

# 基于多传感器的三维目标位姿测量方法

潘旺<sup>1,2\*\*</sup>, 朱枫<sup>1,3,4\*</sup>, 郝颖明<sup>1,3,4</sup>, 张丽敏<sup>1,2</sup>

<sup>1</sup>中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;

<sup>2</sup>中国科学院大学, 北京 100049;

<sup>3</sup>中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;

<sup>4</sup>辽宁省图像理解与视觉计算重点实验室, 辽宁 沈阳 110016

**摘要** 提出了一种基于多传感器的三维目标位姿测量方法,该方法利用多传感器技术,充分发挥深度相机和高分辨率电荷耦合器件(CCD)相机各自的优势,提高了测量的稳健性和效率。利用物体与其固定平面之间的关系,在点云中粗略定位出目标区域,通过预先标定信息将目标区域转换至灰度图像空间。在灰度图像中,利用线段检测器(LSD)算法外加特征约束,筛选出 4 条目标直线,利用透视 4 点(P4P)算法求解出目标的六维位姿。实验验证了该算法的有效性,其测量效率远优于经典模板匹配方法。

**关键词** 机器视觉; 位姿测量; 多传感器; 三维目标; 直线提取

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201939.0212007

## Pose Measurement Method of Three-Dimensional Object Based on Multi-Sensor

Pan Wang<sup>1,2\*\*</sup>, Zhu Feng<sup>1,3,4\*</sup>, Hao Yingming<sup>1,3,4</sup>, Zhang Limin<sup>1,2</sup>

<sup>1</sup>Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China;

<sup>3</sup>Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

<sup>4</sup>Key Laboratory of Image Understanding and Computer Vision of Liaoning Province, Shenyang, Liaoning 110016, China

**Abstract** A method based on multi-sensor to measure the pose of a three-dimensional object is proposed. We employ multi-sensor technology to make the advantages of the depth camera and the high-resolution charge-couple device (CCD) camera, improving the robustness and efficiency of measurement. The target area is coarsely located in the point cloud based on the relationship between the object and its fixed board, and the region is converted to gray image space by the prior calibration information. In the gray image, four target straight lines are extracted by the Line Segment Detector (LSD) and feature constraints. The perspective 4 point (P4P) algorithm is utilized to calculate the six-dimensional pose of the object. The experiment verifies the validity of the algorithm and the measurement efficiency of the algorithm is much better than that of the classical template matching method.

**Key words** machine vision; pose measurement; multi-sensor; three-dimensional object; line extraction

**OCIS codes** 150.4232; 150.6910; 120.3940; 120.4820

## 1 引 言

视觉感知系统通过视觉传感器确定目标物体在三维空间中的位置和姿态,对机器人的抓取操作具有重要意义。目前广泛使用的基于描述子的方

法<sup>[1-3]</sup>通过在物体的特征点周围构造具有区分性的特征描述子来训练分类器。在搜索阶段,将搜索图像与模型进行匹配来确定目标物体。这种方法最大的优点是其执行时间与几何搜索空间的大小无关,在一些应用场景下表现出色,但其仅局限于纹理丰

收稿日期: 2018-08-16; 修回日期: 2018-09-27; 录用日期: 2018-10-08

基金项目: 国家自然科学基金(U1713216)

\* E-mail: fzhu@sia.cn; \*\* E-mail: panwang@sia.cn

富、清晰的物体。这是因为只有当物体的纹理比较清晰时,才能检测到有意义的特征描述子。在实际工程应用中,需要测量三维模型已知的无纹理物体的六维位姿,该类物体表面光滑有光泽,几乎没有纹理,在图像获取点未知的情况下,对三维物体所在的广阔空间进行搜索非常耗时,因此对其精确测量的难度较大。Ulrich 等<sup>[4-5]</sup>提出的基于二维视图形状模板匹配的方法是一种较好的解决方案。该方法的视图模板采样间隔小,匹配精度高,采用了由粗到精的加速策略。但其在在大场景中的稳健性较差,特别是在复杂混乱背景下极易出现误匹配,并且在建模的姿态范围中不能包含物体的退化视图,否则易导致误匹配。Hinterstoisser 等<sup>[6-7]</sup>采用一种改进的模板匹配方法来检测并定位无纹理物体,该方法综合利用了彩色图像中的梯度信息和深度图像中的表面法向量信息。然而,模板的稀疏采样限制了该方法初始估计位置的准确性,不能输出物体精确的六维姿态,且同样存在视图退化情况下误匹配率较高的问题。当待测量物体的几何特征比较明显时,拟采用传统的透视  $n$  点(PnP)问题求解算法<sup>[8-10]</sup>计算物体的精确位姿。当已知物体上 4 个特征点的三维空间坐标及其对应的成像坐标时,采用该方法可以解算出物体的位姿。在图像中搜索并对应目标特征点是问题的关键,然而图像搜索空间较大,且金属表面的反射特性使得物体的成像受光照的影响明显,因此在图像中找到目标特征点的难度较大。

针对这些问题,本文提出了一种基于多传感器的三维目标位姿测量方法,该方法采用由粗到精的定位策略,综合利用了深度相机和电荷耦合器件(CCD)相机各自的优势,使得两种传感技术形成优

云精度有限,且噪声较大,所以无法实现三维目标的高精度位姿测量,且视觉特征没有二维图像直观。CCD 相机采集的二维图像数据的分辨率和精度均优于深度相机,但用于三维目标位姿测量时存在搜索空间巨大、搜索效率低等问题。三维点云用于空间平面拟合具有传统二维相机无法比拟的优势,且对光照变化具有较好的稳健性。因此:首先从点云中拟合出物体所在平面,进而分割出目标区域,完成目标物体的粗略定位;然后将该区域通过预先标定信息转换到灰度图像空间,划定感兴趣区域,从而减小二维图像中的搜索空间,弥补 CCD 相机用于三维目标位姿测量时的劣势。利用线段检测器(LSD)<sup>[11-12]</sup>直线提取算法和特征约束提取出 4 条目标直线,求出它们之间的交点。亚像素级的图像特征点提取弥补了深度相机分辨率低、精度差的劣势。最终利用透视 4 点(P4P)求解算法得到了物体的高精度位姿。两种传感技术的综合利用避免了大面积复杂背景的干扰,解决了三维目标成像退化时的误匹配问题,同时缩小了搜索空间,大幅提高了效率,实现了三维目标位姿的快速高精度测量。

## 2 视觉系统构建

为仿人机器人设计的视觉测量系统主要由基于飞行时间测距的深度相机 SwissRanger SR4000 (Mesa Imaging 公司,瑞士)和紧凑型 CCD 工业相机 Prosilica GC2450C (Allied Vision Technologies 公司,德国)组成。系统的硬件组成及深度相机坐标系  $O_d X_d Y_d Z_d$ 、CCD 相机坐标系  $O_c X_c Y_c Z_c$  和世界坐标系  $O_w X_w Y_w Z_w$  的定义如图 1 所示,其中红色、绿色和蓝色箭头分别代表各坐标系的  $x$  轴、 $y$  轴和

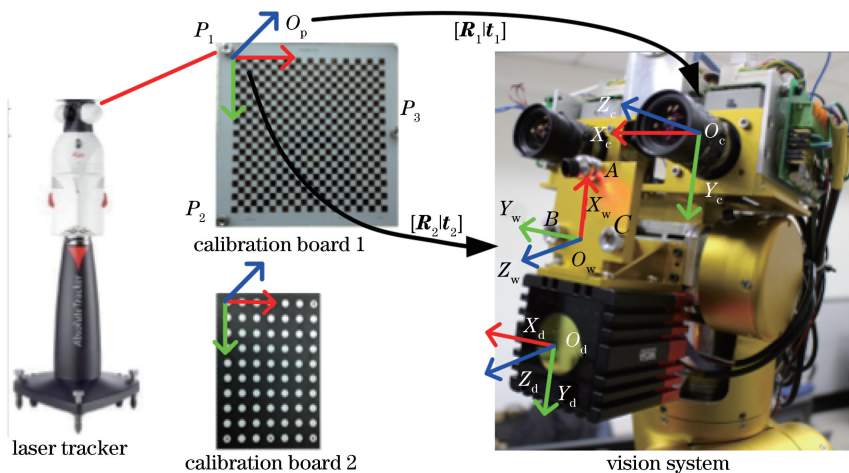


图 1 视觉系统的坐标系定义和标定

Fig. 1 Coordinate system definition and calibration of the visual system

$z$  轴。定义的坐标系均满足右手法则,深度相机坐标系  $O_d$  位于光轴与相机正面的交点处,假设观察者站在相机后面,观察方向与相机视角一致,则从相机出发沿光轴方向为  $z$  轴正方向, $y$  轴正方向垂直向下, $x$  轴正方向水平向右。CCD 相机坐标系  $O_c$  位于相机的光心, $x$  轴和  $y$  轴与图像的  $x$  轴、 $y$  轴平行, $z$  轴为相机光轴,它与图像平面垂直。世界坐标系  $O_w$  定义在与视觉系统固连的金属板上,板上加工有 3 个通孔用于放置激光跟踪仪的反射靶球,以获取目标物体在世界坐标系下的高精度真实位姿,作为视觉测量参考的标准值。为确保多传感器能同时采集数据,深度相机被用作同步触发源。本研究仅使用两个 CCD 相机中的一个,另一个作为备用,当然也可以作为冗余测量来增加系统的可靠性。

### 2.1 多传感器标定

为了准确测量物体的六维位姿,需要先对视觉系统进行标定。深度相机在出厂前已由制造商完成标定,这里仅标定其外参。如图 1 所示,标定板 1 用于标定 CCD 相机的内外参,标定板 2 用于标定深度相机的外参。用张正友标定法<sup>[13]</sup>完成 CCD 相机内参矩阵的标定,透镜模型为

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \quad (1)$$

式中: $Z_c$  为尺度因子; $u$ 、 $v$  为图像坐标; $X_w$ 、 $Y_w$ 、 $Z_w$  为世界坐标系下的坐标; $f_x$ 、 $f_y$ 、 $u_0$ 、 $v_0$  和畸变系数构成相机内参; $\mathbf{R}$ 、 $\mathbf{t}$  分别为  $3 \times 3$  的旋转矩阵和  $3 \times 1$  的平移向量,构成相机外参。外参表示从世界坐标系到相机坐标系的刚体变换<sup>[14]</sup>。激光跟踪仪可以精确获取反射靶球的三维坐标,借助瑞士 Leica Geosystems 公司生产的 AT901 激光跟踪仪标定各相机的外参。多传感器视觉系统的外参标定如图 1 所示。首先将 3 个反射靶球放置在金属板的 3 个靶座位置上,用激光跟踪仪得到这 3 个点的坐标  $A(x_0, y_0, z_0)$ 、 $B(x_1, y_1, z_1)$ 、 $C(x_2, y_2, z_2)$ 。将  $\mathbf{BC}=(x_2-x_1, y_2-y_1, z_2-z_1)$  作为世界坐标系  $O_w X_w Y_w Z_w$  的  $y$  轴,由  $A$  点向直线  $BC$  作垂线得到垂足  $D(x_3, y_3, z_3)$ ,  $\mathbf{AD}=(x_3-x_0, y_3-y_0, z_3-z_0)$  作为世界坐标系  $O_w X_w Y_w Z_w$  的  $x$  轴, $x$  轴又乘  $y$  轴得到  $z$  轴,至此世界坐标系建立完毕。然后将 3 个靶球放到标定板 1 的  $P_1$ 、 $P_2$ 、 $P_3$  三个靶座位置,激光跟踪仪可以直接测得这 3 个点在世界坐标系  $O_w$  下的坐标。已知这 3 个点在标定板 1 坐标系(以

$O_p$  为原点)下的坐标,可以计算该坐标系与  $O_w X_w Y_w Z_w$  之间的变换关系  $[\mathbf{R}_2 | \mathbf{t}_2]$ 。同时,标定板 1 坐标系与 CCD 相机坐标系  $O_c X_c Y_c Z_c$  之间的变换关系  $[\mathbf{R}_1 | \mathbf{t}_1]$  可以通过标定板上的三维点坐标和二维成像点坐标之间的对应关系求得。于是得到了从  $O_w X_w Y_w Z_w$  到  $O_c X_c Y_c Z_c$  的变换矩阵,即 CCD 相机的外参矩阵。使用标定板 2,按照上述方法完成深度相机的外参标定。内外参的标定结果在实验部分给出。完成多传感器的标定后,即可实现空间一点在不同坐标系下的坐标对应。设空间一点在坐标系  $O_c X_c Y_c Z_c$ 、 $O_d X_d Y_d Z_d$ 、 $O_w X_w Y_w Z_w$  下的坐标向量分别为  $\mathbf{P}_c$ 、 $\mathbf{P}_d$ 、 $\mathbf{P}_w$ ,则有

$$\mathbf{P}_c = \mathbf{R}_c \cdot \mathbf{P}_w + \mathbf{t}_c, \quad (2)$$

$$\mathbf{P}_d = \mathbf{R}_d \cdot \mathbf{P}_w + \mathbf{t}_d, \quad (3)$$

式中: $[\mathbf{R}_c | \mathbf{t}_c]$  为 CCD 相机的外参; $[\mathbf{R}_d | \mathbf{t}_d]$  为深度相机的外参。将(3)式两边同时左乘  $\mathbf{R}_d^{-1}$ ,可得

$$\mathbf{R}_d^{-1} \cdot \mathbf{P}_d = \mathbf{P}_w + \mathbf{R}_d^{-1} \cdot \mathbf{t}_d. \quad (4)$$

将  $\mathbf{P}_w$  代入(2)式中,可得到

$$\mathbf{P}_c = \mathbf{R}_c \cdot (\mathbf{R}_d^{-1} \cdot \mathbf{P}_d - \mathbf{R}_d^{-1} \cdot \mathbf{t}_d) + \mathbf{t}_c. \quad (5)$$

将空间一点在深度相机坐标系下的坐标  $\mathbf{P}_d$  代入(5)式即可计算出其在 CCD 相机坐标系下的坐标  $\mathbf{P}_c$ ,再利用 CCD 相机内参可得到对应的成像坐标  $u$ 、 $v$ 。所提方法利用这一对应关系在灰度图像中确定目标特征点的搜索范围。

## 3 定位策略

当机器人与目标有一定距离时,视觉传感器感知的场景较大,会采集到较多复杂混乱的背景信息,这些信息不仅会干扰对目标特征的搜索定位,还会大幅降低搜索效率。如果仅利用灰度图像对无纹理物体进行位姿测量,在整个图像空间中进行穷尽搜索,则需耗费大量无用的计算成本。尤其是在杂乱背景下,搜索效率极低。当背景中存在与物体相似的特征时,它们易被误判为目标,导致算法不稳定,稳健性差。仅利用点云信息时,点云匹配同样面临着待匹配点数量大、遍历耗时长的的问题,且基于飞行时间原理采集的点云数据精度通常较低。所使用的 SR4000 深度相机名义上在 99% 目标反射率和 0.8~8.0 m 的标定范围内,对于相机中心的 11 pixel  $\times$  11 pixel,其绝对精度为 15 mm,远不及基于灰度图像的测量精度。为了解决复杂背景和巨大搜索空间带来的问题,并满足一定的测量精度和效率,提出了一种基于多传感器的位姿测量方法。该方法主要分为两个阶段,首先利用深度相机采集的点云实现感

兴趣区域的分割,再用 CCD 相机实现高精度的位姿测量,整个算法流程如图 2 所示。借助点云的三维信息优势,用随机采样一致性(RANSAC)算法估计物体所在平面,分割出固定物体的面板区域,从场景中移除混乱背景,进而用欧式聚类算法粗略分割出目标区域。然后将分割结果变换到灰度图像中,利用 LSD 直线检测算法外加特征约束,提取出 4 条目标直线,进而求出它们的交点。最后根据这 4 个共面特征点,利用 P4P 算法计算目标位姿,实现高精

度的位姿测量。待测目标物体的三维计算机辅助设计(CAD)模型及其坐标系定义如图 3(a)所示,其中, $O_0$ 为物体坐标系  $O_0x_0y_0z_0$  的原点, $a$ 、 $b$ 、 $c$ 、 $d$  为 P4P 算法用到的 4 个目标特征点,它们在目标物体坐标系下的坐标分别为  $(-19, -175, -38.5)$ ,  $(-19, 175, -38.5)$ ,  $(19, 175, -38.5)$ ,  $(19, -175, -38.5)$ 。位姿测量结果如图 3(b)所示,其中绿色线条是根据位姿测量结果将物体模型反投影到图像中的效果。

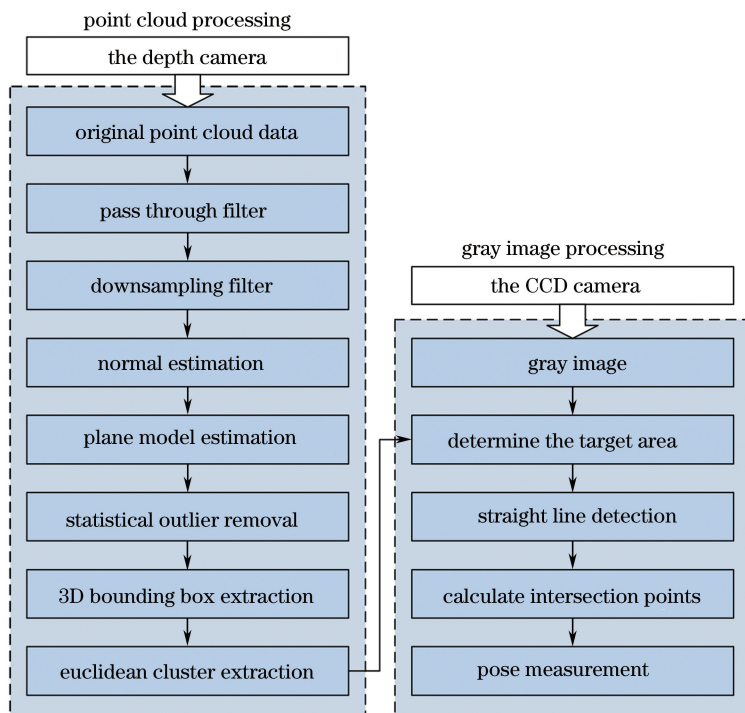


图 2 所提方法的流程图

Fig. 2 Flow chart of the proposed method

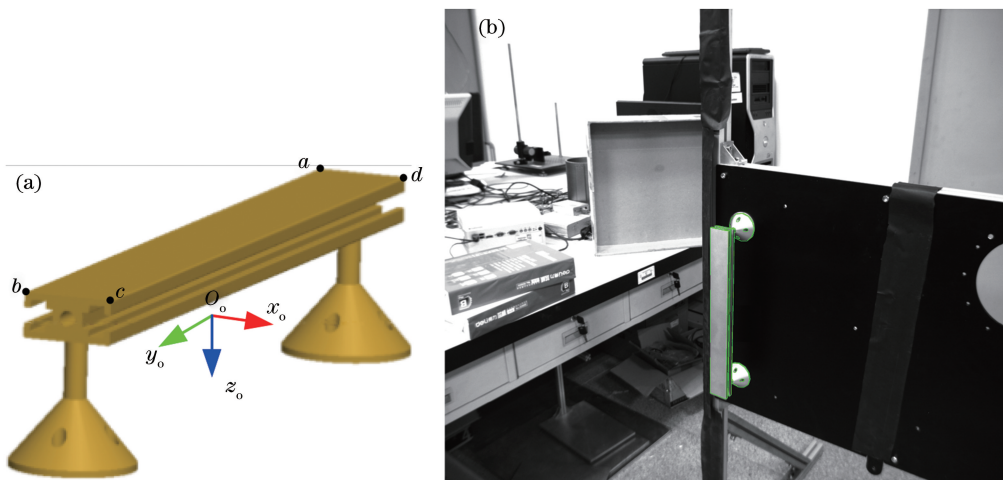


图 3 目标模型和测量结果。(a)物体的三维模型和坐标系定义;(b)位姿测量结果

Fig. 3 Object model and measurement results. (a) Three-dimensional model and definition of coordinate system of the object; (b) pose measurement result

### 3.1 基于点云的目标分割

通过深度相机获取点云数据,估计出物体所在平面,进而实现物体的粗定位。将这些点坐标转换到灰度图像坐标系,一方面可以避免杂乱背景的干扰,减小搜索空间,提高效率,另一方面有利于在图像中筛选得到目标直线。此外,根据分割出的点云的空间坐标值,可与最终位姿求解结果形成冗余信息,进行互验证,提高测量结果的可靠性。分割是点云数据处理中的关键操作<sup>[15]</sup>,点云库(PCL)<sup>[16]</sup>提供了点云处理的基础操作框架。

#### 3.1.1 点云预处理

从SR4000相机输出的 $x$ 、 $y$ 、 $z$ 笛卡儿坐标分别采用 $3\text{ pixel} \times 3\text{ pixel}$ 中值滤波预处理,以消除噪声,然后将它们构成原始点云,大小为 $176 \times 144$ ,共计25344个点。考虑到过多的数据点会增大后续算法的计算开销,且感兴趣点是与机器人的距离在一定范围内的点,因此,首先沿 $z$ 轴进行一次简单的直通滤波,消除距离过近和过远的点:

$$z_{\min} \leq z \leq z_{\max}, \quad (6)$$

设定 $z$ 轴的距离范围为 $0.3 \sim 1.5\text{ m}$ ,这取决于机器人操作范围和实际应用需求。用直通滤波器遍历一遍点云有两个作用:1)删除虚假的无效点;2)移除位于设定区间之外的点,为后续操作减小运算量。

#### 3.1.2 估计平面

首先从点云中得到物体所固定的平面,平面方程定义为 $Ax + By + Cz + D = 0$ ,其法向量 $\mathbf{N} = (A, B, C)$ 。这一步的目标是找到最适合的参数 $(A, B, C, D)$ 来确定目标物体所在的平面。很多点不在平面上,属于离群干扰点,经典的拟合技术(如最小二乘法)在这种情况下处理的效果不佳。RANSAC算法是目前应用广泛的算法之一<sup>[17-19]</sup>,其优势在于能够对模型参数进行稳健估计,即使是在数据集中存在大量异常值的情况下,也可以高准确度地估计参数。理论上三个点即可确定一个平面方程,RANSAC算法的思想是随机选择三个点来确定一个平面,再利用距离阈值判断剩余点中有多少个点属于该平面,迭代此过程,得到包含最多局内点的平面,再对这些局内点用最小二乘法优化求精得到平面方程。

由于迭代过程非常耗时,在此之前使用体素网格方法对点云进行降采样,同时保持其形状特征和细节。位于每个体素内的点都由它们的质心表示。网格的大小对RANSAC算法效率的影响如图4所示。

可以看出,降采样可移除大量数据点,得到具有

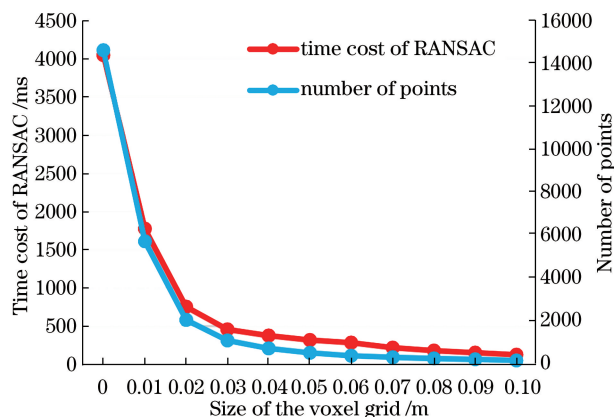


图4 网格大小对采样剩余点数和RANSAC算法效率的影响

Fig. 4 Influence of grid size on the number of sampling points and efficiency of RANSAC

较少点的等效点云,显著提高RANSAC算法的效率。在估计平面时除了点到平面距离的限制,还有法向量夹角的限制,即局内点的法向量与估计平面的法向量应近似平行。每个点周围选取50个邻域点,计算表面法向量。采用RANSAC算法估计平面,在迭代遍历过程中判断点属于平面内点的依据是点到平面的欧式距离和法向量夹角的加权和小于特定阈值。假设点 $i$ 的坐标为 $(x_0, y_0, z_0)$ ,则点到平面的欧式距离为

$$D_{\text{euclid}} = |Ax_0 + By_0 + Cz_0 + D| / \|\mathbf{N}\|. \quad (7)$$

点的法向量 $\mathbf{n}_i$ 与平面法向量 $\mathbf{N}$ 之间的夹角为

$$D_{\text{normal}} = \arccos \frac{\mathbf{n}_i \cdot \mathbf{N}}{\|\mathbf{n}_i\| \cdot \|\mathbf{N}\|}. \quad (8)$$

为了综合考虑这两项约束,实现欧式距离和法向量夹角的协同优化,点与平面模型拟合程度的评价指标为

$$E = \omega \cdot D_{\text{normal}} + (1 - \omega) \cdot D_{\text{euclid}}, \quad (9)$$

式中: $\omega$ 为法向量夹角的加权系数,取值范围为 $0 \sim 1$ 。设 $\omega$ 为 $0.1$ ,误差阈值 $E$ 为 $6\text{ cm}$ ,这意味着当误差不大于 $6\text{ cm}$ 时,该点被认为是平面模型的局内点。阈值的大小与面板的表面反射率密切相关。所用的深度相机是基于飞行时间的,因此测距噪声主要取决于信号幅度和背景照度。信号幅度取决于物距和物体反射率,其中,被测物体的反射率对测量噪声和重复性的影响很大。在实验中,黑色金属表面吸收大部分入射光,即反射率低,测距噪声较大,测量得到的距离在较大范围内振荡,因此选择了较大的阈值( $6\text{ cm}$ )。

#### 3.1.3 分割物体

完成平面估计后,部分属于估计平面的局内点

在目标面板外,这些点不属于面板上的点,会影响感兴趣区域的确定。因此,采用统计离群值滤波算法统计点的邻域来消除这些不需要的点,获得实际的面板区域。得到面板内点后,下一步的目标是只保留固定在面板上的物体点来粗定位物体。先利用估计的平面方程将面板内的点投影到平面上并计算其最小外接矩形。该矩形划定了可固定物体的面板区域,在此矩形基础上,沿平面法向量方向创建一个三维包围盒,通过立方体裁剪滤波器,即三个空间维度上的直通滤波器,仅保留三维包围盒内的点,即

$$\begin{cases} x_{\min} \leq x \leq x_{\max} \\ y_{\min} \leq y \leq y_{\max} \\ z_{\min} \leq z \leq z_{\max} \end{cases} \quad (10)$$

式(10)中边界坐标值  $x_{\min}$ 、 $x_{\max}$ 、 $y_{\min}$ 、 $y_{\max}$  用于限制面板的区域。为了保证目标物体点坐标的精度,裁剪滤波器的输入点云是原始点云,而不是降采样后的点云。由于每个物体对应一片紧凑的点云,因而采用欧式聚类的方法将点云聚类成多个簇,每个簇代表一个物体,同时,将点数少于一定阈值的簇剔除。通过对点云的一系列操作,最终获得了代表目标物体的点云,结果如图 5 所示。其中,红绿蓝三轴表示深度相机坐标系,红色点云表示分割出的目标,绿色点表示输入点云。图 5 还显示了估计出的平面以及裁剪立方体。

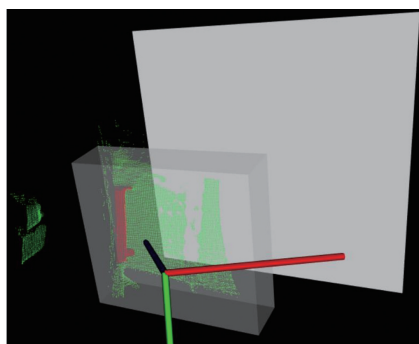


图 5 目标点云分割结果

Fig. 5 Result of target point cloud segmentation

### 3.2 基于灰度图的精确测量

CCD 相机和深度相机的外参已标定至同一世界坐标系下,故点云分割出的点可通过变换矩阵转换到灰度图像坐标,在灰度图中限定物体的区域,如图 6 所示。可以看出,红色的点比较稀疏,可以清晰地分辨出点与点之间的间隔,在局部存在缺失和空洞。因此,仅使用深度相机进行位姿测量,达到的精度有限。但这样的粗定位结果有助于后续在图像中筛选出目标直线,得到交点,进而精确求解出目标位

姿。与基于模板匹配的方法相比,所提的基于直线交点的测量方法速度更快,测量精度更高,且不需要离线训练过程,无需模板的存储空间。

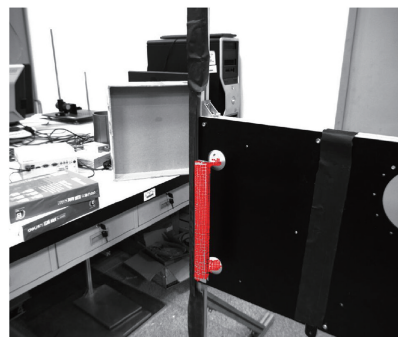


图 6 分割出的目标点云对应的灰度图像中的点

Fig. 6 Points in the gray image corresponding to the segmented target point cloud

#### 3.2.1 提取最大轮廓区域

为减小后续计算量,先计算最小外接矩形,裁剪出这部分灰度图形成小图像。小图像左上角坐标  $(0,0)$  在原图像中记为  $(c_x, c_y)$ ,在新的图像中检测 4 条目标直线并求其交点坐标,加上偏移量  $(c_x, c_y)$  即为 4 个特征点在原图像中的坐标。

为准确提取物体前表面的 4 条直线特征,首先对小图像进行如下操作。采用 Otsu 法自适应地选取阈值对图像进行二值化。以二值图像形态学中的腐蚀和膨胀操作为基础来提取包含目标边缘的局部区域。先进行腐蚀操作,侵蚀二值图像的前景像素区域,提取选择最大轮廓区域,再进行膨胀操作,扩大边界上的前景像素区域。通过二值图像形态学处理可得到物体前表面的局部区域。

#### 3.2.2 提取目标特征点

在该局部区域内采用 LSD 直线检测算法检测直线段,并通过构成直线的像素点的梯度方向来确定该直线的平均梯度方向。由于噪声和光照的影响,LSD 算法最主要的缺陷是图像中同一直线上的直线段可能出现断裂,即一条直线段被 LSD 算法提取成多条直线段。因此,将检测出的直线用极坐标形式表示,并进行直线分组。在极坐标系下将直线  $L$  表示为参数为  $(\rho, \theta)$  的方程,其中  $\rho$  为直线到原点的距离, $\theta$  为夹角,如图 7 所示。具有相似极坐标的直线被划分为一组,将所有直线划分为 4 组。目标直线的梯度方向由边缘内部指向边缘外侧,分别在 4 组中选择满足梯度方向且最长的直线段作为目标直线。通过计算目标直线相应的交点得到 4 个目标特征点。已知 4 个特征点在图像中的坐标和它们在

物体坐标系下的坐标,采用 P4P 位姿求解算法计算出物体在 CCD 相机坐标系下的位姿,最后计算出物体在世界坐标系下的位姿。

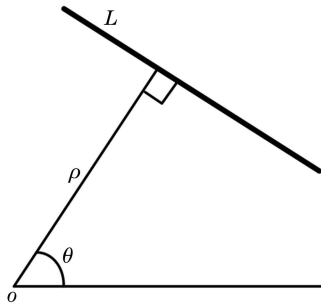


图 7 直线在极坐标系下的表示

Fig. 7 Representation of a line in a polar coordinate system

## 4 实验与结果

### 4.1 实验环境

多传感器视觉测量系统实验环境如图 8(a)所示,主要包括固定在机器人操作面板上的待测物体、两台 CCD 相机、一台深度相机和计算机等。其中: CCD 相机的分辨率为 2448 pixel×2050 pixel,镜头焦距为 8 mm。标定得到的左相机内外参矩阵分别

$$\text{为 } \begin{bmatrix} 2331.10171 & 0 & 1203.2753 \\ 0 & 2330.82152 & 1049.47257 \\ 0 & 0 & 1 \end{bmatrix} \text{ 和 } \begin{bmatrix} -0.00969 & 0.999935 & -0.00598 & 44.38937 \\ -0.99995 & -0.00971 & -0.00193 & 56.51871 \\ -0.00199 & 0.005963 & 0.99998 & 17.90749 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

畸变系数  $(k_1, k_2, p_1, p_2) = (-0.0840758, 0.0807441, 0.0002128, -0.0000821)$ 。右相机内外参矩阵分别

$$\text{为 } \begin{bmatrix} 2314.63265 & 0 & 1195.91035 \\ 0 & 2314.23571 & 991.01816 \\ 0 & 0 & 1 \end{bmatrix} \text{ 和 } \begin{bmatrix} -0.00699 & 0.999968 & -0.00383 & -45.2178 \\ -0.99998 & -0.00699 & 0.000936 & 56.70116 \\ 0.000909 & 0.003833 & 0.999992 & 16.61782 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

畸变系数为  $(k_1, k_2, p_1, p_2) = (-0.0874617, 0.0719448, 0.0003057, 0.0003541)$ 。深度相机的外参矩阵

$$\text{为 } \begin{bmatrix} 0.01009 & 1.011823 & -0.00329 & 0.702914 \\ -1.00921 & 0.008537 & -0.00676 & -47.8274 \\ -0.00659 & 0.003417 & 1.02123 & -31.8345 \\ 0 & 0 & 0 & 1 \end{bmatrix}。$$

将深度相机和左 CCD 相机的外参代入(5)式,可得到两坐标系下的坐标变换关系为

$$\mathbf{P}_c = \begin{bmatrix} 0.9881 & 0.0195 & -0.0025 \\ -0.0180 & 0.9906 & 0.0046 \\ 0.0026 & -0.0044 & 0.9792 \end{bmatrix} \cdot \mathbf{P}_d + \begin{bmatrix} 44.5464 \\ 104.0566 \\ 48.8668 \end{bmatrix}。 \quad (11)$$

为验证所提算法的有效性,以航天扶手和抽屉把手作为待测物体,用螺钉将其固定在面板上,如图 8(b)所示。机器人距离面板 0.8~1.3 m,这是一个较大的测量范围,基于单相机的方法很难满足高精度、高速度的测量要求。

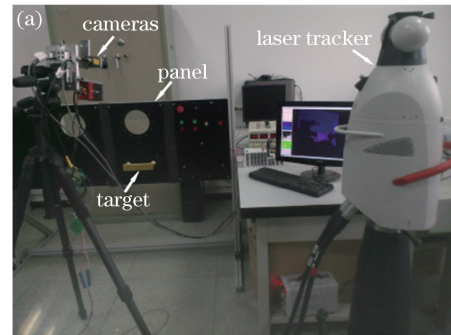


图 8 实验环境搭建。(a)实验环境;(b)待测目标

Fig. 8 Establishment of experimental environment. (a)

Experimental environment; (b) targets to be measured

### 4.2 测量特征提取结果

采用所提方法提取直线特征并求其交点,结果如图 9 所示。

### 4.3 精度评估

为了评估所提方法的测量精度,将视觉系统在距离面板 0.8~1.3 m 的范围内移动,变换不同的视点位置和观察角度。对每个物体在 80 个不同位姿下采集数据,深度相机采集点云,CCD 相机采集 8 位单通道灰度图像。在每个位姿下重复采集 10 次,对于每个物体获取了 800 个点云及其对应的灰度图,将它们作为精度评估数据集。利用激光跟踪仪

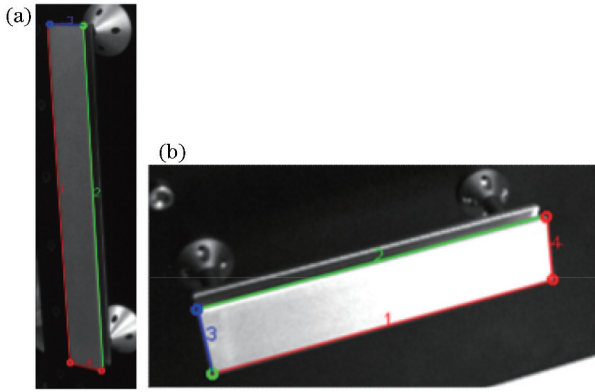


图9 直线特征提取结果。(a)航天扶手;(b)抽屉把手  
Fig. 9 Results of linear feature extraction. (a) Space handrail; (b) drawer handle

获取物体坐标系相对于世界坐标系的真实位姿(平移向量  $\mathbf{T}_{\text{truth}}$  和旋转角度  $\mathbf{A}_{\text{truth}}$ ),然后采用所提方法对目标物体进行位姿测量。为了评估测量的位置误差和姿态误差,以绝对误差作为评价指标,即测量值与真实值之差的绝对值。位置误差的绝对值和姿态误差的绝对值分别为

$$e_{\text{trans}} = | \mathbf{T}_{\text{our}} - \mathbf{T}_{\text{truth}} |, \quad (12)$$

$$e_{\text{ang}} = | \mathbf{A}_{\text{our}} - \mathbf{A}_{\text{truth}} |, \quad (13)$$

表1 测量精度

Table 1 Measurement accuracy

Object	Deviation	$E_x/\text{mm}$	$E_y/\text{mm}$	$E_z/\text{mm}$	$E_\alpha/(\text{^\circ})$	$E_\beta/(\text{^\circ})$	$E_\gamma/(\text{^\circ})$
Space handrail	Mean absolute deviation	0.31	1.20	0.98	0.97	0.41	0.32
	Maximum absolute deviation	0.94	2.76	2.73	1.88	1.56	1.34
Drawer handle	Mean absolute deviation	1.26	0.68	1.35	0.51	0.49	0.37
	Maximum absolute deviation	2.92	1.19	2.51	1.72	1.24	1.42

#### 4.4 效率评估

除测量精度外,衡量算法的另一重要性能指标是算法效率。将所提方法与 Ulrich 的经典模板匹配方法<sup>[5]</sup>进行对比,所有实验均在便携式计算机(Intel Core i7-7700HQ CPU 2.80 GHz,8.00 GB 内存)上进行。所提方法在 Visual Studio 2013 环境 release 模式下运行并记录时间,Ulrich 方法在机器视觉商业软件 Halcon 11 下记录运行时间。表 2 所示为运行时间的对比,可以看出,所提方法具有较高的效率。这主要是因为所提方法通过点云分割规避了复杂背景,节省了在大面积背景上匹配的时间。

表2 运行时间对比

Table 2 Running time comparison

Methods	Average time /ms
Ulrich's method	2120.43
Proposed method	148.14

式中: $\mathbf{T}_{\text{our}}$ 为平移向量的测量值; $\mathbf{A}_{\text{our}}$ 为旋转角度的测量值。

为了评估算法的可重复性和稳健性,对同一位姿下重复采集的 10 次数据,计算其测量位姿的平均值作为该位姿下的测量值,然后计算 80 个不同位姿下测量结果的平均绝对误差和最大绝对误差,统计结果如表 1 所示,其中  $E_x$ 、 $E_y$ 、 $E_z$  分别为  $x$ 、 $y$ 、 $z$  轴上的位置误差, $E_\alpha$ 、 $E_\beta$ 、 $E_\gamma$  分别为绕  $x$ 、 $y$ 、 $z$  轴的旋转角误差。针对航天扶手的误差数据进行分析,可以看出: $y$  轴和  $z$  轴上的位置误差较大,绕  $x$  轴的旋转角误差较大。结合图 3(a)可知,扶手上  $ad$ 、 $bc$  两条边的长度较短,在有些视角下成像不够清晰,无法准确定位出直线  $ad$ ,直线提取存在一定偏差,求取的交点出现偏差,位姿测量误差较大。但总体上平移向量的平均绝对误差小于 2 mm,最大绝对误差小于 3 mm,旋转角平均绝对误差在  $1^\circ$  以内,最大绝对误差在  $2^\circ$  以内。实验结果表明所提方法具有较高的测量精度,同时也证明了所提方法的有效性和准确性。Ulrich 的经典模板匹配方法<sup>[5]</sup>在成像视图退化情况下的稳健性较差,无法正常完成测量,故没有进行测量精度的对比。

另一方面,所提方法没有大量模板的匹配过程,节约了大量时间开销。

## 5 结 论

提出了一种基于多传感器的三维目标位姿测量方法,充分发挥深度相机和 CCD 相机各自的优势,实现了对具有矩形特征三维目标的快速、高精度位姿测量。解决了经典模板匹配算法在成像视图退化情况下无法正确测量的问题,大大提高了测量效率。实验结果证明了该方法的有效性。所提方法借助物体与面板之间位置关系的先验知识,完成了目标物体的粗定位。不利用先验知识,基于多传感器快速准确地完成三维目标位姿测量将是后续的研究方向。

## 参 考 文 献

- [1] Bay H, Ess A, Tuytelaars T, *et al.* Speeded-up



- robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [2] Hinterstoisser S, Benhimane S, Navab N. N<sub>3</sub>M: natural 3D markers for real-time object detection and pose estimation [C]. IEEE 11th International Conference on Computer Vision, 2007: 1-7.
- [3] Lepetit V, Fua P. Keypoint recognition using randomized trees[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(9): 1465-1479.
- [4] Ulrich M, Wiedemann C, Steger C. CAD-based recognition of 3D objects in monocular images[C]. IEEE International Conference on Robotics and Automation, 2009: 1191-1198.
- [5] Ulrich M, Wiedemann C, Steger C. Combining scale-space and similarity-based aspect graphs for fast 3D object recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 1902-1914.
- [6] Hinterstoisser S, Holzer S, Cagniart C, *et al.* Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes [C]. International Conference on Computer Vision, 2011: 858-865.
- [7] Hinterstoisser S, Cagniart C, Ilic S, *et al.* Gradient response maps for real-time detection of textureless objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(5): 876-888.
- [8] Lepetit V, Moreno-Noguer F, Fua P. EPnP: an accurate  $O(n)$  solution to the PnP problem [J]. International Journal of Computer Vision, 2009, 81(2): 155-166.
- [9] Zhang S J, Cao X B, Zhang F, *et al.* Monocular vision-based iterative pose estimation algorithm from corresponding feature points [J]. Science China Information Sciences, 2010, 53(8): 1682-1696.
- [10] Zhu F, Yu F S, Wu Y M, *et al.* Analysis of calibration precision of camera attitude angles [J]. Acta Optica Sinica, 2018, 38(11): 1115005.  
朱帆, 于芳苏, 吴易明, 等. P4P法相机姿态标定精度分析[J]. 光学学报, 2018, 38(11): 1115005.
- [11] von Gioi R G, Jakubowicz J, Morel J M, *et al.* LSD: a line segment detector [J]. Image Processing on Line, 2012, 2: 35-55.
- [12] Zhang H J, Fang Z J, Yang G L. RGB-D visual odometry in dynamic environments using line features [J]. Robot, 2018, 40(5): 1-8.  
张慧娟, 方灶军, 杨桂林. 动态环境下基于线特征的RGB-D视觉里程计[J]. 机器人, 2018, 40(5): 1-8.
- [13] Zhang Z. A flexible new technique for camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [14] Song J H, Ren Y J, Yang S R, *et al.* Extrinsic parameter calibration method based on substitutable target sphere for vision sensors [J]. Acta Optica Sinica, 2017, 37(9): 0915003.  
宋佳慧, 任永杰, 杨守瑞, 等. 基于合作靶球的视觉传感器外参标定方法[J]. 光学学报, 2017, 37(9): 0915003.
- [15] Li R Z, Liu Y Y, Yang M, *et al.* Three-dimensional point cloud segmentation algorithm based on improved region growing [J]. Laser & Optoelectronics Progress, 2018, 55(5): 051502.  
李仁忠, 刘阳阳, 杨曼, 等. 基于改进的区域生长三维点云分割[J]. 激光与光电子学进展, 2018, 55(5): 051502.
- [16] Rusu R B, Cousins S. 3D is here: point cloud library (PCL) [C]. IEEE International Conference on Robotics and Automation, 2011: 1-4.
- [17] Choi S, Kim T, Yu W. Performance evaluation of RANSAC family [C]. Proceedings of the British Machine Vision Conference, 2009: 355.
- [18] Jin J J, Lu W L, Guo X T, *et al.* Position registration method of simultaneous phase-shifting interferograms based on SURF and RANSAC algorithms[J]. Acta Optica Sinica, 2017, 37(10): 1012002.  
靳京京, 卢文龙, 郭小庭, 等. 基于SURF和RANSAC算法的同步相移干涉图位置配准方法[J]. 光学学报, 2017, 37(10): 1012002.
- [19] Zeineldin R A, El-Fishawy N A. Fast and accurate ground plane detection for the visually impaired from 3D organized point clouds[C]. 2016 SAI Computing Conference, 2016: 373-379.