

具有远近视距的两目视觉系统标定技术研究

石世锋, 叶南*, 张丽艳

南京航空航天大学机电学院, 江苏 南京 210016

摘要 工业机械臂在多工位作业时通常需要具有远近视距的双目视觉系统,用于全局目标定位和局域精确测量。由于具有远近视距的两目相机缺乏公共视场,无法由传统双目立体视觉标定内外参数,提出了一种基于点阵编码元和多视图几何的标定技术。首先为具有远近视距的两目系统制作与之视场大小匹配的两种尺寸的点阵编码元标定板,提出了一种基于灰度梯度的亚像素边缘定位方法以提高椭圆中心定位精度,并改进了点阵编码元的解码算法用于图像特征的稳健匹配。再利用不同视角下编码元的高精度中心像素坐标和图像间编码元的对应关系,通过多视图几何解算出两目相机的内参数和相对各自标定板的转换矩阵。然后由转换矩阵构造 $AX=ZB$ 手眼标定方程形式,求解方程式中的 X 即可得到两目相机间的外参数。最后分析了两类优化方法对标定误差的抑制情况。实验结果表明,经优化后标定精度得到提高,角度误差为 0.05° ,位置误差为 0.36 mm 。

关键词 机器视觉; 内外参标定; 多视图几何; 亚像素边缘; 编码元; 手眼标定方程式

中图分类号 TN948.41; TP391

文献标志码 A

doi: 10.3788/AOS202141.2415001

Calibration of Two-Camera Vision System with Far and Near Sight Distance

Shi Shifeng, Ye Nan*, Zhang Liyan

*College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics and Astronautics,
Nanjing, Jiangsu 210016, China*

Abstract A two-camera vision system with far and near sight distance for global target positioning and accurate local measurement is usually required when an industrial robotic arm works in multiple stations. Lacking a common field of view, the intrinsic and extrinsic parameters of the vision system cannot be calibrated using traditional methods; therefore, a new calibration technology based on point-array coded targets and multi-view geometry is proposed. First, two sizes of calibration panels with point-array coded targets matching the field of view size are prepared for the two-camera vision system with far and near sight distance. A sub-pixel edge detection method based on gray gradient is proposed to improve the positioning accuracy of the ellipse centers. Then, a decoding algorithm of the point-array coded target is improved for robust image feature matching. Based on this, using high-precision central pixel coordinates of point-array coded targets under different perspectives and the corresponding relationship of targets between different images, the intrinsic parameters of two cameras and the transformation matrices between cameras and calibration panels can be calculated using multi-view geometry technology. Then, a hand-eye calibration equation, $AX=ZB$, is constructed from the transformation matrix, and the extrinsic parameter X between two cameras can be obtained by solving the equation. Finally, the suppression of calibration error by two kinds of optimization methods is analyzed. Experimental results demonstrate that the calibration accuracy is improved after optimization. The angle error is reduced to 0.05° , and the position error is reduced to 0.36 mm .

Key words machine vision; intrinsic and extrinsic parameter calibration; multi-view geometry; sub-pixel edge; coded target; hand-eye calibration equation

OCIS codes 150.1488; 150.0155; 100.3010; 120.4640

收稿日期: 2021-05-27; **修回日期:** 2021-06-23; **录用日期:** 2021-06-28

基金项目: 国家自然科学基金面上项目(52075260)、中央高校基本科研业务费(3082020NP2020305)

通信作者: *yen@nuaa.edu.cn

1 引言

在航空航天制造领域,工业机械臂被广泛应用于零件的抓取、旋拧等装配任务,或空间大型结构件的在轨组装等^[1-2]。受机械臂定位精度影响,实际应用中常借助各类传感器辅助机械臂完成装配、维护工作,其中视觉传感器以精度高、成本低等特点受到学者们的广泛关注。将单个相机固联于机械臂末端,搭建成单目"Eye-in-Hand"视觉系统,此类系统能灵活高效地完成自动对准过程,但只适用于近距离的装配工作。面对大尺寸组零件,机械臂需要从较远距离开始工作,工作环境横向空间跨度大,纵向距离长,单个相机在某些距离段会出现失焦、离焦现象,无法在整个过程中都清晰成像,导致位姿测量精度降低,甚至装配任务失败。

为使机械臂视觉系统在远近距离都能够捕获到高清晰度图像,部分研究中使用变焦相机采集图像,但变焦相机标定难度大、精度低、实现复杂度高,极大程度限制了其在工业现场的应用^[3-4]。解决该问题的另一种有效途径是使用不同视距相机构成的两目视觉系统,一台相机作为全局相机进行远距离观察定位,另一个相机用于近距离的局域测量^[5]。与传统双目立体视觉系统不同,两目视觉系统视场范围大小不一致,且缺少公共视场,无法使用传统标定方法校准系统^[6],这给标定工作带来挑战。

针对无公共视场视觉系统的外参标定,现存的方法主要包括借助高精度移动装置^[7]、构建大型标定物^[8-9]和基于镜面成像的标定方法^[10]。借助移动装置的标定方法将相机系统或平面标定板放置于移动装置上,如二轴精密转台,通过装置的旋转平移,标定板覆盖所有相机视场。显然,高精度装置增加了系统成本,且移动范围受到严重限制。Liu 等^[8]在一维标定棒上加工多个共线的特征点,由特征点的共线性求解旋转矩阵,由特征点之间的距离解算平移关系,最后进行非线性优化。Xia 等^[9]搭建大尺寸标定板并在标定板上布置环状编码元,由摄影测量技术得到编码元的三维信息,然后每个相机对各自视场内的编码元进行拍摄,最终通过配准编码元的三维信息得到外参。上述两种方法的本质都是通过构建足够覆盖全部相机视场的标定物把各个相机联系起来,标定物的尺寸要根据视觉系统的视场范围决定,缺少灵活性和通用性,当视场范围增大到一定程度,构建尺寸与之匹配的标定物甚至是不实际的。基于镜面成像的标定方法借助平面镜对标定

板进行虚拟成像,使得所有相机都能采集到标定板图像,其中部分相机采集到的是真实的标定板图像,其余相机采集到的是平面镜中的标定板图像。在标定过程中,镜面的摆放位置尤为关键,摆放位置必须满足一定的物理条件,这对操作人员提出了较高要求。此外,标定结果对镜面位置非常敏感,因此难以获得稳定、高精度的标定结果。王刚等^[11]将外参标定问题转换为机器人领域中的手眼标定问题,将两块棋盘格标定板分别用于两台相机的图像采集,通过多次移动视觉系统并同时拍摄各自的标定板图像,从序列图像中估计相机轨迹后建立手眼标定模型 $CX = XD$,求解该模型得到两台相机的外参。但该过程没有对模型求解结果进行优化,位置精度只能达到厘米级。本文的两目视觉系统视距不同,两台相机所观察的视场范围大小不一,上述方法均不能很好地适用于本系统。

针对具有远近视距且视场大小不一致的两目系统,本文制备两种尺寸与相机视场相匹配的点阵编码元标定板,在此基础上提出了一种多视图几何标定技术。提出了一种基于灰度梯度的亚像素边缘中心定位方法,并对点阵编码元的解码算法进行了改进。通过获取不同图像上编码元的高精度中心像素坐标并建立图像间编码元的对应关系,由多视图几何技术解算两目系统的内参数和相对于标定板的转换矩阵。利用转换矩阵建立形如 $AX = ZB$ 手眼标定方程式,解算 X 即可得到两目系统的外参矩阵。使用两类优化方法抑制该方程的求解误差,最终得到精确的标定结果。

2 技术路线

远近视距两目系统包括全局相机和局域相机,两台相机固联工作,如图 1 所示。标定工作包括两个关键步骤,首先通过标定板获得两台相机的内参数和相对标定板的位姿关系(即位置和姿态),然后对两相机之间的位姿(即外参数)进行解算。

在标定之前,根据相机视场,制备两块布置有若干数量的点阵编码元的标定板,其中大尺寸标定板适配于全局相机,小尺寸标定板适配于局域相机。通过整体移动视觉系统改变相机拍摄视角,两台相机在不同视角下分别对各自标定板进行拍摄,这一过程中两台相机同时工作,同时采集图像。需要说明的是,在一个拍摄视角下,相机无需采集到完整的标定板图像,仅需包含至少一个完整的编码元信息即可。

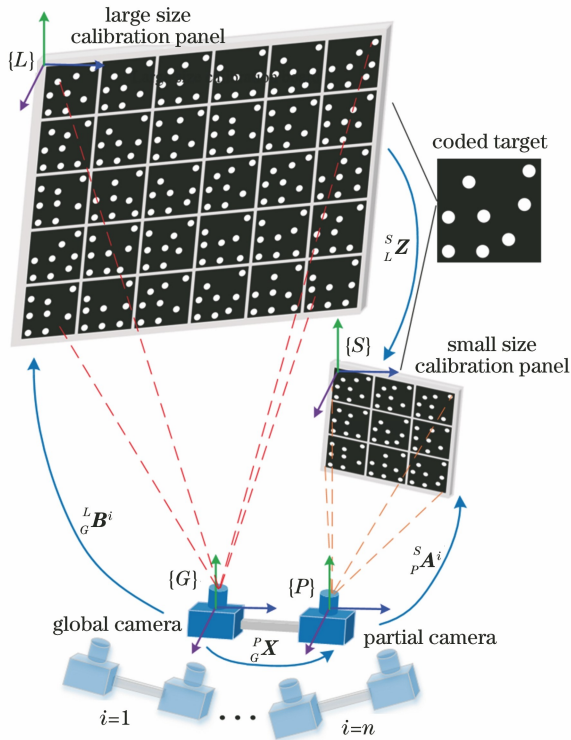


图 1 系统结构

Fig. 1 System structure

在多个视角下采集完图像后,由多视图几何技术^[12]得到相机内参数和每一个拍摄视角下相机相对于标定板的位姿关系。利用获得的位姿数据和图像对关系,建立方程式:

$${}^S_P\mathbf{A}^i \cdot {}^P_G\mathbf{X} = {}^L_S\mathbf{Z} \cdot {}^L_G\mathbf{B}^i, \quad (1)$$

式中: ${}^S_P\mathbf{A}^i$ 表示在第*i*($i=1,2,\dots,n$)个拍摄视角下局域相机坐标系{P}相对于小尺寸标定板坐标系{S}的变换关系; ${}^L_G\mathbf{B}^i$ 表示在第*i*个拍摄视角下全局相机坐标系{G}相对于大尺寸标定板坐标系{L}的变换关系; ${}^P_G\mathbf{X}$ 表示全局相机坐标系{G}相对于局域相机坐标系{P}的变换关系,即视觉系统的外参数; ${}^L_S\mathbf{Z}$ 表示大尺寸标定板坐标系{L}相对于小尺寸标定板坐标系{S}的变换关系。(1)式为机器人手眼标定领域的经典方程,可简写为

$$\mathbf{A}^i \mathbf{X} = \mathbf{Z} \mathbf{B}^i. \quad (2)$$

将(2)式齐次矩阵等式展开,可得

$$\begin{cases} \begin{bmatrix} \mathbf{R}_A^i & \mathbf{t}_A^i \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_X & \mathbf{t}_X \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_Z & \mathbf{t}_Z \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_B^i & \mathbf{t}_B^i \\ \mathbf{0}^T & 1 \end{bmatrix} \\ \mathbf{R}_A^i \mathbf{R}_X = \mathbf{R}_Z \mathbf{R}_B^i \\ \mathbf{R}_A^i \mathbf{t}_X + \mathbf{t}_A^i = \mathbf{R}_Z \mathbf{t}_B^i + \mathbf{t}_Z \end{cases}, \quad (3)$$

式中: \mathbf{R} 表示 3×3 旋转矩阵; \mathbf{t} 表示 3×1 平移向量。求解(3)式中的 $\mathbf{R}_X, \mathbf{t}_X$ 即得到两台相机的外

参数。

3 内参标定和相机位姿估计

相机采集到标定板图像后,椭圆中心定位是关键第一步,得到中心点集后需进行解码与匹配,在此基础上由多视图几何技术计算两目系统的内参数和相机相对于标定板的位姿关系。

3.1 高精度椭圆中心定位

对于图像中的椭圆,传统的中心定位方法一般由Canny算子或二值化操作得到整像素的轮廓边缘信息,在此基础上使用最小二乘拟合椭圆方程并计算中心。整像素点无法精确表示椭圆的真实边缘,所拟合的中心点无法满足高精度的测量要求,因此提出了一种基于灰度梯度的亚像素边缘中心定位方法。

文献[13]指出图像边缘的灰度分布可用三次解析曲线表示,真正边缘处的灰度导数值最大。由于灰度的导数可用灰度梯度表示,因此对图像边缘处的灰度梯度进行二次多项式拟合,取其极值点判定为真正的边缘点。基于灰度梯度的亚像素边缘中心定位法主要包括整像素点的灰度梯度计算、沿灰度梯度方向采样和灰度梯度幅值曲线拟合3个步骤。

1) 使用Canny算子^[14]进行边缘检测,确定整像素边缘点;使用Sobel算子^[15]计算图像中每个整像素点沿*U*和*V*方向的灰度梯度(幅值和方向),如图2(a)所示。

2) 设某整像素点为*P*(*u, v*),沿其灰度梯度方向分别在*P*点两侧以一定步长*L*进行采样。由于本成像系统可控制椭圆边缘过渡区域为5~7个像素,因此取步长*L*=1.5像素,共采样4个点*A, B, C, D*,如图2(b)所示。采样点*A, B, C, D*一般不在整像素位置上,无法直接获得非整像素点的梯度幅值,需采用插值的方法来确定其梯度幅值。使用双线性法进行梯度幅值插值。

3) 对整像素点*P*和四个采样点*A, B, C, D*的梯度幅值进行二次多项式拟合,取其极值点作为亚像素精度的边缘位置,如图2(c)所示。最后对得到的亚像素边缘使用最小二乘拟合法求出椭圆中心。

图3展示了整像素边缘与所提方法确定的亚像素边缘的对比效果,明显可以看出,亚像素边缘比整像素边缘更为平滑光滑,更加接近真实的椭圆轮廓。

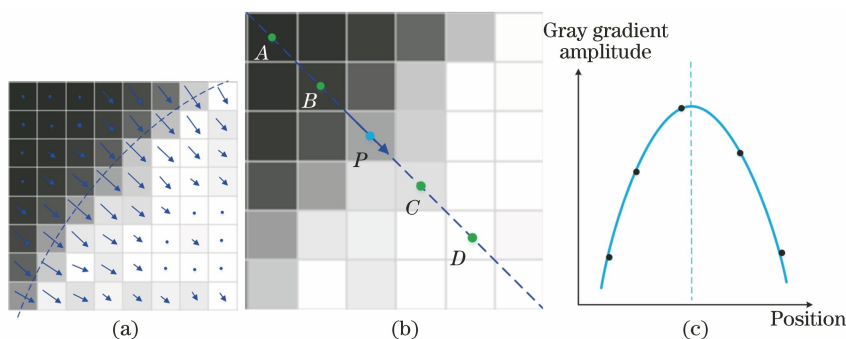


图 2 确定亚像素边缘的步骤。(a)图像灰度梯度计算;(b)沿灰度梯度方向采样;(c)灰度梯度曲线拟合

Fig. 2 Three steps for determining the sub-pixel edge. (a) Calculating image gray gradient; (b) sampling along the gray gradient direction; (c) gray gradient amplitude curve fitting

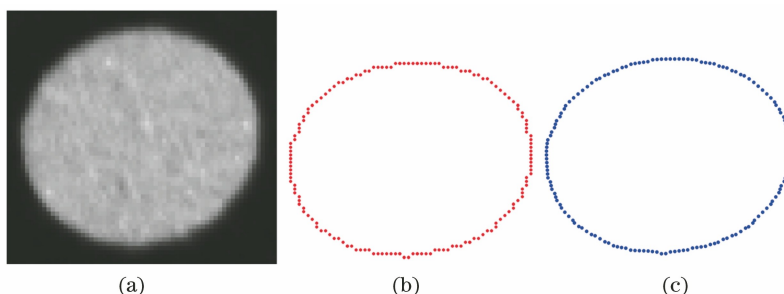


图 3 边缘对比。(a)椭圆图像;(b)整像素边缘;(c)亚像素边缘

Fig. 3 Edge comparison. (a) Ellipse image; (b) pixel edge; (c) sub-pixel edge

3.2 点阵编码元解码与匹配

使用的点阵编码元由 8 个圆点组成,包括 5 个固定点和 3 个浮动点。图 4(a)为编码元模板, A、B、C、E、F 五个圆点为固定点,相对位置关系已知, A、B、C 三点在一条直线上,直线 EF 与直线 AB 交于点 P。A、P、B、C 四点的交比是确定值,为

1.5276。从编号为 1~25 的点中任取三个作为浮动点,这三点确定了编码号,为提高解码的准确性,这三个浮动点不能处于相邻位置。图 4(b)、(c)展示了编码元的实际成像情况,浮动点编号为 10, 11, 25,则该编码元的编码号为 $101125 = 10 \times 10^4 + 11 \times 10^2 + 25$ 。

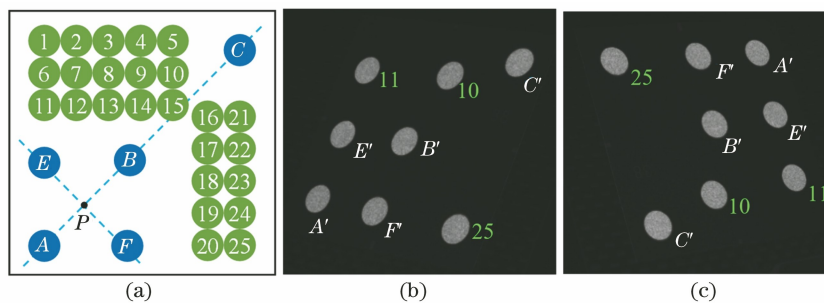


图 4 点阵式编码元。(a)点阵编码元模板;(b)(c)不同视角下的编码元图像

Fig. 4 Point-array coded target. (a) Point-array coded target template; (b)(c) images of point-array coded target from different perspectives

点阵编码元的解码过程如下。

1) 图像中所有椭圆经中心定位后得到点集 $P = \{p_i, i = 1, 2, \dots, n\}$,并以 KD-Tree 的形式存储。

2) 使用 K 邻近算法(KNN)对点集 P 中任意一点 p_i 进行搜索,获得距离 p_i 最近的 K 个点(包括其自身点),其中 $K = 8$,构成一个八点点集 $Q =$

$\{q_i, i = 1, 2, \dots, n\}$ 。

3) 遍历点集 Q,从中取三点 q_m, q_u, q_k ,构造两条向量 $n_1 = \frac{q_m - q_k}{|q_m - q_k|}, n_2 = \frac{q_u - q_k}{|q_u - q_k|}$,若满足(4)式,则认为这三点共线。受编码元加工误差和椭圆中心定位误差的影响,任意三点不可能完全在一条

直线上, 给定阈值 $\lambda = 0.99$ 。分别计算这三点间的距离, 找到距离最大的两点, 则另一点就暂定为 B' 点, 根据距离关系, A' 和 C' 两点也就暂时确定了。若遍历完点集 Q 还未找到满足条件的三点, 则说明该八点不属于同一编码元, 返回步骤 2)。

$$|n_1 \cdot n_2| \geq \lambda. \quad (4)$$

4) 遍历点集 Q 中的剩余五点, 在直线 $A'B'C'$ 的逆时针方向寻找一点 q_x , 顺时针方向寻找一点 q_y , 两点连线与直线 $A'B'C'$ 交于点 P' , 计算 A' 、 P' 、 B' 、 C' 四点的交比, 若满足 (5) 式, 则确定 q_x 为点 E' , q_y 为点 F' 。在实际解码过程中, 受编码元加工误差和图像噪声等影响, 计算出的交比与模板的交比 (1.5276) 存在些许偏差, 给定阈值 $\eta = 0.05$ 。若遍历完点集 Q 还未找到满足条件的 E' 、 F' 两点, 则返回步骤 3)。

$$\begin{cases} \text{cr}(A', P', B', C') = \frac{\overline{A'B'} \cdot \overline{P'C'}}{\overline{P'B'} \cdot \overline{A'C'}} \\ |\text{cr}(A', P', B', C') - 1.5276| \leq \eta \end{cases}, \quad (5)$$

式中: $\text{cr}(\cdot)$ 为交比。

5) 图像中五个固定点 A' 、 B' 、 C' 、 E' 、 F' 已知, 且与模板中的 A 、 B 、 C 、 E 、 F 一一对应, 通过单应性计算图像点与模板点的变换关系 H 。用此变换关系对点集 Q 中剩余三点进行投影变换, 并与模板中的 25 个浮动点作比较, 最终确定编码号。

6) 重复步骤 2)~5), 直至点集 P 中所有点完成解码。

上述算法的伪代码如图 5 所示。

```

Algorithm 1 Identify Coded Targets
Input: Image {I}
Output: Coded Targets {C_j}
1: {P} ← CenterPosition({I})
2: for each p_i from {P} do
3:   {Q} ← KNNSearch(p_i, 8)
4:   for each q_m, q_u, q_k from {Q} do
5:     if FitLine(q_m, q_u, q_k) then
6:       for each q_x, q_y from {Q} do
7:         cr ← CalCrossRatio(q_m, q_u, q_k, q_x, q_y)
8:         if |1.5276 - cr| ≤ 0.05 then
9:           H ← FindHomograph(q_m, q_u, q_k, q_x, q_y)
10:          {q_1, q_2, q_3} ← PerspectiveTransform(H)
11:          C ← GetCodedTargetInfo({q_1, q_2, q_3})
12:          {C_j} ← C
13:        end if
14:      end for
15:    end if
16:  end for
17: end for
    
```

图 5 算法伪代码

Fig. 5 Algorithm pseudo code

图 6 展示了不同视角下两幅图像间的编码元匹配关系, 解码完成后, 每个编码号都对应对着八个圆点中心坐标, 这八个中心在解码过程中已有序排列。因此在不同图像间, 只要寻找到相同的编码号, 其对应的八个圆点就一一匹配好了。

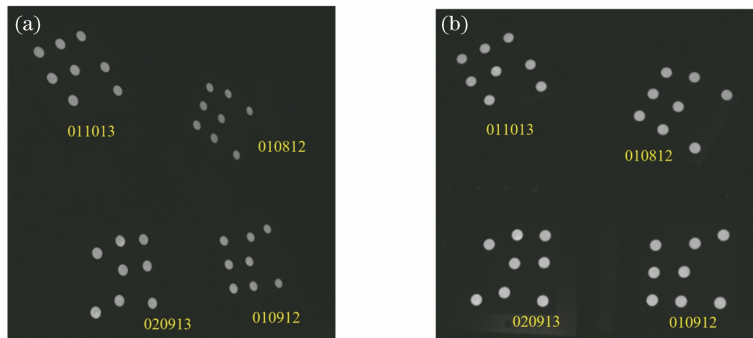


图 6 点阵编码元匹配。(a)(b)不同视角下的点阵编码元图像

Fig. 6 Point-array coded target matching. (a)(b) Images of point-array coded targets in different perspectives

3.3 多视图几何技术

多视图几何技术的流程一般包括特征提取、特征匹配、structure from motion(SfM)、总体优化 4 个步骤^[16-17]。特征提取和匹配在本文中分别指椭圆中心定位和编码元解码匹配, 由 3.1 节和 3.2 节完成。在此基础上, 采用增量式 SfM^[18] 恢复标定板的三维信息, 具体过程如图 7 所示。首先在图像序列中选择最优的两张图像, 由两视图几何技术计算初始结构信息; 然后不断新增图像, 利用新增图像中的

二维点信息和初始结构中的三维点信息, 由 perspective-n-point(PnP)解算相机位姿, 使用三角法重建结构信息并进行光束平差优化; 最后对重建出来的结构进行总体优化。在光束平差过程中, 相机内参也参与优化, 因此最终能够获得内参的精确值。

本文中, 两目视觉系统在不同视角下同时采集各自的标定板图像, 在使用增量式 SfM 重建标定板的过程中得到相机内参数, 并获得不同视角下局域

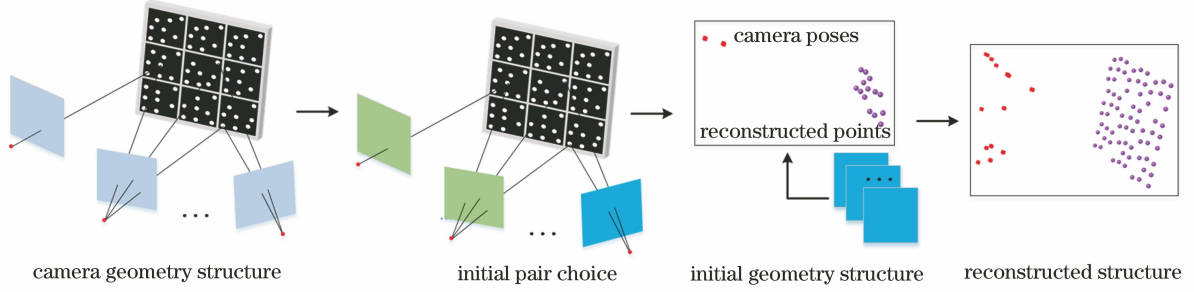


图 7 增量式 SfM 重建

Fig. 7 Incremental SfM reconstruction

相机坐标系相对于小尺寸标定板坐标系的位姿关系即 \mathbf{R}_A^i 和 \mathbf{t}_A^i , 及全局相机坐标系相对于大尺寸标定板坐标系的位姿关系即 \mathbf{R}_B^i 和 \mathbf{t}_B^i 。

4 外参估计

获得 $\mathbf{R}_A^i, \mathbf{t}_A^i, \mathbf{R}_B^i, \mathbf{t}_B^i$ 后可建立(3)式, 两目视觉系统的外参标定实质是求解(3)式中的 $\mathbf{R}_X, \mathbf{t}_X$ 。以矩阵直积形式表示旋转量, 同时求解旋转和平移, 即 Li 等^[19]提出的方法, 多数研究将此结果作为最终值, 而本文在此基础上将该结果作为初值进行进一步迭代优化。

对于 3×3 的旋转矩阵 \mathbf{R}_A^i 和 \mathbf{R}_B^i , 有

$$\text{vec}(\mathbf{R}_A^i) = [a_{11}, a_{12}, a_{13}, \dots, a_{33}]^T. \quad (6)$$

(6)式表示将矩阵 \mathbf{R}_A^i 以行优先形式转换为一维向量。

$$\mathbf{R}_A^i \otimes \mathbf{R}_B^i = [a_{mm} \mathbf{R}_B^i] = \begin{pmatrix} a_{11} \mathbf{R}_B^i & \dots & a_{13} \mathbf{R}_B^i \\ \vdots & & \vdots \\ a_{31} \mathbf{R}_B^i & \dots & a_{33} \mathbf{R}_B^i \end{pmatrix}_{9 \times 9}. \quad (7)$$

借助(6)、(7)式, 将(3)式扩展为

$$[\mathbf{R}_A^i \otimes \mathbf{I}_3 \quad -\mathbf{I}_3 \otimes \mathbf{R}_B^{iT}]_{9 \times 18} \begin{bmatrix} \text{vec}(\mathbf{R}_X) \\ \text{vec}(\mathbf{R}_Z) \end{bmatrix}_{18 \times 1} = \mathbf{0}_{9 \times 1}, \quad (8)$$

$$[\mathbf{I}_3 \otimes \mathbf{t}_B^{iT} \quad -\mathbf{R}_A^i \quad \mathbf{I}_3]_{3 \times 15} \begin{bmatrix} \text{vec}(\mathbf{R}_Z) \\ \mathbf{t}_X \\ \mathbf{t}_Z \end{bmatrix}_{15 \times 1} = \mathbf{t}_A^i, \quad (9)$$

式中: \mathbf{I}_3 表示 3×3 的单位矩阵。合并(8)、(9)式, 可得

$$\begin{bmatrix} \mathbf{R}_A^i \otimes \mathbf{I}_3 & -\mathbf{I}_3 \otimes \mathbf{R}_B^{iT} & \mathbf{0}_{9 \times 3} & \mathbf{0}_{9 \times 3} \\ \mathbf{0}_{3 \times 9} & \mathbf{I}_3 \otimes \mathbf{t}_B^{iT} & -\mathbf{R}_A^i & \mathbf{I}_3 \end{bmatrix}_{12 \times 24} \times$$

$$\begin{bmatrix} \text{vec}(\mathbf{R}_X) \\ \text{vec}(\mathbf{R}_Z) \\ \mathbf{t}_X \\ \mathbf{t}_Z \end{bmatrix}_{24 \times 1} = \begin{bmatrix} \mathbf{0}_{9 \times 1} \\ \mathbf{t}_A^i \end{bmatrix}_{12 \times 1}. \quad (10)$$

由(10)式可知, 两目视觉系统在一个视角下所得到的位姿关系可建立 12 个方程, 欲求解 $\mathbf{R}_X, \mathbf{t}_X, \mathbf{R}_Z, \mathbf{t}_Z$ 共 24 个未知量, 系统至少需要在两个不同视角采集图像, 且这两个视角不能是纯旋转或纯平移的。通过奇异值分解(SVD)求解系统外参数, 对旋转部分进行正交化, 确定最终结果。

将上述结果作为初值, 研究了二类优化方法^[20], 第一类为基于最小化方程求解误差, 第二类优化方法为基于最小化重投影误差。

第一类优化方法的损失函数定义为

$$c_1 = \sum_{i=1}^n \|A_i \mathbf{X} - \mathbf{Z} \mathbf{B}_i\|_F^2. \quad (11)$$

把旋转和平移代入(11)式中, 损失函数可转换为

$$\text{argmin}_{\mathbf{R}_X, \mathbf{t}_X, \mathbf{R}_Z, \mathbf{t}_Z} \sum_{i=1}^n \|A_i \begin{bmatrix} \mathbf{R}_X & \mathbf{t}_X \\ \mathbf{0}^T & \mathbf{1} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_Z & \mathbf{t}_Z \\ \mathbf{0}^T & \mathbf{1} \end{bmatrix} \mathbf{B}_i\|_F^2. \quad (12)$$

第二类优化方法是基于最小化重投影误差。对于小尺寸标定板上的三维点 ${}^S \mathbf{P}_j$, 经过 ${}^P \mathbf{A}^i = {}^P_G \mathbf{X} ({}^L_G \mathbf{B}^i)^{-1} {}^L_S \mathbf{Z}^{-1} = \mathbf{X} (\mathbf{B}^i)^{-1} \mathbf{Z}^{-1}$ 变换到局域相机坐标系下, 再由局域相机内参数变换到图像坐标系下, 得到二维点 $\widetilde{\mathbf{p}}_{ij}$, 与其相对应的原图像中的二维点为 \mathbf{p}_{ij} , 因此损失函数可定义为

$$c_2 = \sum_{i=1}^n \sum_{j=1}^{n'} \| \mathbf{p}_{ij} - \widetilde{\mathbf{p}}_{ij} \|^2 = \sum_{i=1}^n \sum_{j=1}^{n'} \| \mathbf{p}_{ij} - \text{Proj}(\mathbf{K}_p, [\mathbf{X} (\mathbf{B}^i)^{-1} \mathbf{Z}^{-1}]_{3 \times 4} {}^S \mathbf{P}_j) \|^2, \quad (13)$$

式中： K_p 表示局域相机的内参数。用 \hat{B}^i 表示 $(B^i)^{-1}$ ， \hat{Z} 表示 Z^{-1} ，将旋转和平移代入(13)式中，可得

$$\text{Proj}\left(K, \begin{bmatrix} R_X & t_X \\ \mathbf{0}^T & 1 \end{bmatrix} \hat{B}^i \begin{bmatrix} R_Z & t_Z \\ \mathbf{0}^T & 1 \end{bmatrix}_{3 \times 4} sP_j\right) \parallel^2. \quad (14)$$

5 实验验证

使用两台型号为 MER-502-79U3M，分辨率为 2448×2048 的大恒工业相机，其中一台搭配 6 mm 镜头作为全局相机使用，视场大小约为 3 m×3 m×2 m，另一台搭配 8 mm 镜头作为局域相机使用，视场大小约为 0.5 m×0.5 m×0.4 m，相机固联在夹具上，实验平台如图 8 所示。实验中的点阵编码元由反光材料制成，因此使用红外频光源，其闪烁频率与相机采集频率同步，在相机前安装滤镜，便于获得良好的图像质量。制备两种尺寸的点阵编码元标定板，使用多视图几何技术重建标定板三维信息时，一般需要事先设置两个标准距离，一个称为尺度距离，用于恢复尺度信息，另一个称为评价距离，用于评价重建误差。本实验中两块标定板上的标准距离事先由高精度光学仪器测量确定。在实验过程中，将大尺寸标定板摆放在离全局相机约 1.5 m 处，小尺寸标定板离局域相机约 0.3 m 处。多次移动两

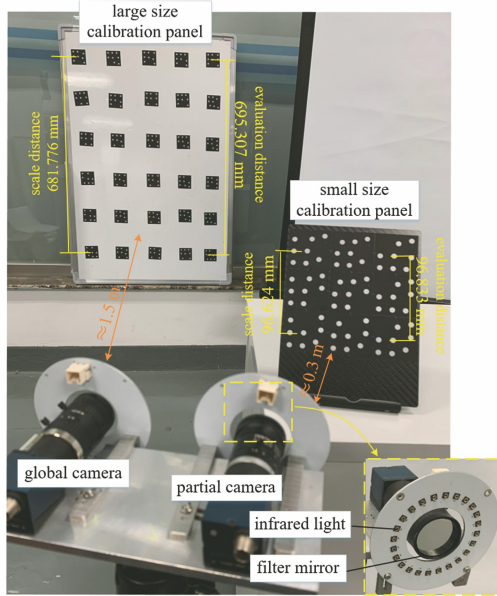


图 8 实验装置

Fig. 8 Experimental device

目视觉系统，在不同视角下两台相机同时采集各自的标定板图像。

5.1 中心定位精度验证

为比较所提基于灰度梯度的亚像素边缘中心定位法和基于整像素边缘的中心定位法的精度，分别采用这两种方法定位大尺寸标定板图像椭圆中心，然后进行解码和匹配，最后由多视图几何技术重建标定板三维信息，并获得评价距离的实测值，共进行 5 组实验。表 1 展示这 5 组实验的评价距离重建误差，即将两种方法重建得到的实测值分别与图 8 中的评价距离 695.307 mm 比较，可以看出，基于整像素边缘的中心定位方法的评价距离重建误差在 0.040 mm 左右，标准差为 0.044 mm，而所提方法误差为 0.010 mm，标准差为 0.011 mm。

表 1 测量距离与评价距离比较结果

Table 1 Comparison of measured distance and evaluation distance

No.	Pixel edge		Sub-pixel edge (proposed algorithm)	
	Reconstructed distance /mm	Error / mm	Reconstructed distance /mm	Error / mm
1	695.347	0.040	695.315	0.008
2	695.348	0.041	695.316	0.009
3	695.268	-0.039	695.298	-0.009
4	695.265	-0.042	695.294	-0.013
5	695.346	0.039	695.317	0.010

5.2 两目视觉系统内外参标定实验

全局相机和局域相机在 10 个不同视角下同时采集各自标定板图像，结果如图 9 所示。经椭圆中心定位、解码和匹配，由多视图几何技术获得两台相机内参数，结果如表 2 所示；每个视角下，获得局域相机坐标系相对于小尺寸标定板坐标系的位姿关系，即 R_A^i 和 t_A^i ；全局相机坐标系相对于大尺寸标定板坐标系的位姿关系，即 R_B^i 和 t_B^i ，其中 $i=1,2,\dots,10$ 。至此，可根据得到的位姿关系建立(3)式，分别使用 Li 等的方法和所提两类优化方法进行求解，通过角度误差(e_R)、位置误差(e_t)和重投影均方根误差(RRMSE)评价外参求解精度，结果如图 10 所示，其中 Li 表示文献[19]提出的方法， c_1 和 c_2 分别表示第一类优化方法和第二类优化方法。

$$\begin{cases} \Delta R^i = (R_Z R_B^i)^T (R_A^i R_X) \\ e_R = \parallel \text{axis-angle}(\Delta R^i) \parallel_F^2, \end{cases} \quad (15)$$

$$e_t = \parallel (R_A^i t_X + t_A^i) - (R_Z t_B^i + t_Z) \parallel_F^2, \quad (16)$$

表 2 相机内参数

Table 2 Camera intrinsic parameters

Intrinsic parameter	Global camera	Partial camera
(α, β)	(1753.089, 1753.089)	(2358.690, 2358.690)
(u_0, v_0)	(1219.075, 1051.850)	(1256.748, 1063.047)
(k_1, k_2)	(-0.201, 0.218)	(-0.114, 0.182)
(p_1, p_2)	(0.000624, -0.000826)	(0.000295, 0.000240)

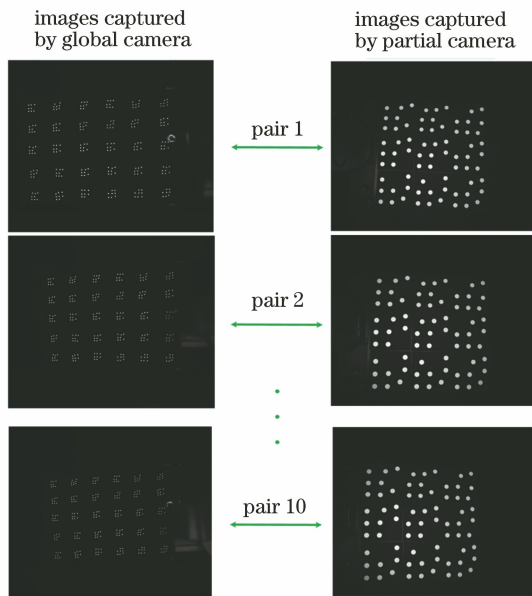


图 9 两目视觉系统采集的图像

Fig. 9 Images captured by two-camera vision system

$$E_{RRMS} =$$

$$\sqrt{\frac{1}{n} \sum_{j=1}^n \| \mathbf{p}_{ij} - \text{Proj}(\mathbf{K}_P, [\mathbf{X}(\mathbf{B}^i)^{-1} \mathbf{Z}^{-1}]_{3 \times 4} \mathbf{S} \mathbf{P}_j) \|^2}, \quad (17)$$

式中: $\text{axis-angle}(\cdot)$ 表示将旋转矩阵转换为轴角向量的形式。

由图 10 可以看出, 经所提方法优化后, 位置误差和重投影均方根误差明显降低。使用第一类优化

方法后, 位置误差的均值从 4.45 mm 降至 1.08 mm, 重投影均方根误差的均值从 29.66 pixel 降至 10.23 pixel; 经第二类方法优化后, 位置误差的均值为 0.77 mm, 重投影均方根误差的均值为 1.72 pixel。但无论哪一种优化方法, 角度误差都增大了, 均值从 0.02° 增大到 0.33° , 这是因为所提两类优化方法将旋转和平移全都参与优化, 而初值求解误差主要集中于平移部分, 因此本文将旋转量固定, 只对平移量进行优化, 即在 (12) 式中只优化 t_x 和 t_z , (14) 式中只优化 t_x 和 t_z 。图 11 展示了只优化平移量的结果, 与图 10 相比较, 只优化平移量后最终所得误差更小。由于旋转量不参与优化, 因此角度误差不变, 误差均值为 0.02° 。使用第一类优化方法, 位置误差的均值为 0.22 mm, 重投影均方根误差的均值为 2.00 pixel; 经第二类方法优化后, 位置误差的均值为 0.35 mm, 重投影均方根误差的均值为 1.74 pixel。可以看出, 第一类优化方法在提高位置精度方面更具优势, 而第二类优化方法侧重于减小重投影误差量, 这两类方法在一定程度上均提高了外参求解精度。

最后, 通过改变两块标定板的相对位姿关系, 对两目视觉系统的外参数进行了 15 次标定。每次标定中, 使用 Li 的方法获得外参初值, 由所提两类方法迭代优化求解, 15 次标定结果的误差如图 12 所

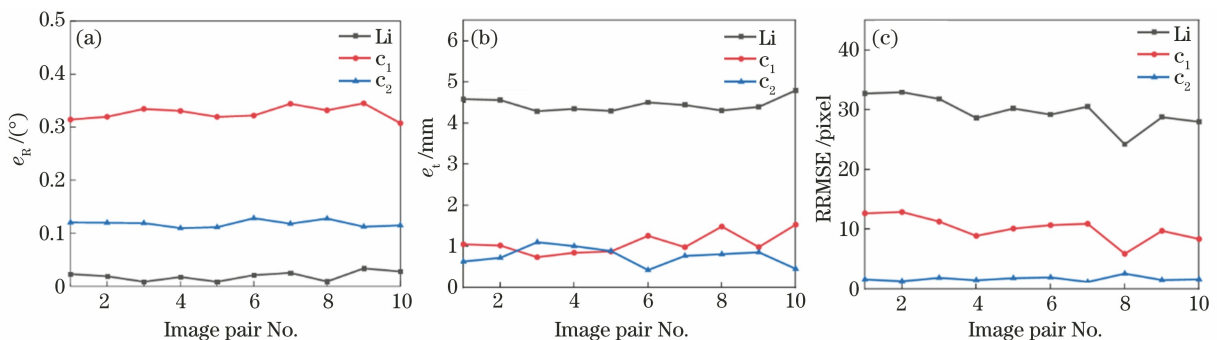


图 10 旋转和平移同时参与优化时的外参数求解误差。(a) 角度误差; (b) 位置误差; (c) 重投影均方根误差

Fig. 10 Error of extrinsic parameters when rotation and translation components optimize simultaneously.

(a) Angle error; (b) position error; (c) reprojection root-mean-squared error

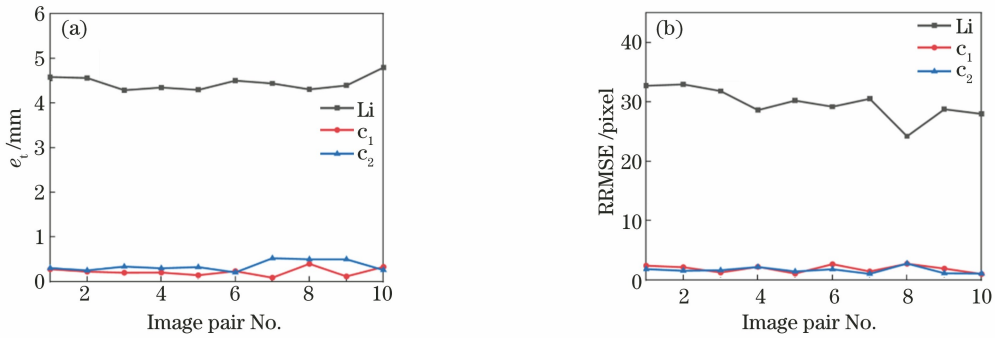


图 11 只有平移参与优化时的外参数求解误差。(a) 位置误差;(b) 重投影均方根误差

Fig. 11 Error of extrinsic parameters when only translation component optimizes. (a) Position error; (b) reprojection root-mean-squared error

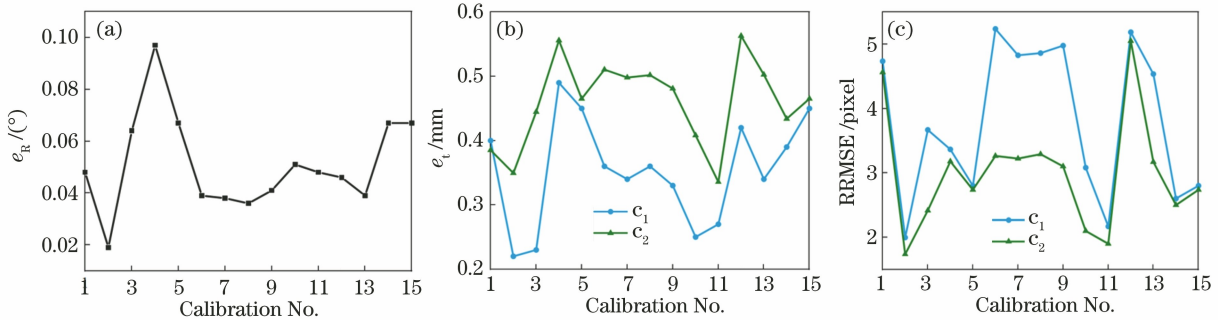


图 12 15 次标定实验结果的误差。(a) 角度误差;(b) 位置误差;(c) 重投影均方根误差

Fig. 12 Error of fifteen calibration experiments. (a) Angle error; (b) position error; (c) reprojection root-mean-squared error

示,由于角度部分不参与优化,因此两种方法所求外参的角度误差相同。其中,角度误差的均值为 0.05° ,标准差为 0.019° ;位置误差方面,第一类优化方法所得误差均值为 0.36 mm ,标准差为 0.083 mm ,第二类优化方法分别是 0.46 mm 和 0.066 mm ;在重投影均方根误差方面,第一类优化方法所得误差均值为 3.73 pixel ,标准差为 1.152 pixel ,第二类优化方法分别是 2.98 pixel 和 0.872 pixel 。由此可见,无论使用哪一种误差评价方式,第二类优化方法都能获得更稳定的外参标定结果。因此,取第二类优化方法求解的 15 个外参数的平均值,作为最终的外参标定结果,结果为

$$\mathbf{X} = \begin{bmatrix} 0.983 & -0.008 & -0.185 & -129.710 \\ 0.009 & 0.100 & 0.001 & 0.089 \\ 0.185 & -0.003 & 0.983 & -21.060 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

5.3 机械臂视觉系统定位实验

为进一步验证外参标定结果的有效性,在 UR 机械臂上搭建了所提远近视距两目视觉系统,两目系统与机械臂固联。放置一块编码元在实验平台

上,编码元上八个圆点中心的相对位置事先由高精度光学仪器测得,如图 13 所示。实验过程如下。

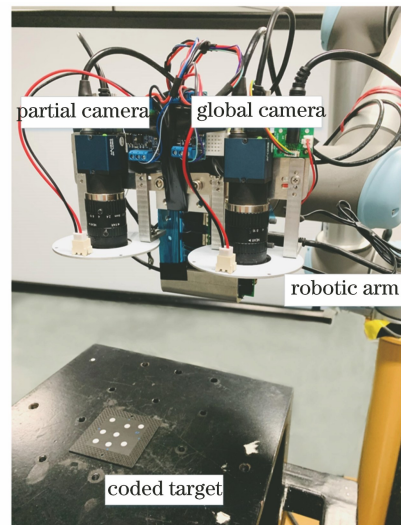


图 13 机械臂视觉系统

Fig. 13 Vision-based robotic arm system

1) 调整机械臂使得局域相机在光轴方向上距编码元约 0.3 m 远,且编码元处于局域相机视场中间位置,局域相机采集编码元图像。经椭圆中心定位后由 PnP 算法^[21]解算编码元坐标系 $\{C\}$ 相对于

局域相机坐标系 $\{P\}$ 的位姿关系, 将其定义为期望位姿 ${}^P_c T_E$ 。本实验中 ${}^P_c T_E = (0.032, 0.008, 0.136, -1.043, -0.841, 31.609)$, 前三个分量以轴角形式表示角度, 单位为弧度, 后三个分量表示位置, 单位为 mm, 如图 14(a) 所示。

2) 移动机械臂使得全局相机在光轴方向上距编码元约 1.5 m, 全局相机采集编码元图像, 并计算编码元坐标系 $\{C\}$ 相对于全局相机坐标系 $\{G\}$ 的位姿关系, 将其定义为全局相机的初始位姿 ${}^G_c T_1$ 。本实验设置了 5 个不同的初始位姿 ${}^G_c T_1$, 分别为 ${}^G_c T_1^1 = (0.144, 0.136, 0.007, -67.020, 244.811, 1489.016)$, ${}^G_c T_1^2 = (0.200, -0.087, 0.0120, 268.691, 58.918, 1552.250)$, ${}^G_c T_1^3 = (0.106, 0.289, -0.015, -71.044, -23.679, 1417.887)$, ${}^G_c T_1^4 = (-0.056, 0.068, -0.014, 65.893, -176.780, 1365.383)$, ${}^G_c T_1^5 = (0.303, 0.157, 0.012, -68.952, 285.48, 1472.019)$, 如图 14(b) 所示。

3) 分别使用 Li 方法和所提两类优化方法所标定的外参 ${}^P_c T$ 计算出编码元坐标系相对于局域相机坐标系的位姿关系 ${}^P_c T_1 = {}^P_c T \cdot {}^G_c T_1$, 定义为局域相机的初始位姿。

4) 局域相机的初始位姿 ${}^P_c T_1$ 与期望位姿 ${}^P_c T_E$ 存在偏差, 欲使其达到期望位姿状态, 可通过偏差值解算出机械臂的运动量^[22]。机械臂根据该运动量调整姿态后, 编码元坐标系相对于局域相机坐标系的实际位姿关系为 ${}^P_c T_A$, 如图 14(c) 所示。

通过计算 $\Delta_c T = {}^P_c T_A - {}^P_c T_E$ 来反映外参标定精度, 统计 $\Delta_c T$ 中的角度误差和平移误差, 结果如图 15 所示。从理论上分析, Li 方法和所提两类优化方法所标定外参中的旋转量相同, 视觉定位结果中的角度误差也应相同。受机械臂定位精度等影响, 实际定位结果中角度误差略有不同, 但基本处于同一数量级。而位置误差可明显反映出所提方法所标定的外参精度更高, 定位精度可提高约 50%。

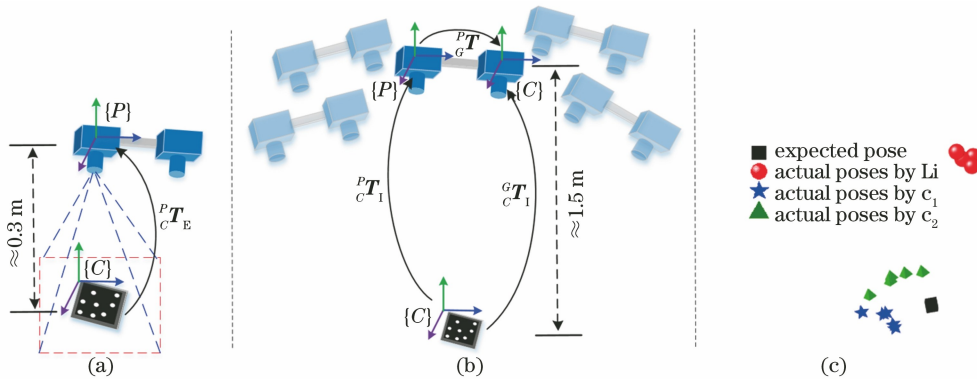


图 14 实验过程示意图。(a)期望位姿;(b)初始位姿;(c)实际位姿

Fig. 14 Schematic of experiment procedure. (a) Expected pose; (b) initial poses; (c) actual poses

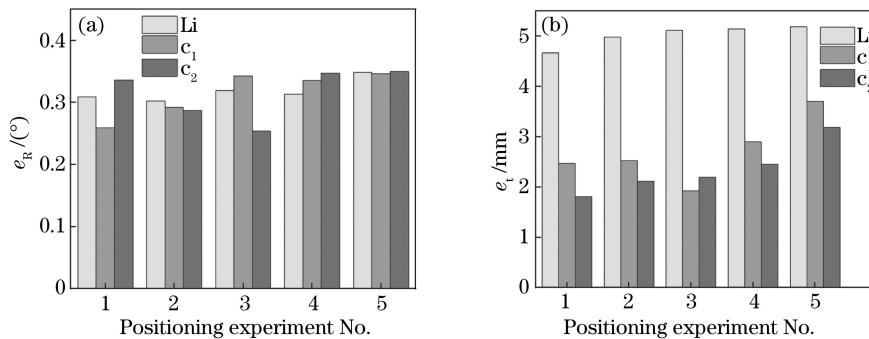


图 15 视觉系统定位误差。(a)角度误差;(b)位置误差

Fig. 15 Positioning error of vision system. (a) Angle error; (b) position error

6 结 论

针对视距不同的两目视觉系统无法由传统方法校准内外参数这一问题, 研究了一种基于多视图几

何技术的标定方法, 结果表明该方法有效可靠, 标定精度满足需求。首先, 提出了一种基于灰度梯度的亚像素边缘定位方法, 所拟合的椭圆边缘更加平滑, 实验证明该方法能获得更精确的中心定位结

果。其次,改进了点阵编码元的解码算法,验证了算法的正确性,在此基础上由多视图几何技术获取相机内参数和相机相对于标定板的位姿关系。最后,通过构建手眼标定方程式解算两目系统的外参数,并对两种外参优化方法展开分析。与传统求解 $AX=ZB$ 方法相比较,所提方法得到的外参数精度更高,其中角度误差为 0.05° ,位置误差降至 0.36 mm ,重投影均方根误差降至 2.98 pixel 。在机械臂定位实验中,所提方法所标定的外参提高了位置定位精度约 50% 。所提方法不仅限于两目视觉系统的内外参标定,可扩展到多目系统的标定,但后续工作主要集中于提高标定结果的精度。

参 考 文 献

- [1] Xie Z X, Chen W Z, Chi S K, et al. Industrial robot positioning system based on the guidance of the structured-light vision[J]. *Acta Optica Sinica*, 2016, 36(10): 1015001.
解则晓, 陈文柱, 迟书凯, 等. 基于结构光视觉引导的工业机器人定位系统[J]. *光学学报*, 2016, 36(10): 1015001.
- [2] Lei J Z, Zeng L B, Ye N. Research on industrial robot alignment technique with monocular vision[J]. *Optics and Precision Engineering*, 2018, 26(3): 733-741.
雷金周, 曾令斌, 叶南. 工业机器人单目视觉对准技术研究[J]. *光学精密工程*, 2018, 26(3): 733-741.
- [3] Qi N M, Xia Q, Guo Y B, et al. Pose measurement model of space cooperative target capture based on zoom vision system [J]. *Advances in Mechanical Engineering*, 2016, 8(7): 1-8.
- [4] Zhou Q F, Liu J H. Rapid nonlinear distortion correction of aerial optical zoom lens system[J]. *Acta Optica Sinica*, 2015, 35(4): 0411001.
周前飞, 刘晶红. 航空变焦镜头非线性畸变快速校正方法[J]. *光学学报*, 2015, 35(4): 0411001.
- [5] Qin Z K, Wang P, Sun J, et al. Precise robotic assembly for large-scale objects based on automatic guidance and alignment [J]. *IEEE Transactions on Instrumentation and Measurement*, 2016, 65 (6): 1398-1411.
- [6] Sun N, Zhang L Y, Ye N, et al. Calibration of binocular stereo measurement system based on calibration cross [J]. *Acta Optica Sinica*, 2012, 32 (4): 0412001.
孙楠, 张丽艳, 叶南, 等. 基于十字靶标的双目立体测量系统标定 [J]. *光学学报*, 2012, 32 (4): 0412001.
- [7] Xu Q Y, Yang L H, Niu Z Y, et al. High-accuracy calibration of multi-camera positioning system based on precision angle reference [J]. *Acta Optica Sinica*, 2019, 39(1): 0115001.
徐秋宇, 杨凌辉, 牛志远, 等. 精密角度基准下的多相机定位系统高精度标定 [J]. *光学学报*, 2019, 39 (1): 0115001.
- [8] Liu Z, Zhang G J, Wei Z Z, et al. Novel calibration method for non-overlapping multiple vision sensors based on 1D target [J]. *Optics and Lasers in Engineering*, 2011, 49(4): 570-577.
- [9] Xia R B, Hu M B, Zhao J B, et al. Global calibration of multi-cameras with non-overlapping fields of view based on photogrammetry and reconfigurable target [J]. *Measurement Science and Technology*, 2018, 29 (6): 065005.
- [10] Genovese K, Chi Y X, Pan B. Stereo-camera calibration for large-scale DIC measurements with active phase targets and planar mirrors [J]. *Optics Express*, 2019, 27(6): 9040-9053.
- [11] Wang G, Shang Y, Guan B L, et al. Flexible calibration of setting relation of a multi-camera rig for non-overlapping views [J]. *Chinese Journal of Lasers*, 2017, 44(6): 0604004.
王刚, 尚洋, 关棒磊, 等. 无重叠视场多相机组安装关系的灵活标定 [J]. *中国激光*, 2017, 44 (6): 0604004.
- [12] Agarwal S, Snavely N, Simon I, et al. Building Rome in a day [C] // 2009 IEEE 12th International Conference on Computer Vision, September 29-October 2, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 72-79.
- [13] Kisworo M, Venkatesh S, West G. Modeling edges at subpixel accuracy using the local energy approach [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(4): 405-410.
- [14] Canny J. A computational approach to edge detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6): 679-698.
- [15] Sanida T, Sideris A, Dasygenis M. A heterogeneous implementation of the Sobel edge detection filter using OpenCL [C] // 2020 9th International Conference on Modern Circuits and Systems Technologies (MOCASST), September 7-9, 2020, Bremen, Germany. New York: IEEE Press, 2020.
- [16] Yan J, Ye N, Li T C, et al. Research and implementation of industrial photogrammetry without coded points [J]. *Acta Optica Sinica*, 2019, 39(10): 1015002.
严俊, 叶南, 李廷成, 等. 无编码点的工业摄影测量技术的研究及实现 [J]. *光学学报*, 2019, 39(10): 1015002.

- [17] Ye N, Zhu H Y, Wei M Q, et al. Accurate and dense point cloud generation for industrial measurement via target-free photogrammetry [J]. *Optics and Lasers in Engineering*, 2021, 140: 106521.
- [18] Moulon P, Monasse P, Marlet R. Adaptive structure from motion with a contrario model estimation[M]//Fitzgibbon A, Lazebnik S, Perona P, et al. *Computer vision-ACCV 2012. Lecture notes in computer science*. Heidelberg: Springer, 2013, 7727: 257-270.
- [19] Li A G, Lin W, Wu D F. Simultaneous robot-world and hand-eye calibration using dual-quaternions and Kronecker product [J]. *International Journal of Physical Sciences*, 2010, 5(10): 1530-1536.
- [20] Tabb A, Ahmad Yousef K M. Solving the robot-world hand-eye(s) calibration problem with iterative methods [J]. *Machine Vision and Applications*, 2017, 28(5/6): 569-590.
- [21] Terzakis G, Lourakis M. A consistently fast and globally optimal solution to the perspective-n-point problem[M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12346: 478-494.
- [22] Zhu H Y, Zeng L B, Shi S F, et al. Research on visual alignment technology of industrial robotic arm in multi-station operation [J]. *Aeronautical Manufacturing Technology*, 2020, 63(19): 56-65.
祝鸿宇, 曾令斌, 石世锋, 等. 工业机械臂多工位作业下的视觉对准技术研究[J]. *航空制造技术*, 2020, 63(19): 56-65.