室内测量; 基于惯性导航^[8-9]单元的三维姿态测量精度随价格波动变化很大。近年来新型的三维姿态测量方法逐渐被提出, 但精准、简洁且适合各类场景的三维

姿态感知方法还有很大研究价值。

孙国燕等^[10]利用两个相同的二维准直仪和一个反射部件构建的系统测量物体三维姿态角, 该类人工搭建光学系统测量姿态的方法的测角精度易约束视场范围。如今双目视觉技术^[11-13]已成为机器视觉的核心, 通过左右镜头掌握物体立体信息, 模拟人眼, 快速、灵活、准确地感知三维姿态。有许多研究者已将双目视觉融入位姿测量中, 例如魏弦等^[14]提出一种基于双目视觉识别确定工件姿态的算法, 通过分析双摄像机成像平面上的投影得出工件特征轴线, 求解相机光心和

收稿日期: 2021-09-23; 修回日期: 2021-10-08; 录用日期: 2021-10-19

通信作者: hmh@njust.edu.cn

像面轴线所组成的两平面,并利用空间交汇原理计算工件俯仰角和偏航角,若提取到的工件特征与目标工件特征相符,则实现工件的识别,这种方法难以分辨角度特征相近的工件。束安等^[15]以两台可见光相机构建双目视觉,选用双边滤波法进行图像预处理,目的是为了检测对接环椭圆与目标边框直线,以解算目标坐标系与世界坐标系的相对位姿。这种方法中图像预处理的精度较依赖于拍摄图像的质量,这种方法主要对形态简单而规整的目标有效,对于特征错综复杂的部件,拟合椭圆与直线时会引入很多误差。阴雷等^[16]搭建双目视觉系统并利用双目正交迭代算法获取系统运动中的位姿信息,该方法需事先精准标定相机外参,且需全局优化系统的累积误差,缺乏灵活性。综合比较各类方法可知,高效性、普适性、可行性是视觉计算物体位姿面临的问题,而快速准确且适用于多种目标的图像处理方法和姿态计算方法是解决问题的关键。

本文分析深度相机中以双目视觉技术重建的深度图,提出了一种三维图像处理方法以获取同名点坐标值,通过计算同名点在像素、图像、相机与物体坐标系的关系计算目标物体的姿态。选用一款适合工作环境的深度相机,采集深度图和 RGB 图,处理三维点云数据并分析点与点间的深度关系,准确、高效地获取具有特征的同名点。通过逐一搜索点、确定阈值、最小二乘法、简单平均法以及已知的实际工件尺寸等抑制各种图像噪声,与这些方法相比,所提方法具有更强的灵活性、稳定性与高效性。一系列的实验结果表明,整个系统能在允许误差范围内实现对目标物体姿态的感知。

2 图像处理

三维物体上各个点在像素、图像、相机和物体坐标系上有不同的坐标值,物体上固定点在各个坐标系下统称为同名点。三维物体有独特的形态特征,本文图像处理的目的是准确获取物体上最具特征的三个同名点。而深度图反映图像上每个像素点与相机间的距离,以此距离(深度)作为物体特征的评判标准。

深度相机根据左右两个相机的图像构建数学模型,从而重建一幅具有三维点云数据的深度图像。对于一个内参已标定的深度相机,像素坐标系和图像坐标系的关系为

$$\begin{cases} u = \frac{x}{dx} + u_0 \\ v = \frac{y}{dy} + v_0 \end{cases}, \quad (1)$$

式中: (u, v) 和 (x, y) 分别是像素坐标和图像坐标; (u_0, v_0) 是图像坐标系原点在像素坐标系中的坐标值; dx 和 dy 表示每一行和每一列一个像素分别代表多少毫米。图像坐标系转换到相机坐标系后,得到

$$\begin{cases} X_c = \frac{x \times Z_c}{f} \\ Y_c = \frac{y \times Z_c}{f} \end{cases}, \quad (2)$$

式中: (X_c, Y_c) 是相机坐标值; Z_c 是深度值; f 是相机焦距。

2.1 图像处理过程

在三维物体表面选取具有明显深度特征的三个同名点,例如点在轮廓边缘、最近端面或是物体上有规律深度变化的部件等。所提方法在图像上以直线搜索每个像素点对应的深度值,对深度均匀变化或是突然变化的各个点进行进一步判断,去除噪声后获取具有特征的同名点。以深度相机感知待切割钢筋的目标工件为例,对于此工件,钢筋与端面之间有深度变化,因此选取同名点在端面边界线上且与钢筋端点处于同一平面的位置。相机拍摄工件 RGB 图和欲选择的特征同名点如图 1(a)所示,对三维数值场中深度值不为 0 的点进行着色,绘制的深度图像如图 1(b)所示。根据深度图像,易观察到工件距离相机最前端面含有更多深度值大于 0 的点,基于此观察对该 640×400 深度图像上每个点取深度整数值,以方便在程序上统计深度直方图。直方图显示不同深度整数值所包含的像素点数。根据直方图提取出现次数最多的深度整数值作为工件离相机最近端面深度估计值,对此最近端面进行二值化处理,结果如图 1(c)所示。

欲准确获取图 1(a)四个特征点作为待选同名点,需拟合出图 1(b)中离相机最近的四根钢筋所在直线。首先将最近端面深度估计值与实际测量尺测得工件端面相对于钢筋的垂直距离取和,得到钢筋在三维数值场中的深度特征数,设定 10 mm 为阈值,有 $N_{\text{feature}} - 10 < D < N_{\text{feature}} + 10$,其中 N_{feature} 为钢筋深度特征数, D 为钢筋实际深度。利用上述约束,在图像中心区域获取每根钢筋上第一个边界点。根据迭代思想在首个边界点附近找寻第二个边界点,第二个边界点附近找寻第三个边界点,依此类推得到足以代表四根钢筋的许多点,再用最小二乘法拟合出钢筋所在直线,并得到斜率。

三维数值场中获取第一根和第四根钢筋(从左到右)所在直线上每个像素点的深度值,两条直线上深度值突变的四个像素点即是图 1(a)绘出点,根据式(2)对它们进行计算,所得四个相机坐标值是图像处理结果。图 1(d)展现最近钢筋与端面二值化图像,包括图像处理过程中四根钢筋左边界特征点、所拟合的直线及最终获得的四个特征点。从四个特征点中选取三个,共有四种组合,但注意选取特征点时需与物体上所选项一一对应,从而构成三个同名点。

所提方法在三维数值场中逐一搜索点来确保深度图中点与实际物体上点一一对应,并拟合特征部位所在直线,以大幅度缩小搜索范围。归纳的图像处理核心步骤:在目标深度图中找到一个特征部位上(如本实

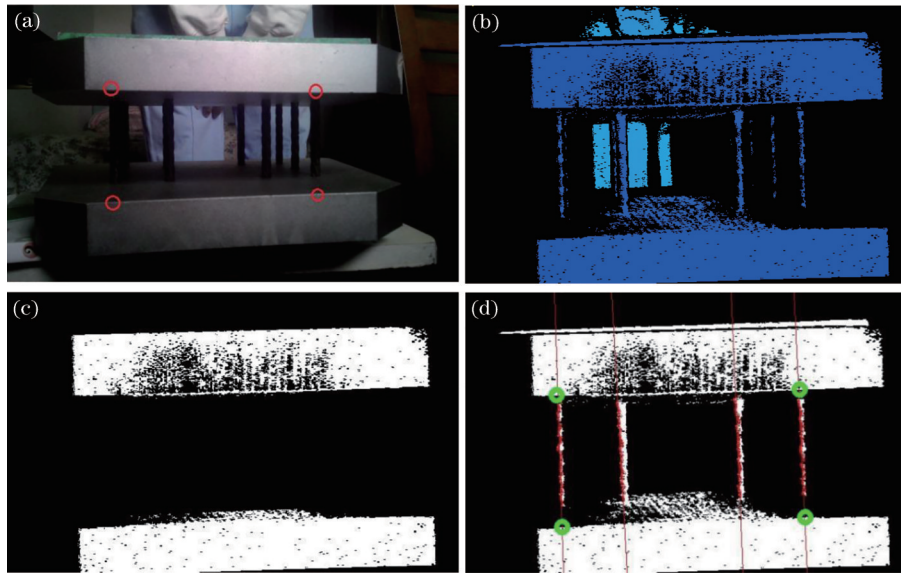


图1 轨枕切割图像处理示例。(a) RGB图;(b)深度图;(c)端面二值化图;(d)图像处理结果图

Fig. 1 Example diagram of sleeper cutting image processing. (a) RGB diagram; (b) depth map; (c) end face binarization diagram; (d) diagram of image processing result

验钢筋)的点作为种子点(搜索起点);在第一个种子点周围找寻与该点具有相同特征的点,依此迭代产生许多点,用这些点拟合出轮廓、边缘或特征部位;在特征线上设定阈值,找到深度突变点。

2.2 误差分析

误差分析一直以来都是特征点匹配算法的重点,结合深度相机计算的核心也在于抑制各种噪声,减小必然误差与偶然误差。例如杂质飞入视场,目标物体具有红外吸光材料或是不规则反光形状,导致深度信息缺乏、错误。阈值精准度与直线拟合也会带来计算误差。文献[17]在标定深度相机与惯性测量单元(IMU)相对姿态过程中利用最小二乘法求解最优姿态。此类算法上的误差分析无法从根本处理噪声,导致测量精度下降。

本实验消除误差的方法和解释如下。

1) 首先在使用深度相机时,在计算机上根据环境手动控制曝光时间、增益、对比度等提高图像质量,确保图像中不出现过暗点(深度值为0)。

2) 通过二值化最前端面与工件实际尺寸,找到种子点的深度特征数,并规定获得的种子点的深度范围不超过 $(N_{\text{feature}} \pm 10)$ mm。阈值选取包含多种因素。首先,需要考虑到钢筋本身存在3~4 mm厚度,阈值需大于此范围才能囊括完整钢筋。其次,阈值不可选取过大,观察图1(a)可得,上下端面之间除离相机最近的四根钢筋外,还有其他钢筋,则在工件相对相机有较大倾斜时,阈值若选取过大,易囊括进后方不在预想情况中的部件。经实验验证,阈值设定20 mm时,图1(a)中工件沿中心轴(假设中心轴自下往上穿过工件中心)旋转的最大角度为 $\pm 6.8^\circ$,超过该范围,将囊括5或6根钢筋,不符合实验需求。即使待处理工件不含有

本实验所列工件的特征,阈值选取过大时也容易引入噪声,本实验后方预期外的钢筋也属一类噪声。再则,阈值选取不能过小,如本实验阈值需大于3~4 mm,如若选取5 mm作为阈值,经实验验证,很容易丢失特征部件。这是因为实验选取的特征数是根据测量尺测得的实际尺寸与最近端面的深度计算得到的,它是钢筋深度大概数值,本身存在一定误差,并且工件若存在较大角度倾斜,阈值过小也易丢失欲拟合的部件。综合上述多种因素,本实验选取10 mm作为阈值,实验结果表明在该阈值下,工件沿中心轴旋转角度可达 $\pm 12^\circ$,在此角度范围内任意改变工件方向拍照万次,结果足够稳定,满足实验要求。

3) 拟合特征部件时对每个即将绘出的点判断其与上一个点的空间距离,滤除视场范围内飞入杂质带来的噪声。

4) 为避免表面上细微不平整带来的误差,对最终得到的同名点,取其周围50~100个同一平面上的点计算深度平均值,以该平均值替换同名点深度值。

2.3 普适性分析

所提方法在图像坐标系中进行直线搜索后分析各个像素点在实际空间中的深度值关系,进而提取出具有特征的同名点。该图像处理方法对本实验工件具有较好的效果,能够拟合出预想的特征部件与同名点。现探讨该方法的普适性:任何一个三维物体都有其独特形态特征,不同物体在深度相机下产生不同三维数值场,因此理论上以深度为特征的同名点存在于每一个三维物体上。实际实验时,对各种具有明显深度特征的工件执行所述图像处理,处理效果良好,程序工作稳定。图2展示处理效果良好的两种工件RGB图和它们各自可选取的同名点。

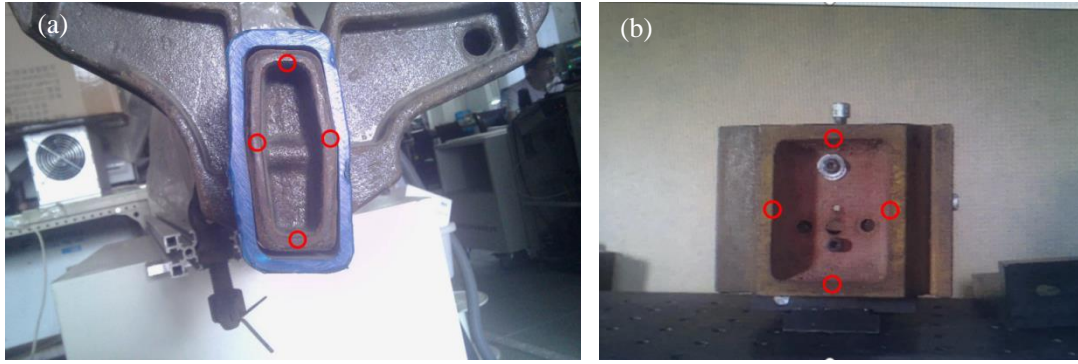


图 2 图像处理成功的工件 RGB 图。(a)工件 1 的 RGB 图;(b)工件 2 的 RGB 图

Fig. 2 RGB diagrams of workpiece with successful image processing. (a) RGB drawing of workpiece 1; (b) RGB drawing of workpiece 2

从图 2 可以清楚地看出,当工件具有凹陷、凸出、中空等明显深度特征时,图像处理效果更好。根据理论推导与实验验证得出,在物体深度特征性下降到一定程度时,处理结果误差会随着深度特征性的下降而开始增加,这个反相关大概在工件变为类似球形的状态开始。在这种形态下,工件深度值的变化呈均匀梯度上升或下降,图像处理很难捕捉某一个特征点,这对阈值选取问题提出了更高要求。但对于大部分工件而言,图像处理结果的误差为 1~10 mm,所提方法普适性较好,可适用于多种工件。

3 相机与物体相对姿态计算

相机坐标系与物体坐标系间是刚体变换,二者之间变换的表达式为

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & \mathbf{1} \end{bmatrix} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix}, \quad (3)$$

式中:\$(X_C, Y_C, Z_C)\$是相机坐标系的坐标值, \$(X_T, Y_T, Z_T)\$是物体坐标系的坐标值; \$\mathbf{R}, \mathbf{t}\$ 分别是旋转矩阵和平移矩阵。求取 \$\mathbf{R}, \mathbf{t}\$ 即相机感知物体姿态的过程,也是求解相机外参矩阵的过程。

3.1 算法原理

设 \$M = \{m_1, m_2, m_3\}\$ 为物体坐标系的三个点, \$N = \{n_1, n_2, n_3\}\$ 为相机坐标系上同名的三个点,则有

$$\mathbf{n}_i = \mathbf{R}\mathbf{m}_i + \mathbf{t} \quad (i = 1, 2, 3), \quad (4)$$

式中: \$\mathbf{n}_i\$ 和 \$\mathbf{m}_i\$ 都是 \$3 \times 1\$ 的列向量。为得到最佳 \$\mathbf{R}\$ 和 \$\mathbf{t}\$, 在各点权重相同的情况下,参考 Besl 等^[18]和 Umeyama 等^[19]分别于 1992 年和 1991 年提出的算法,将求解式(4)转换为求解最优解的问题(以下公式省略 argmin 下标变量 \$\mathbf{R}\$),并展开后续算法推导:

$$F(\mathbf{t}) = \arg \min \sum_{i=1}^3 \|\mathbf{R}\mathbf{m}_i + \mathbf{t} - \mathbf{n}_i\|^2, \quad (5)$$

$$\mathbf{0} = \frac{\partial F}{\partial \mathbf{t}} = \sum_{i=1}^3 2(\mathbf{R}\mathbf{m}_i + \mathbf{t} - \mathbf{n}_i) = 6\mathbf{t} + 2\mathbf{R}(\sum_{i=1}^3 \mathbf{m}_i) - 2(\sum_{i=1}^3 \mathbf{n}_i), \quad (6)$$

设

$$\bar{\mathbf{m}} = \frac{\sum_{i=1}^3 \mathbf{m}_i}{3}, \quad \bar{\mathbf{n}} = \frac{\sum_{i=1}^3 \mathbf{n}_i}{3}, \quad (7)$$

将式(7)代入式(6),得

$$\mathbf{t} = \bar{\mathbf{n}} - \mathbf{R}\bar{\mathbf{m}}. \quad (8)$$

将式(8)代入式(5),可消除平移向量 \$\mathbf{t}\$,从而可以先计算旋转矩阵 \$\mathbf{R}\$。设

$$\mathbf{x}_i = \mathbf{m}_i - \bar{\mathbf{m}}, \quad \mathbf{y}_i = \mathbf{n}_i - \bar{\mathbf{n}}, \quad (9)$$

则式(5)变为

$$\mathbf{R} = \arg \min \sum_{i=1}^3 \|\mathbf{R}\mathbf{x}_i - \mathbf{y}_i\|^2. \quad (10)$$

将式(10)所得的旋转矩阵 \$\mathbf{R}\$ 代入式(8),即可获得坐标系间的平移。对矩阵 \$\mathbf{R}\$ 反算,可获取旋转的欧拉角,实现深度相机对三维物体的姿态感知。下文叙述对式(10)的求解过程与反算三轴欧拉角的方法。

3.2 旋转矩阵计算

坐标系的旋转可看作是分别绕 \$X, Y, Z\$ 三轴的旋转,旋转矩阵 \$\mathbf{R}\$ 则为依次绕 \$X, Y, Z\$ 三轴旋转的效果之和,其旋转分别用 \$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\$ 表示。绕 \$X\$ 轴旋转的结果为

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (11)$$

绕 \$Y\$ 轴旋转的结果为

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (12)$$

绕 \$Z\$ 轴旋转的结果为

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (13)$$

以上旋转均以逆时针旋转为正,旋转矩阵 \$\mathbf{R}\$ 为

$$\mathbf{R} = \mathbf{r}_1 \times \mathbf{r}_2 \times \mathbf{r}_3 = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \cos \beta \cos \gamma & \cos \beta \sin \gamma & -\sin \beta \\ \sin \alpha \sin \beta \cos \gamma - \sin \gamma \cos \alpha & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \cos \beta \\ \sin \beta \cos \alpha \cos \gamma + \sin \alpha \sin \gamma & \sin \beta \sin \gamma \cos \alpha - \sin \alpha \cos \gamma & \cos \alpha \cos \beta \end{bmatrix} \quad (14)$$

由式(11)~(14)可知,旋转矩阵 \mathbf{R} 是正交矩阵,根据向量二范数定义将式(10)展开,得

$$\begin{aligned} \|\mathbf{R}\mathbf{x}_i - \mathbf{y}_i\|^2 &= (\mathbf{R}\mathbf{x}_i - \mathbf{y}_i)^T (\mathbf{R}\mathbf{x}_i - \mathbf{y}_i) = \\ & \mathbf{x}_i^T \mathbf{R}^T \mathbf{R} \mathbf{x}_i - \mathbf{y}_i^T \mathbf{R} \mathbf{x}_i - \mathbf{x}_i^T \mathbf{R}^T \mathbf{y}_i + \\ & \mathbf{y}_i^T \mathbf{y}_i = \mathbf{x}_i^T \mathbf{x}_i - 2\mathbf{y}_i^T \mathbf{R} \mathbf{x}_i + \mathbf{y}_i^T \mathbf{y}_i \quad (15) \end{aligned}$$

式(15)的计算是因正交矩阵 $\mathbf{R}^T \mathbf{R} = \mathbf{I}$ (\mathbf{I} 是单位矩阵)且 $\mathbf{x}_i^T \mathbf{R}^T \mathbf{y}_i$ 是标量,则 $\mathbf{x}_i^T \mathbf{R}^T \mathbf{y}_i = \mathbf{y}_i^T \mathbf{R} \mathbf{x}_i$ 。因只关心式(10)最优时 \mathbf{R} 的结果,式(15)中的标量部分可忽略,将式(15)代入式(10),并重塑最优解问题,得

$$\begin{aligned} \mathbf{R} &= \arg \min \left(\sum_{i=1}^3 \mathbf{x}_i^T \mathbf{x}_i - 2 \sum_{i=1}^3 \mathbf{y}_i^T \mathbf{R} \mathbf{x}_i + \sum_{i=1}^3 \mathbf{y}_i^T \mathbf{y}_i \right) = \\ & \arg \min \left(-2 \sum_{i=1}^3 \mathbf{y}_i^T \mathbf{R} \mathbf{x}_i \right) = \arg \max \sum_{i=1}^3 \mathbf{y}_i^T \mathbf{R} \mathbf{x}_i = \\ & \arg \max \left[\text{tr}(\mathbf{Y}^T \mathbf{R} \mathbf{X}) \right] = \arg \max \left[\text{tr}(\mathbf{R} \mathbf{X} \mathbf{Y}^T) \right] \quad (16) \end{aligned}$$

式中: $\mathbf{X} = (x_1, x_2, x_3)$, $\mathbf{Y} = (y_1, y_2, y_3)$; $\mathbf{X}\mathbf{Y}^T$ 称为协方差矩阵 \mathbf{H} 。对 \mathbf{H} 进行奇异值分解,得到

$$\mathbf{H} = \mathbf{X}\mathbf{Y}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (17)$$

将式(17)代入式(16)中,得

$$\mathbf{R} = \arg \max \left[\text{tr}(\mathbf{\Sigma}\mathbf{V}^T \mathbf{R} \mathbf{U}) \right] \quad (18)$$

设 $\mathbf{P} = \mathbf{V}^T \mathbf{R} \mathbf{U}$,因 \mathbf{V} 、 \mathbf{R} 、 \mathbf{U} 都是正交矩阵,则 \mathbf{P} 的列向量都是正交向量,有

$$\mathbf{p}_j^T \mathbf{p}_j = \sum_{i=1}^3 p_{ij}^2 = 1 \quad (19)$$

由式(19)可清楚看出,对于 \mathbf{P} 的每一个元素 p_{ij} ,都有 $|p_{ij}| \leq 1$ 。将式(18)部分展开,得到

$$\text{tr}(\mathbf{\Sigma}\mathbf{V}^T \mathbf{R} \mathbf{U}) = \text{tr}(\mathbf{\Sigma}\mathbf{M}) = \sum_{i=1}^3 \sigma_i p_{ii} \leq \sum_{i=1}^3 \sigma_i \quad (20)$$

当且仅当 $p_{ii} = 1$ 时,式(20)取得最大值。而对于正交矩阵 \mathbf{P} ,当它的对角元素为1时,有

$$\mathbf{P} = \mathbf{I} = \mathbf{V}^T \mathbf{R} \mathbf{U} \rightarrow \mathbf{R} = \mathbf{V} \left\{ \text{diag}[1, 1, \det(\mathbf{V}\mathbf{U}^T)] \right\} \mathbf{U}^T \quad (21)$$

此中加入 $\text{diag}[1, 1, \det(\mathbf{V}\mathbf{U}^T)]$ 这一项是为了定向矫正旋转过程中点集之间的反射,使之产生完美对点旋转。

3.3 反算欧拉角

平移向量 \mathbf{t} 的三个元素构成相对姿态的平移三自由度,旋转三自由度则要通过旋转矩阵反算欧拉角获得。根据式(14),可得

$$\begin{cases} \alpha = \arctan\left(\frac{r_{12}}{r_{22}}\right) \\ \beta = \arctan\left(\frac{-r_{02}}{\sqrt{r_{00}^2 + r_{01}^2}}\right) \\ \gamma = \arctan\left(\frac{r_{01}}{r_{00}}\right) \end{cases} \quad (22)$$

根据反正切函数的性质,其值域需在 180° 范围内,在计算机上可获取反正切函数的唯一值。设置物体坐标系的方向与相机坐标系一致,在旋转角度不超过 180° 的情况下,根据式(22)获取三轴唯一欧拉角。

3.4 算法总结与分析

通过计算机编程,完成3.1节、3.2节和3.3节所述计算,步骤归纳如下。

1) 求三个同名点在物体坐标系与相机坐标系上的质心:

$$\begin{cases} \mathbf{M}_{\text{center}} = \frac{1}{3} \sum_{i=1}^3 \mathbf{X}_i^T \\ \mathbf{N}_{\text{center}} = \frac{1}{3} \sum_{i=1}^3 \mathbf{X}_i^C \end{cases} \quad (23)$$

2) 构成协方差矩阵,并对其进行奇异值分解,计算旋转矩阵 \mathbf{R} 和平移向量 \mathbf{t} :

$$\begin{cases} \mathbf{H} = \sum_{i=1}^3 (\mathbf{X}_i^T - \mathbf{M}_{\text{center}})(\mathbf{X}_i^C - \mathbf{N}_{\text{center}})^T \\ [\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \text{SVD}(\mathbf{H}) \\ \mathbf{R} = \mathbf{V} \left\{ \text{diag}[1, 1, \det(\mathbf{V}\mathbf{U}^T)] \right\} \mathbf{U}^T \\ \mathbf{t} = \mathbf{N}_{\text{center}} - \mathbf{R} \times \mathbf{M}_{\text{center}} \end{cases} \quad (24)$$

3) 将旋转矩阵 \mathbf{R} 中的元素值代入式(22)中,计算所得的三轴欧拉角与平移向量 \mathbf{t} 的三个元素,构成六自由度,完成深度相机对物体的三维姿态感知。

算法整体包含坐标系刚体变换、正交矩阵性质求解、奇异值分解、协方差矩阵计算、反三角函数计算等。实验基于几个方面,仅代入三个特征同名点参与计算:对于每一次图像处理后的同名点,都需求取质心与协方差矩阵,若代入点数太多,会增大计算机负载,降低程序工作效率,与本实验要求的高效执行姿态感知不符;图像处理难以做到绝对理想状态,意味着每个同名点坐标值都将产生或多或少的误差,这取决于多种因素,若输入同名点数过多,在算法理论上结果更加稳定,但在图像处理中引入了更多误差;深度相机感知的是三维物体,物体上含有许多倾斜平面,而距离远的同名点更能代表一个平面,对于本实验工件,选取三个最远点皆可;图像处理获取了四个同名点,只选用三个是因考虑到三个点与四个点代入计算的结果的稳定性都很高,且三个点代入计算所得结果的误差早已满足实验需求。再则,在四个点中选取三个点进行计算,即可组合为四次结果的数据,而四次结果之间可以比较,也可以进行平均,这对实验结果的分析有益。

大量实验结果表明,计算精度几乎取决于输入同名点坐标值的准确性。将图1(d)中获取到的四个特征点组合成四组,根据上述分析所讲,将每组三个同名点所对应相机坐标系与物体坐标系上的坐标值分别代

入式(23)和式(24)中,计算六自由度,将它们平均值作为姿态测量结果,从而进一步减小误差。

4 实验结果与分析

本实验系统主要由六轴机器人、深度相机、待测工件与载物台搭建而成。六轴机器人末端装上一个夹持器用来固定相机,通过PC机控制机器人六轴进行旋转或平移,从而移动相机位置,使得相机可以从任何方向对准待测工件。待测工件固定于平整的载物台上。软件全程控制系统运作,使用VS2017作为软件开发环境,搭建opencv用作部分图像处理,配置QT设计软件UI。图3是实际工作中机器人手夹持深度相机与激光器,图4是设计的软件操作界面。



图3 机器人手夹持深度相机与激光器

Fig. 3 Robot hand holding depth camera and laser

实验采用型号为Surface 60的深度相机,测量距离为200~300 mm,被测工件尺寸为200 mm×120 mm。程序循环工作,周期约为1秒/次。

对于本实验整个系统,评价其性能的参数是绝对精

度与相对精度。绝对精度是指深度相机测量所得物体姿态参数与真实空间中深度相机与被测物体相对姿态参数之间的差值。而深度相机拍摄的深度图像是由左右两幅图像重建形成的,因而很难得知相机坐标系原点在相机哪个部位,探讨实验结果时也只能大致估计其原点在相机中部,绝对精度测量下,会产生1~5 mm的误差。其次,如果要准确测量真实空间中相机与目标物体间的六个自由度,需使用非常标准的平台夹持相机和物体,也需用更为精准的测量尺和量角器来测量X、Y、Z三个方向的平移和旋转,此方法可将绝对精度测量误差降低到1 mm左右,本实验对绝对精度的要求并没有达到这个程度。虽降低绝对精度测量误差意义较小,但须探讨系统对于姿态感知的正确性:首先默认相机坐标系原点位于左右镜头中点,通过测量尺测量相机坐标系和工件坐标系中心在X、Y、Z三个方向上的平移量,将该参量与软件计算所得平移量对比,若差值在10 mm以内,认为平移量计算正确;其次,在深度相机旁夹持一个激光器,程序拍照计算后,使激光器移动到拍照时相机所处位置,并将软件计算所得的六个自由度位姿参量传输到机器人程序中,令机器人根据此参量进行旋转和平移。打开激光器,观察到激光光点或激光发射头大致位于工件坐标系的中心轴(Z轴)上,如图3所示,则本实验位姿参量计算正确。

不论使用何种深度相机,拍摄物体深度图像时,相机本身会带来难以避免的误差。即当相机与目标物体相对位置没有发生变化时,每次拍摄获取的三维数值场并不相同。鉴于此,图像处理能否克服此类不稳定性是实验结果分析的一个重点,而本实验相对精度指在相机与被测物体不移动时,多次拍照计算后分别得到的位姿参量之间的差值。对同一工件多次拍照,绘出折线图,并观察计算结果的相对重复性。工件在三种不同姿态下的处理结果如图5~7所示。

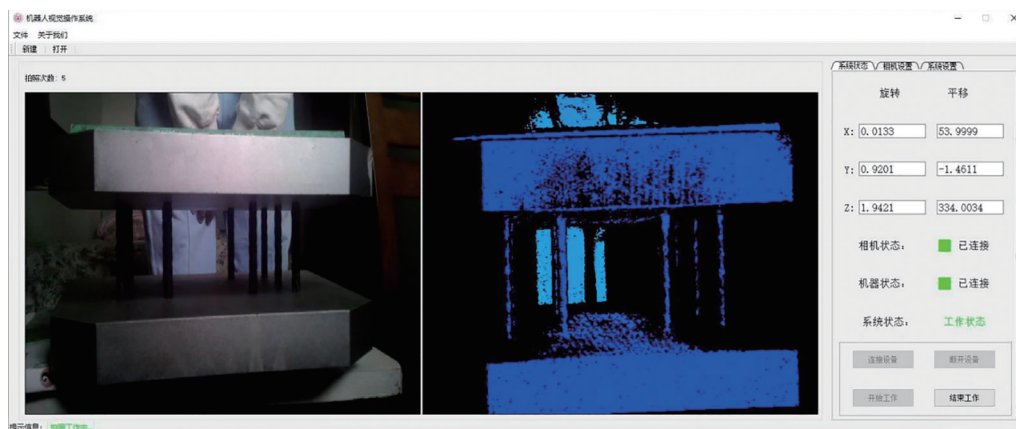


图4 软件操作界面

Fig. 4 Software operation interface

折线图表示每次拍下一帧图像的处理结果(六自由度)与该种姿态下初次拍照计算结果相减所得的欧拉角

与平移量之差,用以探讨系统相对重复性。三种姿态初次拍照后计算所得的六个自由度参量列于表1。

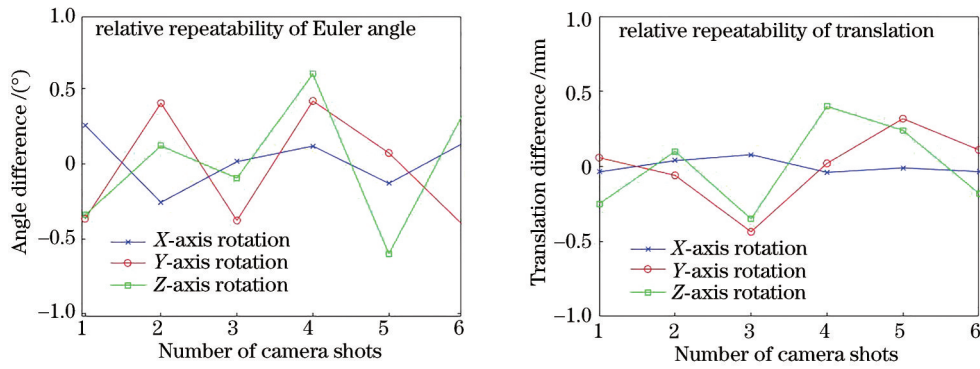


图 5 物体在第一种姿态下的欧拉角与平移量的重复性折线图

Fig. 5 Repetitive line graph of Euler angle and translation of object in the first attitude

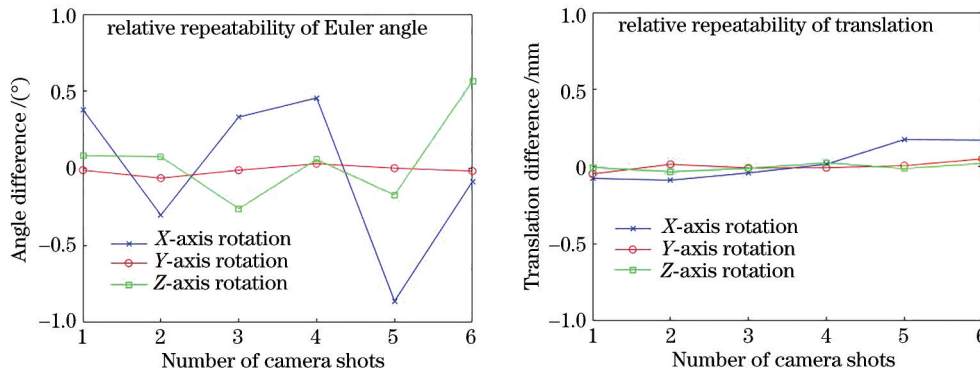


图 6 物体在第二种姿态下的欧拉角与平移量的重复性折线图

Fig. 6 Repetitive line graph of Euler angle and translation of object in the second attitude

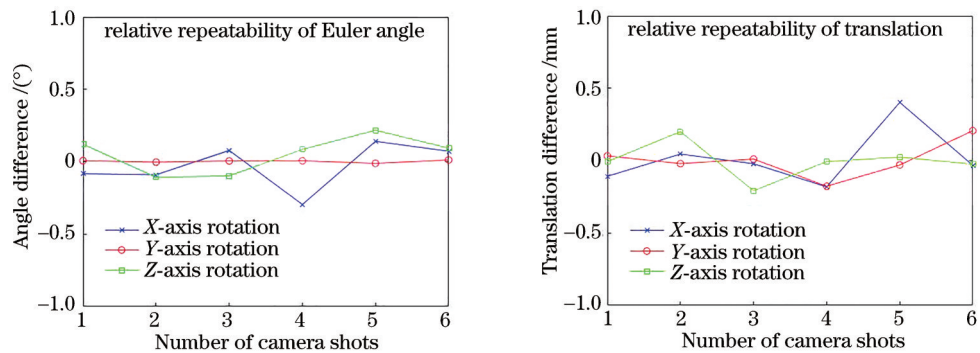


图 7 物体在第三种姿态下的欧拉角与平移量的重复性折线图

Fig. 7 Repetitive line graph of Euler angle and translation of object in the third attitude

表 1 实验中物体相对于相机的三种姿态

Table 1 Three postures of the object relative to the camera in the experiment

Attitude	Rotation (X) / (°)	Rotation (Y) / (°)	Rotation (Z) / (°)	translation (X) / mm	translation (Y) / mm	translation (Z) / mm
Attitude 1	-7.714	1.727	1.663	-5.807	-28.484	228.57
Attitude 2	68.110	88.973	72.399	47.436	-3.086	270.11
Attitude 3	-1.055	0.322	4.365	47.314	3.110	270.06

实验采用的深度相机处理该工件的效果良好,相对重复性高,旋转角度几乎不超过 0.5°,平移量最大不超过 1 mm,对各种噪声有较为良好的抑制,对三维物体的姿态感知误差也在允许范围内。但相机坐标系与

物体坐标系间的六自由度真值难以精确测量,绝对误差的结果采用一些测量器件进行估计,实验结果表明了姿态测量的正确性。但目标工件若距相机太远,易丢失深度信息,产生过多暗点,影响实验结果。

5 结 论

提出一种计算物体上三个同名点在物体坐标系和相机坐标系下的坐标值进而通过协方差矩阵、SVD、反三角函数等数学应用获取六个自由度位姿参数的方法。详细阐述了在三维数值场中利用深度信息准确匹配特征点,同时合理抑制各种噪声,减小了误差的内容。实验结果表明,在不同姿态下欧拉角的相对误差在 $\pm 0.5^\circ$ 内,平移量的相对误差在 ± 1 mm 内。成功计算出物体坐标系与相机坐标系之间的旋转与平移,实现了相机对物体三维姿态的感知。所提方法具有原理简单、成本低廉、易于操作、适用于各类环境、无需接触测量等特点,可应用于车间机器运作,无需人工标定。但相对姿态测量受物体特征的影响,对于特征变化十分不明显的物体,所提方法对姿态的感知精度会下降,且对于深度范围有一定的需求。未来的研究工作将围绕如何降低对物体特征信息的需求及如何加强图像去噪,从而进一步提高三维数值场中同名点提取的准确性来进行,这是提高三维姿态感知精度的核心问题。

参 考 文 献

- [1] 尹仕斌,任永杰,刘涛,等. 机器视觉技术在现代汽车制造中的应用综述[J]. 光学学报, 2018, 38(8): 0815001. Yin S B, Ren Y J, Liu T, et al. Review on application of machine vision in modern automobile manufacturing[J]. Acta Optica Sinica, 2018, 38(8): 0815001.
- [2] Tarabani K A, Allen P K, Tsai R Y. A survey of sensor planning in computer vision[J]. IEEE Transactions on Robotics and Automation, 1995, 11(1): 86-104.
- [3] Dong J T, Lu R S, Shi Y Q, et al. Optical design of color light-emitting diode ring light for machine vision inspection[J]. Optical Engineering, 2011, 50(4): 043001.
- [4] Yi S K, Haralick R M, Shapiro L G. Optimal sensor and light source positioning for machine vision[J]. Computer Vision and Image Understanding, 1995, 61(1): 122-137.
- [5] Quan W, Kubota N. Multi-view 3D human pose tracking based on evolutionary robot vision[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2021, 25(4): 432-441.
- [6] 张磊,徐孝彬,曹晨飞,等. 基于动态特征剔除图像与点云融合的机器人位姿估计方法[J]. 中国激光, 2022, 49(6): 0610001. Zhang L, Xu X B, Cao C F, et al. Robot pose estimation method based on image and point cloud fusion with dynamic feature elimination[J]. Chinese Journal of Lasers, 2022, 49(6): 0610001.
- [7] 梁挺,韩瑞,丁昨凯,等. GPS 三维姿态测量技术研究[J]. 中国新通信, 2017, 19(9): 75-76. Liang T, Han R, Ding Z K, et al. Research on GPS 3D attitude measurement technology[J]. China New Telecommunications, 2017, 19(9): 75-76.
- [8] Qayyum U, Kim J. Depth-camera-aided inertial navigation utilizing directional constraints[J]. Sensors, 2021, 21(17): 5913.
- [9] 吴禹彤,张林. 新型优化策略在视觉惯性导航系统中的应用[J]. 计算机应用, 2020, 40(S2): 215-220. Wu Y T, Zhang L. Application of novel optimization strategy in visual inertial navigation system[J]. Journal of Computer Applications, 2020, 40(S2): 215-220.
- [10] 孙国燕,高立民,白建明,等. 三维姿态角高精度测量装置[J]. 光学精密工程, 2016, 24(5): 963-970. Sun G Y, Gao L M, Bai J M, et al. High accuracy three-dimensional attitude angle measuring device[J]. Optics and Precision Engineering, 2016, 24(5): 963-970.
- [11] 束安,裴浩东,丁雷,等. 空间非合作目标的双目视觉位姿测量方法[J]. 光学学报, 2020, 40(17): 1712003. Shu A, Pei H D, Ding L, et al. Binocular visual position and attitude measurement method for a spatial non-cooperative target[J]. Acta Optica Sinica, 2020, 40(17): 1712003.
- [12] 邓烈刚,李文峰,杨雁南. 基于双目机器视觉的高精度激光充电瞄准系统[J]. 激光与光电子学进展, 2021, 58(14): 1415006. Deng L G, Li W F, Yang Y N. High-precision aiming system for laser charging based on binocular machine vision[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1415006.
- [13] 解则晓,邵玮婧,高翔,等. 基于离散外极曲线模型的水下多线结构光双目测量方法[J]. 中国激光, 2021, 48(13): 1304003. Xie Z X, Shao W J, Gao X, et al. Underwater multiple line-structured light binocular measuring method based on discrete epipolar curve model[J]. Chinese Journal of Lasers, 2021, 48(13): 1304003.
- [14] 魏弦,张敬东,吴明卿. 基于3D视觉识别的工件姿态研究[J]. 煤矿机械, 2013, 34(8): 72-73. Wei X, Zhang J D, Wu M Q. Research on workpiece posture based on 3D visual recognition[J]. Coal Mine Machinery, 2013, 34(8): 72-73.
- [15] 束安,裴浩东,周姗姗,等. 非合作航天器的立体视觉位姿测量[J]. 光学精密工程, 2021, 29(3): 493-502. Shu A, Pei H D, Zhou S S, et al. Stereo measurement of position and attitude of non-cooperative spacecraft[J]. Optics and Precision Engineering, 2021, 29(3): 493-502.
- [16] 阴雷,王向军,秦贯宇. 基于视觉结构光的大型曲面动态测量方法[J]. 激光与光电子学进展, 2020, 57(13): 131202. Yin L, Wang X J, Qin G Y. Dynamic measurement method of large-scale surface based on visual structured light technology[J]. Laser & Optoelectronics Progress, 2020, 57(13): 131202.
- [17] 陈集辉,朱海飞,谷世超,等. 深度相机与惯性测量单元的相对姿态标定[J]. 哈尔滨工业大学学报, 2018, 50(11): 131-136. Chen J H, Zhu H F, Gu S C, et al. Relative orientation calibration of a depth camera and an inertial measurement unit[J]. Journal of Harbin Institute of Technology, 2018, 50(11): 131-136.
- [18] Besl P J, McKay N D. A method for registration of 3-D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [19] Umeyama S. Least-squares estimation of transformation parameters between two point patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(4): 376-380.