

# 老年服务机器人视觉定位方法研究

王俊<sup>1,2</sup> 梅涛<sup>1</sup> 孔斌<sup>1</sup> 董翔<sup>1</sup>

(1 中国科学院合肥智能机械研究所, 合肥 230031;

2 中国科学技术大学自动化系, 合肥 230026)

**摘要** 提出一种视觉定位方法对影响老年服务机器人的视觉定位精度的摄像机标定、手眼系统标定以及三维重建 3 个部分进行优化。首先对多幅标定图像的结果进行最小二乘处理, 剔除大权重的误差样本以减小标定的统计误差; 通过建立过渡空间坐标系的方法对手眼系统进行标定, 减少多次计算转移矩阵所造成的累积误差; 利用最小二乘拟合精确确定目标的匹配点, 减小三维重建所带来的误差。实验结果表明: 使用该方法使机器人获取目标空间位置的误差小于 2 mm, 满足老年服务机器人路径规划及使用手爪抓取物体的需要。

**关键词** 服务机器人; 摄像机标定; 三维重建; 手眼系统标定; 转移矩阵

**中图分类号** TP24 **文献标志码** A **文章编号** 1671-4512(2011)S2-0255-04

## Research on vision localization method of robot for elderly

Wang Jun<sup>1,2</sup> Mei Tao<sup>1</sup> Kong Bin<sup>1</sup> Dong Xiang<sup>1</sup>

(1 Institute of Intelligent Machines, Chinese Academy of Science, Hefei 230031, China;

2 Department of Automation, University of Science and Technology, Hefei 230026, China)

**Abstract** Vision localization precision of robot for elderly was influenced by error of camera calibration, eye-in-hand frame calibration, and three-dimensional reconstruction. A method to optimize process of the three parts base on the reason was proposed. Frist process the least square of many calibration image, eliminate the big weight error stylebook to reduce statistic error of calibration. Calibrate the eye-in-hand frame by setting up transition space coordinate frame, reduce the cumulate error caused by repeated computation transfer matrix. Use least square fitting to fix on the matching point of object exactly, reduce the error of three-dimensional reconstruction. The result of experiment show that the error of robot obtain the space position of object less than 2 mm by using this method, meet the needs of path programming and grab objects with gripper of robot for elderly.

**Key words** camera calibration; three-dimensional reconstruction; eye-in-hand frame calibration; transfer matrix

在老年服务机器人系统中, 在机器人进行路径规划前, 须计算出目标相对于机器人坐标系的位置; 当机器人执行抓水杯等动作时, 须通过视觉系统准确给出目标相对于机器人手臂坐标系的空间位置, 因此目标定位的精度直接影响机器人完成路径规划及抓取的能力。

老年服务机器人采用双目视觉系统实现目标定位, 通过双目视觉系统计算出目标相对于机器人手臂坐标系的空间位置。整个过程分 3 个步骤

实现: 一是通过摄像机标定得出摄像机的内外参数; 二是通过寻找匹配点计算出匹配点的空间坐标; 三是通过计算空间转移矩阵将视觉计算的坐标转化到机器人手臂坐标系。因此老年服务机器人的目标定位精度由摄像机标定<sup>[1]</sup>、三维重建<sup>[2]</sup>以及手眼系统坐标系转换<sup>[3]</sup> 3 部分误差决定。本研究在此基础上提出一种机器人目标定位方法, 该方法利用最小二乘法剔除大权重的误差标定样本对相机外参数计算进行优化, 减小相机标定所

收稿日期 2011-06-18.

作者简介 王俊(1982-), 男, 博士研究生, E-mail: wangjun@iim.ac.cn.

基金项目 中国科学院知识创新工程重要方向性项目 (KGCX2-YW-154); 国家自然科学基金资助项目 (61005075).

带来的误差;利用建立中间过渡空间坐标系的方法对手眼系统进行标定,减小求取转移矩阵的误差;利用最小二乘拟合精确确定目标的匹配点,减小在三维重建中所带来的误差.

### 1 双目视觉坐标系标定方法

#### 1.1 双目视觉系统标定基本原理

双目立体视觉是基于视差、根据三角法原理进行三维信息获取<sup>[4]</sup>,即由 2 个摄像机的图像平面和被测物体之间构成一个三角形.如图 1 所示,已知左右 2 摄像机之间的位置关系,便可获取 2 摄像机公共视场内物体的三维尺寸及空间物体特征点的三维坐标.

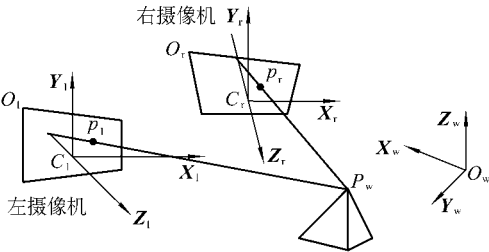


图 1 双目立体视觉系统结构示意图

双目立体视觉系统的标定是指左右摄像机的内部参数标定后,确定视觉系统的结构参数  $R$  和  $T$ ,即左右 2 个摄像机坐标系之间的旋转平移关系<sup>[5]</sup>.双目立体视觉系统中左右摄像机的外部参数分别为  $R_l, T_l$  和  $R_r, T_r$ , $R_l$  和  $T_l$  表示左摄像机与世界坐标系的相对位置, $R_r$  和  $T_r$  表示右摄像机与世界坐标系的相对位置.对于任意一点,若它在世界坐标系、左摄像机坐标系和右摄像机坐标系下的非齐次坐标分别为  $X_w, X_l$  和  $X_r$ ,则有  $X_l = R_l X_w + T_l$ ,  $X_r = R_r X_w + T_r$ . 消去  $X_w$ , 得  $X_r = R_r R_l^{-1} X_l + T_r - R_r R_l^{-1} T_l$ , 因此 2 个摄像机之间的结构参数  $R$  和  $T$  可以用关系式表示为:

$$R = R_r R_l^{-1}; \quad T = T_r - R_r R_l^{-1} T_l.$$

若对双摄像机分别标定,得到  $R_l, T_l$  和  $R_r, T_r$ , 则 2 个摄像机的结构参数就可以由上式计算得到.

#### 1.2 多幅标定图像处理优化方法

本研究采用的是亚像素提取棋盘格角点的方法对摄像机进行标定的<sup>[6]</sup>.在摄像机标定过程中,左右摄像机均拍摄  $m$  幅标定图像,每幅标定图像上有  $n$  个角点,如图 2 所示.

设定  $p_{ij2}$  为右摄像机第  $i$  幅图像上的点  $p_j$  的真实图像坐标,  $p_{ij1}$  为左摄像机第  $i$  幅图像上的点  $p_j$  的真实图像坐标,  $\hat{p}(R, T, p_{ij1})$  为点  $p_{ij1}$  通过  $R$

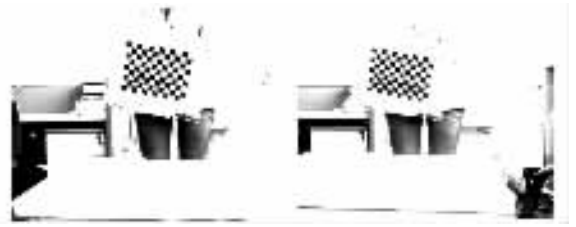


图 2 2 个相机标定图像角点提取情况

和  $T$  矩阵转换到右摄像机上的图像坐标,利用最小二乘法可以求得使  $\sum_{i=1}^n \sum_{j=1}^m |p_{ij2} - \hat{p}(R, T, p_{ij1})|^2$  最小的  $R_1$  和  $T_1$ <sup>[7]</sup>.

因最小二乘法对大误差赋予大权重,某些与统计结果偏差大的标定图像对视觉系统标定误差影响较大,须将这些标定图像筛选出来.通过  $R_1$  和  $T_1$  求得点  $p_{ij1}$  转换到右摄像机上的图像坐标  $\hat{p}(R_1, T_1, p_{ij1})$ , 令  $S_i = \sum_{j=1}^m |p_{ij2} - \hat{p}(R_1, T_1, p_{ij1})|^2$ , 如果  $|p_{ij2} - \hat{p}(R_1, T_1, p_{ij1})| > S_i K$ ,  $K$  为某一判定系数,就将第  $i$  幅标定图像筛选掉,处理完后还剩下  $m'$  幅标定图像,再对这  $m'$  幅标定图像进行最小二乘处理得到  $R'$  和  $T'$ ,  $R'$  和  $T'$  就是通过最小二乘优化后的双目视觉系统标定结果.

### 2 手眼系标标定方法

通过标定得到的视觉坐标系不能具体确定其原点及空间坐标轴方向,因此不能直接通过刚体转换的方式得到视觉坐标系与机器人坐标系的空间关系.解决办法可以参照双目视觉系统标定方法<sup>[8]</sup>,即找 4 个同一平面空间点,4 个点到机器人坐标系位置关系已知,组成一个类似标定板的坐标系  $S$ .先计算视觉坐标系到  $S$  坐标系的  $R_v$  和  $T_v$  与机器人坐标系到  $S$  坐标系的  $R_l$  和  $T_l$ ,再计算得到视觉坐标系到机器人坐标系的  $R$  和  $T$ .此方法须要计算 2 次转移矩阵,并且选取 4 个在同一平面的点,这在实际操作中很难保证,通常须要对 4 个点进行平面拟合,从而增加了计算的复杂度及累积误差.

本研究采用建立过渡空间坐标系的方法,即找 4 个不在同一平面的点,4 个点相对机器人坐标系坐标已知,为  $(t_{x_i}, t_{y_i}, t_{z_i}) (i=1, 2, \dots, 4)$ ,然后再通过双目视觉系统得到这 4 个点相对视觉坐标系的坐标为  $(v_{x_i}, v_{y_i}, v_{z_i}) (i=1, 2, \dots, 4)$ ,令视觉坐标系到机器人坐标系的转移矩阵为  $C$ ,则视觉坐标系到机器人坐标系的变换关系为

$$(t_{x_i}, t_{y_i}, t_{z_i}, 1)^T = C(v_{x_i}, v_{y_i}, v_{z_i}, 1)^T.$$

由于 4 个点不在同一平面上,所以  $(v_{x_i}, v_{y_i},$

$v_{z_i}, 1)$  ( $i=1, 2, \dots, 4$ )可逆,有

$$C = (t_{x_i}, t_{y_i}, t_{z_i}, 1)^T [(v_{x_i}, v_{y_i}, v_{z_i}, 1)^T]^{-1}$$

利用此方法只须计算 1 次转移矩阵就能计算出视觉坐标系到机器人坐标系转移矩阵,从而降低了多次计算转移矩阵带来的累积计算误差及计算复杂度.取  $n$  组点(每组 4 个,且不在同一平面),对  $n$  组转移矩阵  $C_1, C_2, \dots, C_n$  进行最小二乘处理,即得到视觉坐标系到机器人坐标系转移矩阵  $C'$ .

### 3 最小二乘拟合匹配点选取

匹配点的选取精确与否对三维重建的结果有直接的影响,匹配点的选取通常是在目标提取的基础上再进行特征点的提取.传统的目标提取方法(如 shift-mean 算法<sup>[9]</sup>、分类器算法<sup>[10]</sup>等)虽然能进行有效识别,却很难精确地确定目标的位置,从而确定匹配点的精度不够高,因此本研究采用最小二乘拟合<sup>[11]</sup>确定匹配点的方法,该方法能对具有规则形状(如椭圆及多边形等)目标或标志精确地提取匹配点.

在双目视觉系统中,目标在左右 2 个相机的投影角度不同,导致目标在 2 个相机中的投影结果也不一样.对于水杯、药瓶等目标,它们的口径部分是圆形的,由于相机成像平面和水杯杯口平面存在一定的夹角,在不考虑相机畸变影响的情况下,它们投射到相机成像平面上都是标准的椭圆.通过透视投影原理可知:由于目标与左右 2 个相机的成像平面的夹角及距离不同,对应投影的椭圆也不同,即使提取椭圆边界上的点进行匹配也很困难.但是 2 个相机投影平面投影的椭圆的圆心都是圆形口径中心的投影,因此选择这 2 个点作为匹配点.

目标边缘检测的结果是计算机获得目标轮廓边缘的部分数据,对这些数据进行拟合可得到整个边界的精确数据.本研究采用最小二乘拟合的方法来获取目标轮廓的边界及中心点数据.设椭圆中心点坐标为  $(x_0, y_0)$ ,边缘检测得到的边界点坐标为  $(x_i, y_i)$ ,拟合的边界点到中心的欧氏距离表示为  $h$ ,则边缘检测得到的边界点到中心点的距离为

$$h_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2};$$

拟合点到中心点坐标为

$$h = \sqrt{(x - x_0)^2 + (y - y_0)^2};$$

拟合点与边界点的距离为

$$\Delta h = \left| \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - \sqrt{(x - x_0)^2 + (y - y_0)^2} \right|.$$

取所有边界点到拟合点距离平方和最小,即为最小二乘意义下的最好拟合.拟合后可得到精确的椭圆中心点坐标上下左右 4 个边界点的坐标,从而可以精确地得到目标的空间信息,图 3 为水杯边缘提取后对杯口进行椭圆拟合的示例.

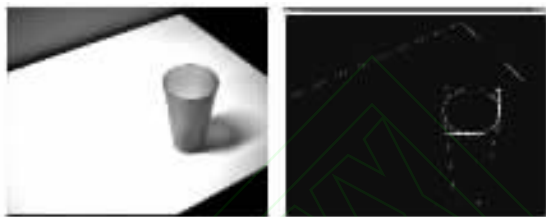


图 3 水杯进行边缘提取后杯口椭圆拟合示例

## 4 实验

### 4.1 实验设计

实验设计为老年服务机器人在展台场景下演示搜寻水杯,然后用手抓取水杯,最后完成倒水这一系列动作.由于老年服务机器人路径规划以及用手爪抓取功能的需要,视觉系统在机器人初始位置抓取物品以及到饮水机前倒水时,须要对目标进行定位,图 4 展示机器人完成倒水这一系列动作过程中利用视觉进行目标定位的情况.



图 4 机器人目标定位实验情况

### 4.2 实验数据分析

将本研究方法计算目标的空间位置数据与 NDI 公司 OPTOTRAK CERTUSTM 三维测量系统测量的数据作对比,如表 1 所示,OPTO-

表 1 视觉测量数据与 NDI 测量数据对比

视觉			NDI			误差		
$x$	$y$	$z$	$x$	$y$	$z$	$\Delta x$	$\Delta y$	$\Delta z$
-28.096	523.641	-454.879	-27.866	523.908	-456.439	0.15	0.26	1.55
2.953	609.957	-397.198	2.178	611.128	-398.206	0.78	1.17	1.01
-191.168	640.094	-248.459	-190.986	641.657	-248.848	0.18	1.56	0.39
4.428	676.349	-347.697	3.131	676.643	-346.785	1.29	0.29	0.91
-140.754	710.700	-153.353	-141.399	711.937	-151.930	0.64	1.23	1.42

mm

TRAK CERTUSTM 三维测量系统测量标准标记点坐标的精度可达到 0.1 mm,所有的测量数据都以机器人坐标系为参照系,测量得到 5 组数据.从表 1 中的数据可知:采用本研究方法计算出来目标中心点三维坐标的误差都控制在 2 mm 以内,超过老年服务机器人路径规划所要求的 1 cm 及手抓取物体所要求的 5 mm 误差的要求.

参 考 文 献

[1] Zhang Z Y. A flexible new technique for camera calibration[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.  
 [2] 陈棣湘,罗飞路,潘孟春.立体视觉测量中的图像匹配策略研究[J]. 光学技术, 2002, 28(4): 392-394.  
 [3] Frank C P, Bryan J M. Robot sensor calibration; solving  $AX = XB$  on the euclidean group. IEEE Trans on Robotics and Automation, 1994, 10 (5): 717-721.  
 [4] 张剑清,潘励,王树根. 摄影测量学[M]. 武汉: 武汉大学出版社, 2003.

[5] 张广军. 机器视觉[M]. 北京: 科学出版社, 2006.  
 [6] Sojka E. A new algorithm for detecting corners in digital images[C]// Pattern Recognition 24th DAGM Symposium Proceedings. [s. n.], 2002, 2449: 125-132.  
 [7] Behrooz K P. Minimization of the quantization error in camera calibration[C]// Proceedings of the DARPA Image Understanding Workshop. [s. n.], 1987, 2: 671~880.  
 [8] Radu H, Fadi D. Hand-eye calibration[J]. The International Journal of Robotics Research, 1995, 14 (3): 195-210.  
 [9] Zhou Shangbo, Hu Peng, Liu Yujiong. Target tracking based on improved mean-shift and adaptive Kalman filter [J]. Journal of Computer Applications, 2010, 30(6): 1573-1576.  
 [10] 潘志松,陈斌,缪志敏,等. One-Class 分类器研究[J]. 电子学报, 2009, 37(11): 2496-2503.  
 [11] 邹益民,江渤. 一种基于最小二乘的不完整椭圆拟合算法[J]. 仪器仪表学报, 2006, 27(7): 808-809.

(上接第 247 页) Conference on Communication Systems, Networks and Applications. Washington DC: IEEE Computer Society, 2010: 52-55.  
 [4] Deb K, Kang S J. Statistical characteristics in HIS color model and position histogram based vehicle license plate detection[J]. Intelligent Service Robotics, 2009, 2(3): 173-186.  
 [5] Mutch J, Lowe D G. Object class recognition and localization using sparse features with limited receptive fields[J]. International Journal of Computer Vision,

2008, 80(1): 45-57.  
 [6] Zhang Zhentao, Xu De. High-speed vision system based on smart camera and its target tracking algorithm[J]. Robot, 2009, 31(3): 229-234.  
 [7] Chen Xiaopeng, Huang Qiang. Rapid and precise object detection based on color histograms and adaptive bandwidth mean shift[C]// 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, Washington DC: IEEE Computer Society, 2009: 4281-4286.