

稀疏立体视觉算法在自动包装机器人中的应用

王杰琼, 张印辉, 何自芬
(昆明理工大学, 昆明 650500)

摘要: **目的** 为实现包装机器人自动测距的功能, 提出将双目稀疏立体视觉算法应用到自动包装机器人视觉系统中。 **方法** 为验证算法的实用性, 利用深圳市元创兴科技有限公司生产的配备双目视觉系统的REBot-V-6R六自由度机器人进行实验, 计算实验台上包装件与摄像机的距离, 通过计算值与测量值之间的误差评估算法的精确度。 **结果** 实验证明双目稀疏立体视觉算法在600~1000 mm范围内深度计算误差小于7%。 **结论** 该稀疏立体视觉算法可以应用于自动包装机器人测距。

关键词: 立体视觉; 自动测距; 标定技术; 机器人

中图分类号: TB486 **文献标识码:** A **文章编号:** 1001-3563(2016)07-0008-05

Application of Sparse Stereo in Automatic Packaging Robot

WANG Jie-qiong, ZHANG Yin-hui, HE Zi-fen
(Kunming University of Science and Technology, Kunming 650500, China)

ABSTRACT: To realize the automatic ranging function of packaging robot, this experiment proposed to apply sparse stereo vision to the packaging robot arm. To evaluate the accuracy of this algorithm, a robot arm REBot-V-6R which was equipped with stereo cameras was used to carry out the experiment to calculate the distance between the package pieces on the experimental table and the camera, and the accuracy of the algorithm was evaluated by the error between the value computed by sparse stereo and the measured value. Experimental results showed that the computing error was below 7% in a range of 600~1000 mm. In conclusion, the algorithm method of sparse stereo can be used in the ranging system of automatic packaging robot.

KEY WORDS: stereo vision; automatic ranging; calibration technology; robot

双目立体视觉是计算机视觉领域的一个重要分支, 其模拟人类双眼观察景物的方式, 通过2台或1台摄像机经过移动或旋转观察同一个场景, 根据拍摄的2幅图像中空间点的视差值恢复其三维信息。双目立体视觉技术可以应用在三维重建、距离测量和目标定位等领域。工业应用中的主动测距技术使用结构光^[1-2]、激光雷达^[3-4]等传感技术获取距离信息。在包装工业尤其是自动化包装中, 利用双目立体视觉技术计算包装件, 或者待包装产品与包装机械手

之间距离的被动测距技术^[5-10]结构简单易实现, 且对环境要求少, 具有广阔的应用前景。文中提出将稀疏立体视觉算法应用于包装工业, 通过双目立体视觉技术实现包装机器人自动测距。

实验包括图像采集、摄像机标定、特征提取、特征匹配、深度恢复等5个部分。图像采集即利用双目摄像机系统通过不同角度拍摄同一个场景的2幅图像。摄像机标定的目的是计算所用摄像机的内参数和外参数, 从而得到相机的透视投影模型, 即空间点和其像点

收稿日期: 2015-08-25

基金项目: 国家自然科学基金(61302173, 61461022)

作者简介: 王杰琼(1989—), 女, 河北石家庄人, 昆明理工大学硕士生, 主攻双目立体视觉。

通讯作者: 张印辉(1977—), 男, 河北故城人, 博士, 昆明理工大学副教授、硕士生导师, 主要研究方向为图像处理、机器视觉与机器智能。

之间的投影关系。文中采用介于传统标定算法和自标定算法之间的张正友标定法^[12]计算双目摄像机参数。应用鲁棒加强特征(SURF)^[11,13]算法检测图像稀疏的特征点,并且构造特征点的64维特征描述算子,根据描述符号间的欧式距离确定立体匹配。完成了摄像机标定和立体匹配得到空间点和像点之间的投影关系,以及左、右2幅图的特征点对应关系之后,利用射线交汇法^[14]计算匹配点所对应空间点的坐标值。

1 图像采集

REBot-V-6R 机器人采用的双目摄像机系统是2个型号为 acA1300-30gc 的 BASLER 工业摄像机,该相机装配 Sony ICX445 CCD 感光芯片,见图1。文中将左摄像机坐标系视为世界坐标系,在实验台距左摄像机 600~1000 mm 范围内,用双目摄像机每隔 100 mm 拍摄包装件图像对,图像大小为像素。实验用包装件高度为 160 mm。图2为实验采集图像,第1行为每组立体图像对的左图像,第2行为右图像。从左至右包装件上表面与左摄像机距离分别为 440, 540,

640, 740, 840 mm。

2 摄像机标定

摄像机标定的目的是计算文中所用摄像机的焦距以及摄像机坐标系相对于世界坐标系的旋转矩阵和平移矩阵,即获取摄像机的内、外参数,从而得到摄像机的透视投影模型。摄像机标定结果的好坏会直接影响后续的立体匹配和三维坐标计算。现有的摄像机标定技术分为传统的摄像机标定技术和自标定技术。传统的摄像机标定技术使用标定参照物,利用物体上一些已知点的三维坐标和它们相应的图像坐标,来计算摄像机的内、外参数。这种方法主要包括利用最优化算法的标定方法、利用摄像机变换矩阵的标定方法、两步法、双平面标定方法等。自标定技术不需要使用标定块,利用多幅图像对应点之间的关系进行标定。相比于自标定技术,传统的摄像机标定技术过程比较复杂,需要使用高精度的标定块,而实际情况中经常无法使用标定块。自标定技术克服了传统标定方法操作繁琐、设备要求高的缺点,同时标定精度高于自标定方法。

张氏平面标定法假设标定模板平面位于 $z=0$ 的世界坐标系上,摄像机通过不同角度采集模板平面的图像,提取图像中的特征点奇次坐标 $\tilde{m} = [u, v, 1]^T$,且已知特征点的物理坐标 $\tilde{M} = [X, Y, Z, 1]^T$,由张氏标定法^[12]中根据空间点及其像点之间的关系计算相机参数:

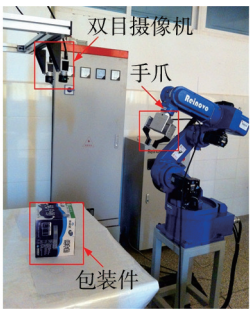


图1 包装机器人
Fig.1 Packaging robot



图2 立体图像对
Fig.2 Stereo image pair

$$s\tilde{m} = K[Rt]\tilde{M} \quad (1)$$

式中: s 为尺度因子; R, t 分别为摄像机坐标系相对于空间坐标系的旋转矩阵和平移向量; K 为摄像机内参数矩阵,定义为:

$$K = \begin{bmatrix} f_u & c & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

式中: (u_0, v_0) 为摄像机主点坐标; f_u, f_v 分别为 u 轴和 v 轴的比例因子; c 为因摄像机两轴不垂直造成的偏移。由于张氏标定法^[12]假设模板平面位于 $z=0$ 的世界坐标系,则有:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K[r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K[r_1 \ r_2 \ t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3)$$

式中: r_i 为旋转矩阵 R 的列向量。张氏标定^[12]令 $\tilde{M} = [X, Y, 1]^T$, $\tilde{m} = [u, v, 1]^T$, 得到:

$$s\tilde{m} = H\tilde{M} \quad (4)$$

其中:

$$H = K[r_1 \ r_2 \ t] = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \quad (5)$$

H 为 3×3 的单应性矩阵。如果所使用的模板平面的尺寸已知,那么可以得到空间点 M 和像点 m 的坐标值,则根据式(3)和(5)就可以计算单应性矩阵 H 。式(6)为张氏标定^[12]表示单应性矩阵:

$$H = [h_1 \ h_2 \ h_3] = \lambda K[r_1 \ r_2 \ t] \quad (6)$$

式中: λ 为比例因子; r_1 和 r_2 为单位矩阵且正交。

于是有 $r_1^T r_1 = r_2^T r_2 = 1$ 和 $r_1^T r_2 = 0$, 即:

$$\begin{cases} h_1^T K^{-T} K^{-1} h_2 = 0 \\ h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \end{cases} \quad (7)$$

式(7)为张氏标定法^[12]计算内参数矩阵的2个约束条件。每一个摄像机包括5个未知内参数,当采用 $n \geq 3$ 张图像时,便可计算得到摄像机的内参数矩阵 K 。计算得到摄像机内参数后,通过已知参数计算两视角之间的本质矩阵,通过分解本质矩阵得到摄像机外参数^[11]。

加州理工大学的Matlab标定工具箱是基于张氏标定法开发的。文中使用该工具箱对所用立体摄像机进行标定,其中左、右摄像机均通过不同角度采集标定板的7幅图像用于内参数标定。且文中计算包装件的深度信息以左摄像机坐标系为世界坐标系,则外

参数标定为计算右摄像机坐标系相对于左摄像机坐标系的旋转矩阵和平移向量。完成摄像机内参数标定后的后投影误差见图3,图3a,b分别代表左、右摄像机后投影误差。图3中“+”代表标定模板7幅图像的后投影误差点,可看出左摄像机后投影误差密集且均在像素范围内,右摄像机后投影误差较左摄像机密集度差,但大多数后投影误差也均在像素范围内,可以认为标定结果可靠。

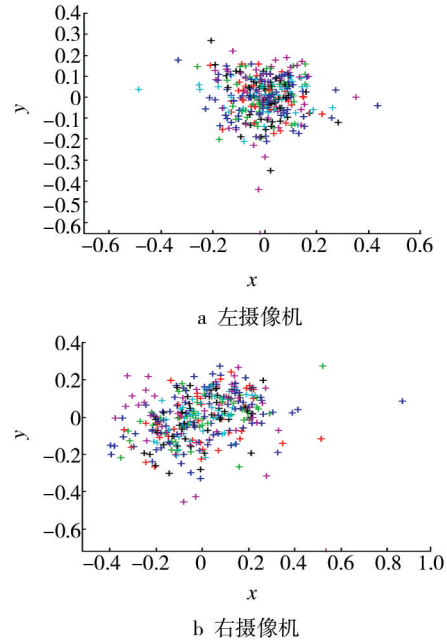


图3 后投影误差

Fig.3 Reprojection error

文中使用Matlab标定工具箱计算左、右摄像机内参数标定结果分别为 $f_{lu}=2238.24$, $f_{lv}=2243.26$, $f_{ru}=2239.43$, $f_{rv}=2242.58$, 所用摄像机像素尺寸为 $d_i=3.75 \mu m$ 。实验所用双目立体视觉系统焦距 $f=d_i \times (f_{lu}+f_{lv}+f_{ru}+f_{rv})/4=8.40 mm$, 计算得到右摄像机坐标系相对于左摄像机坐标系(世界坐标系)的旋转矩阵和平移向量分别为:

$$R = \begin{bmatrix} 0.9979 & -0.0606 & -0.0246 \\ 0.0616 & 0.9973 & 0.0402 \\ 0.0221 & -0.0417 & 0.9989 \end{bmatrix}, T = \begin{bmatrix} 0.3000 \\ 0.6674 \\ 0.8917 \end{bmatrix}$$

其中:旋转矩阵 R 的列向量 r_1, r_2 和 r_3 分别为右摄像机坐标系相对于左摄像机坐标系 x, y 和 z 轴的旋转向量;平移向量 T 的参数分别为右摄像机相对于左摄像机坐标系3个轴的平移量。

3 特征匹配

立体匹配是双目立体视觉的关键技术,立体匹配

常用的特征有点特征、线特征和区域特征。文中采用SURF^[11,13]算法提取图像特征点,SURF算法是在尺度不变特征变换(SIFT)^[14]算法的基础上提出的,与SIFT不同之处在于,SURF将积分图用于图像卷积大大加快了运算速度,并且采用近似Hessian矩阵检测特征点以及构造基于Haar小波值分布的特征描述符号。表1是分别利用SIFT算法和SURF算法对5组实验所用图像对完成匹配所用时间,可以看出SURF算法速度较快。该算法包括特征检测、特征描述和特征匹配等3个部分。在完成了前2步得到特征描述符后,利用描述符之间的欧式距离完成立体匹配,得到立体图像对的稀疏特征匹配。图4分别为5组立体图像对中左图像提取SURF特征点的结果,亮点表示特征点位置。

表1 SIFT和SURF匹配运算时间
Tab.1 Computation time by SIFT and SURF

组类	SIFT/s	SURF/s
1	25.3	17.3
2	25.3	17.7
3	26.5	18.4
4	27.3	18.2
5	26.4	17.8

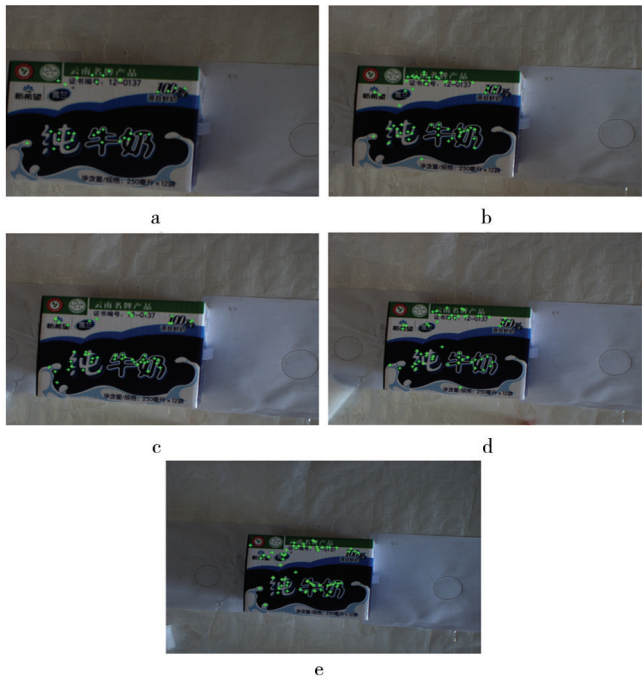


图4 SURF特征
Fig.4 Images of SURF feature

4 深度计算

在完成了摄像机标定和特征匹配得到立体摄像

机的透视投影模型和匹配点之间相互关系后,提取匹配特征点对应的空间点的三维坐标就相对容易了。根据双目立体视觉的几何原理^[15],利用射线交汇法^[11],即空间点位于其2个像点和相应摄像机中心所构成的2条射线相交的位置,因此应用已知的摄像机透视投影模型和匹配特征点的相互关系,便可得到空间点的三维坐标。文中提取每组立体图像对中50个特征点,计算每个点在左摄像机坐标系(世界坐标系)下的三维坐标,从而恢复空间点的深度信息。图5为5组立体图像对的深度计算结果,考虑到图片清晰度,仅提取每组图像对的20个特征点,每个特征点标记其对应空间点的深度值(m)。

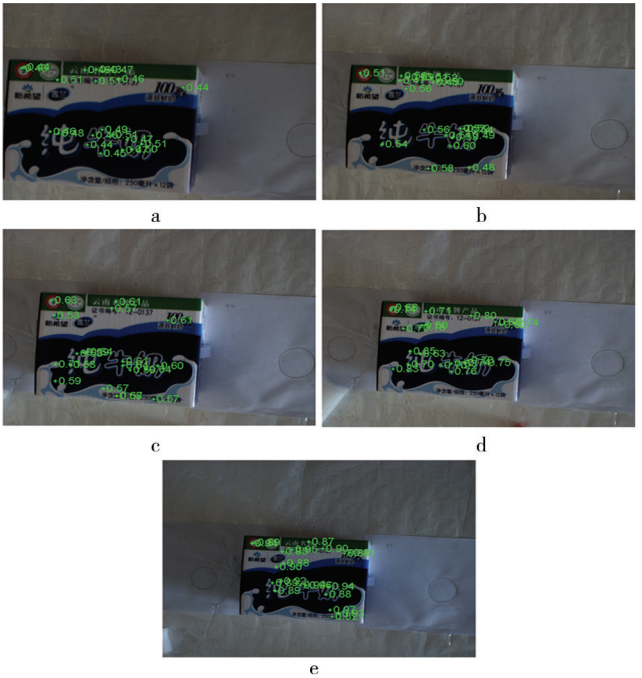


图5 稀疏深度
Fig.5 Images of sparse depth

由SURF特征点提取结果可知特征点均位于包装件上表面,因此为评估稀疏立体视觉算法应用于包装机器人测距的精度。文中计算每组提取的50个特征点深度值的平均值,将该值视为包装件与双目摄像机之间的距离,且已知包装件的高度为160 mm。5组实验中计算值和测量值之间的误差分析见表2。

表2 稀疏立体视觉深度提取误差
Tab.2 Error of depth estimation by sparse stereo

组类	深度估计值/mm	真实值/mm	误差/%
1	464.3	440	5.51 ± 0.23
2	568.8	540	5.33 ± 0.16
3	608.4	640	4.92 ± 0.18
4	752.9	740	1.74 ± 0.65
5	894.7	840	6.52 ± 0.23

5 结语

为解决包装机器人自动测距问题,文中提出将稀疏立体视觉算法应用到自动包装机器人,进行双目立体视觉测距研究。利用包装机器人配置的双目摄像机从不同范围内采集实验台上的包装件的图像,经过张正友摄像机标定和SURF特征点匹配等步骤获得包装件上表面特征点的深度值,用于指导机械手抓取的移动距离。实验结果表明,双目摄像机与实验台距离在600~1000 mm范围内,测量的包装件距离误差小于7%,满足一般情况下包装作业的要求。文中采用的匹配方法为稀疏的SURF特征点匹配,与SIFT算法相比速度快,但其稀疏性决定了基于SURF特征点的匹配算法不能提取足够多的图像特征。对于精度要求较高的包装作业,如包装件体积小或者结构复杂的情况下就需要较高的测距精度,因此,今后还需要在立体匹配方面做大量的研究工作和实际应用工作。

参考文献:

- [1] SCHARSTEIN D, SZELISKI R. High-accuracy Stereo Depth Maps Using Structured Light[C]// Proceedings of the 2003 IEEE Computer Society Conference Computer Vision and Pattern Recognition, 2003: 195—202.
- [2] LING Y, JIAHE Y. The 3D Surface Measurement and Simulation for Turbine Blade Surface Based on Color Encoding Structural Light[J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8(3): 273—280.
- [3] ZHANG L, GRIFT T E. A Monocular Vision-based Diameter Sensor for Miscanthus Giganteus[J]. Biosystems Engineering, 2012, 111(3): 298—304.
- [4] JIMÉNEZ A R, CERES R, PONS J L. A Vision System Based on A Laser Range-finder Applied to Robotic Fruit Harvesting[J]. Machine Vision and Applications, 2000, 11(6): 321—329.
- [5] 项荣, 应义斌, 蒋焕煜, 等. 基于双目立体视觉的番茄定位[J]. 农业工程学报, 2012, 28(5): 161—167.
- XIANG Rong, YING Yi-bin, JIANG Huan-yu, et al. The Localization of Tomatoes Based on Binocular Stereo Vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(5): 161—167.
- [6] 蒋焕煜, 彭永石, 应义斌, 等. 双目立体视觉技术在果蔬采摘机器人中的应用[J]. 江苏大学学报, 2008, 29(5): 377—380.
- JIANG Huan-yu, PENG Yong-shi, YING Yi-bin, et al. Binocular Stereo Vision Applied to Harvesting Robots[J]. Journal of Jiangsu University (Natural Science Edition), 2008, 29(5): 377—380.
- [7] XIANG R, HONG T, ZHOU M. Analysis of Depth Measurement Errors of Tomatoes Using Binocular Stereo Vision Based on Single Factor Experiments[C]// Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on, IEEE, 2014: 88—93.
- [8] RAKUN J, STAJNKO D, ZAZULA D. Detecting Fruits in Natural Scenes by Using Spatial-frequency Based Texture Analysis and Multiview Geometry[J]. Computers and Electronics in Agriculture, 2011, 76(1): 80—88.
- [9] XIANG R, YING Y, JIANG H, et al. Localization of Tomatoes Based on Binocular Stereo Vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(5): 161—167.
- [10] LIU Y, CAO X, DAI Q, et al. Continuous Depth Estimation for Multi-view Stereo[C]// IEEE Computer Vision and Pattern Recognition, 2009: 2121—2128.
- [11] CORKE P. Robotics, Vision and Control: Fundamental Algorithms in MATLAB[M]. Berlin Heidelberg: Springer Science & Business Media, 2011.
- [12] ZHANG Z. A Flexible New Technique for Camera Calibration[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000, 22(11): 1330—1334.
- [13] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up Robust Features (SURF)[J]. Computer Vision and Image Understanding, 2008, 110(3): 346—359.
- [14] DAVID G. Lowe, Distinctive Image Features from Scale-invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91—110.
- [15] HARTLEY R, ZISSERMAN A. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press, 2003.