

DOI:10.3969/j.issn.1001-4551.2019.05.020

# 基于 RGB-D 相机的球体三维重建及定位方法研究\*

冯 春, 花 省, 王树磊, 尹飞鸿, 江 炜  
(常州工学院 机械与车辆工程学院, 江苏 常州 213032)

**摘要:**针对单个和多个球体识别和定位问题,采用一种利用 RGB-D 相机进行三维重建的方法,对 RGB-D 相机的标定、深度图像和彩色图像的配准、彩色三维点云的生成等进行了研究。提出了一种阈值分割结合随机采样一致性算法的定位方法,利用色彩阈值分割和深度阈值分割方法实现了背景分割,得到了整体的球体区域的点云数据;利用欧几里得点云分割方法将整体的球体区域分割成了独立的球体点云,并辅以基于曲面法线的阈值分割方法分离相接触球体的点云;采用随机采样一致性算法分别计算出了每个球体的参数,得到了球体的球心坐标、半径以及深度信息。研究表明:在不同光照条件下该方法具有较好的实时性和定位精度,在 0.5 m ~ 1.5 m 的深度范围内,半径 30 mm 球体的半径平均测量误差为 4 mm,球心坐标在  $x, y, z$  方向的平均偏差为 1.3 mm、2.6 mm、1.4 mm,单帧图像处理平均时间约为 1 s。

**关键词:**RGB-D 相机;三维重建;识别;定位;点云分割

中图分类号:TP391

文献标志码:A

文章编号:1001-4551(2019)05-0549-06

## 3D reconstruction and positioning for spheres based on RGB-D camera

FENG Chun, HUA Xing, WANG Shu-lei, YIN Fei-hong, JIANG Wei

(Department of Mechanical and Vehicular Engineering, Changzhou Institute of Technology, Changzhou 213032, China)

**Abstract:** Aiming at single and multiple sphere recognition and localization problems, a method of three-dimensional reconstruction using an RGB-D camera was employed. The calibration of RGB-D camera, the registration of depth image and color image, and the generation of color 3D point cloud were studied. A positioning method based on threshold segmentation combined with random sampling consistency algorithm was proposed. The method of color threshold segmentation and the depth threshold segmentation method was used to realize the background segmentation, and the point cloud data of the whole sphere region was obtained. The Euclidean point cloud segmentation method was applied to divide the whole sphere region into independent sphere point clouds, and the threshold segmentation method based on surface normal was used for separating the point cloud of the contact sphere. The random sampling consistency algorithm was used to calculate the parameters of each sphere separately, and get the spherical coordinates, radius and depth information of the sphere. The results indicate that the method has good real-time and positioning accuracy under different illumination conditions. In the range of 0.5 m ~ 1.5 m depth, the average radius error of the sphere with a radius of 30 mm is 4 mm, and the spherical coordinates error at  $x, y$  and  $z$  direction is 2.3 mm, 4.5 mm, 3.7 mm in average. Meanwhile single frame image average processing time is about 1 s.

**Key words:** RGB-D camera; 3D reconstruction; recognition; positioning; point cloud segmentation

## 0 引 言

在工件的形状聚类<sup>[1]</sup>,以及苹果、西红柿等的采摘中<sup>[2-3]</sup>,对球体(或类球体)的识别和定位具有重要

的意义。近些年,利用 3D 重建方法对目标物进行识别和定位是一种有效的技术。3D 重建可以作为物体尺寸测量、形状识别、目标跟踪等的基础<sup>[4]</sup>,通常利用 3D 系统或者 3D 传感器来实现。立体视觉系统

收稿日期:2018-10-29

基金项目:江苏省自然科学基金青年项目(BK20140252, BK20150256)

作者简介:冯春(1981-),男,内蒙古赤峰人,博士,副教授,主要从事机器视觉、机器人技术方面的研究。E-mail:chun0471@163.com

常用的是双目立体视觉,它从不同角度进行拍摄,获取图像并对图像处理从而完成 3D 重建。由于图像处理涉及的计算量比较大,该方法通常涉及较为复杂的图像处理算法,实时性不强,同时彩色相机对外界光照的抗干扰性差。采用激光雷达传感器是 3D 重建的有效手段,但是存在价格昂贵且点云数据量大的问题。

RGB-D 相机不但能够获取彩色图像,同时可以得到深度图像信息,具有价格低、分辨率高等优势,因此,近几年获得了较为广泛的应用。利用 RGB-D 相机进行三维重建主要包含以下几个步骤:图像信息的获取及预处理,三维点云的生成、配准和拼接,点云数据的分割、对象检测和表面重建等。

IZADI S 等<sup>[5]</sup>实现了实时的 Kinect Fusion 三维重建系统,但是该算法不能适应过快的表面变化;NEW-COMBE R A 等<sup>[6]</sup>提出了改进的 Dynamic Fusion 三维重建系统,实现了对快速变形的有效反映;在 Kinect 进行三维重建中的关键步骤深度图像处理中,针对缺陷问题,FLEISHMAN S 等<sup>[7]</sup>提出了利用三维空间中的双边滤波,对曲面附近的噪声进行了平顺化,但是对离群噪声点效果有限;ZHENG Y 等<sup>[8]</sup>提出了一种基于一阶特征即小平面的法线的简单而有效的各向异性网络去噪框架,在设计新的双边滤波器的基础上实现了更好的去噪效果;点云分割方面,麦春艳等<sup>[9]</sup>利用随机采样一致性算法,简单高效地实时完成了苹果果实点云的分割与检测;除此之外,欧几里得点群提取算法、区域生产分割算法、凸凹边分割算法等也都有各自的应用。上述的 3D 重建技术均有各自的应用场合,但是它们实时性相对较差,处理单帧图像的时间在十几甚至二十几秒<sup>[10-12]</sup>。

针对球体等简单形状的物体,本文基于 Kinect v2 相机搭建实时精度较高的通用三维重建系统,在此基础上实现目标物的识别和定位。

## 1 3D 重建系统的硬件和软件

本文使用 Kinect V2 传感器获取环境的 RGB 图像信息和深度图像信息。借助 Kinect for Windows SDK 2.0 可以高效地从 Kinect V2 传感器中获取环境的彩色图像和深度图像。其中深度信息基于 TOF 技术,在 0.5 m ~ 4.5 m 视野内精度为 4 mm。

软件开发环境为 Visual Studio + Qt + OpenCV + PCL + Kinect for Windows SDK 2.0,采用 C++

编程语言。其中 Visual Studio 作为集成开发环境,Qt 用于交互界面开发,OpenCV 和 PCL 分别是二维图像和三维点云处理的 C++ 库,Kinect for Windows SDK 2.0 辅助从 Kinect V2 传感器获取环境图像。

## 2 球体的 3D 重建和定位算法研究

实现球体的识别和定位需要在进行三维重建的基础上完成球体的坐标和半径的计算,整体算法的流程图如图 1 所示。

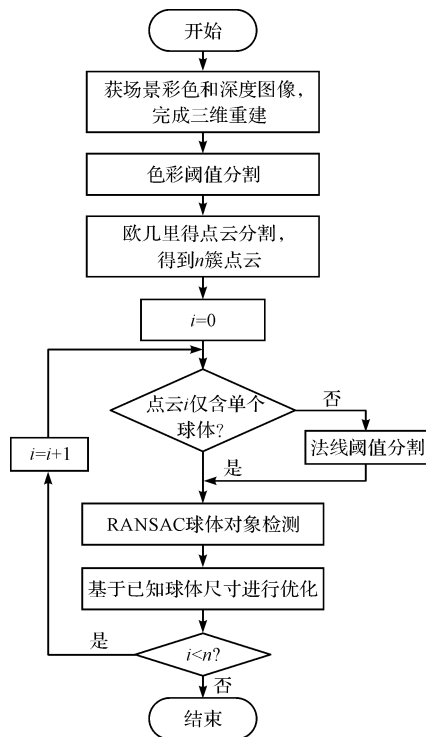


图 1 算法流程图

在利用 RGB-D 相机进行三维重建的过程中,首先需要获取彩色图像和深度图像信息,并进行彩色相机和深度相机的标定以及二者的配准;然后,本研究对目标物进行背景分割,采用颜色作为分割特征;最后,对单个及多个球体进行三维模型的确定以及位姿参数的确定。

### 2.1 球体图像的三维重建

球体图像的三维重建需要融合 Kinect 传感器的深度图像和彩色图像,得到场景的三维点云数据。在三维重建过程中最关键的两个步骤是完成相机标定和相机配准。

首先,本研究建立 kinect 相机的坐标系,如图 2 所示。

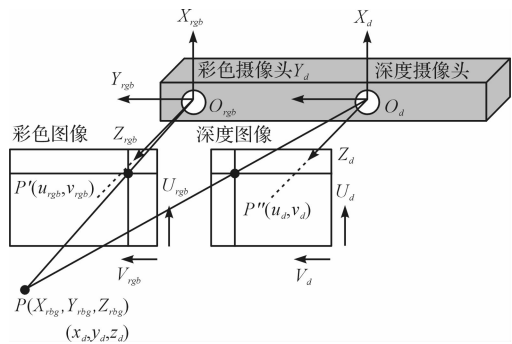


图2 Kinect相机坐标系

笔者利用相机标定技术,得到深度相机坐标系和深度图像像素坐标系之间的映射关系:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (1)$$

式中:  $(X_c, Y_c, Z_c)$ —空间一点在相机坐标系中坐标;  $(u, v)$ —该空间点对应像素坐标;  $f_u, f_v$ —相机归一化参数;  $s$ —相机内参数系数;  $(u_0, v_0)$ —光学中心坐标。

融合深度图像和彩色图像生成三维点云,需要确定深度图像像素坐标系与彩色图像像素坐标系之间的映射关系(即彩色摄像头与深度摄像头的配准),即:

$$Z_{rgb} \begin{bmatrix} u_{rgb} \\ v_{rgb} \\ 1 \end{bmatrix} = \mathbf{K}_{rgb} \mathbf{T}_{rgb-d} \mathbf{K}_d^{-1} Z_d \begin{bmatrix} u_d \\ v_d \\ 1 \end{bmatrix} \quad (2)$$

式中:  $Z_{rgb}, Z_d$ —空间一点分别在彩色和深度相机坐标系中的  $Z$  坐标;  $(u_{rgb}, v_{rgb}), (u_d, v_d)$ —该点在彩色和深度图像的像素坐标;  $\mathbf{T}_{rgb-d}$ —彩色相机向深度相机的位姿变换矩阵;  $\mathbf{K}_{rgb}, \mathbf{K}_d^{-1}$ —彩色和深度相机内参矩阵。

相机标定的算法有很多<sup>[13-14]</sup>,其中张正友标定法最为常用。本文在该方法的基础上,利用Kinectfor

$$d = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + \dots + (\Delta x_n)^2 + \dots + (\Delta x_k)^2} \quad (4)$$

式中:  $d$ —欧式距离;  $\Delta x_n (1 \leq n \leq k)$ —两点在第  $n$  维度上的距离。

由上式可知:在二维和三维空间中,欧式距离就是点的几何距离。该分割算法从点云中一点出发,计算其邻近点的欧式距离,当距离小于设定的阈值,就将其归入同一簇点云,并以新加入的点为基点进行运算,迭代直至不再有新的点加入点云。

### 2.4 基于法线的阈值分割

当两球体互相接触时,两球体的点云互相粘连,会

Windows SDK 2.0 中的函数为基础,分别得到深度相机的内部参数、彩色图像向深度图像的映射以及深度图像向相机坐标系的转换矩阵等参数,完成相机的标定和相机配准。

### 2.2 背景分割

原始的三维点云数据包含大量的背景点云,需要将目标球体的点云从背景中分割出来。色彩阈值分割是一种简便、高效的分割方法。根据目标物体与背景环境色彩的不同选择合适的色彩阈值(阈值的确定主要是通过实验测试的方法获取),可以快速地将背景与目标分割开。本文的目标球体为黄色,故选择了  $R-G, R-B$  分割方法如下式所示:

$$\begin{cases} R-B \geq a \text{ 且 } R-G \leq b \text{ 且 } Z \leq d, & \text{目标物体} \\ R-B < a \text{ 且 } R-G > b \text{ 且 } Z \leq d, & \text{背景} \end{cases} \quad (3)$$

式中:  $R, G, B$ —彩色通道灰度值;  $a, b, d$ —阈值;  $Z$ —目标深度。

由于彩色摄像头与深度摄像头视角不同,存在前景与背景的重叠与遮挡问题,融合彩色图像和深度图像时会错误的将前景色附着到背景上。尤其在前景与背景空间距离较远时,目标球体会在背景上产生大片投影,影响色彩阈值分割效果。本文在色彩阈值分割的基础上,增加了基于三维点云  $Z$  坐标的阈值分割方法,很好地解决了这一问题。

### 2.3 点云簇分割

场景中存在多个球体时,为了方便计算单个球体的位置和尺寸信息,需要将每个球体的点云从整体中分割出来。由于球体内点云间距小于球体间点云间距,可以采用基于欧几里得距离的点云分割方法。该分割方法在分割球体点云的同时,还有对离群噪声点进行滤波的作用。

欧几里得点云分割算法通过计算点与点之间的欧式距离判断其亲疏关系,  $k$  维空间的欧式距离为:

导致欧几里得点云分割方法无法有效分割两球体。因而在欧几里得点云分割之前,本研究使用基于表面法线的点云滤波方法去除两球体间粘连的点。

在相机坐标系中,球体的可视半球的边缘几乎与  $Z$  方向相切,即表面法线的单位向量的  $Z$  坐标值很小,因而可以使用  $Z$  坐标阈值分割方法切除半球体的边缘,阈值分割为:

$$e_n = (x_n, y_n, z_n) = n(p_0, K), \begin{cases} |z_n| \geq t, p_0 \in S_{in} \\ |z_n| < t, p_0 \in S_{ex} \end{cases} \quad (5)$$

式中： $e_n(x_n, y_n, z_n)$ ,  $n(p_0, K)$ —点云拟合曲面在  $p_0$  点的法线的单位向量； $(x_n, y_n, z_n)$ —点云坐标； $t$ —阈值； $K$ — $p_0$  点的临近点维度； $p_0$ —点云； $S_m, S_{ex}$ —球体内和球体外。

$e_n$  可通过  $p_0$  点及其  $K$  维邻近点计算得来，并且当  $e_n$  的  $Z$  坐标小于阈值  $t$  时，点  $p_0$  在球体点云的边缘，可以将其去除，进而切除两个接触球体的粘连部分。

## 2.5 球体参数估计

RANSAC 算法是一种经典的数据处理算法<sup>[15]</sup>。本文研究的球体的数学模型为：

$$(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 = r^2 \quad (6)$$

式中： $(x_0, y_0, z_0)$ —相机坐标系中的球心坐标； $r$ —球面半径。

笔者在分割得到单个球体的点云后，对点云进行 RANSAC 运算，得到该球体的模型参数。依次对每个球体点云进行运算，即可得到场景中所有球体的位置和尺寸。

对于已知半径或半径范围的球体，还可以对 RANSAC 计算结果进一步优化。可以在 RANSAC 计算结果超出范围的时候调整 RANSAC 参数迭代地计算，一般可以获得更好的结果，由此可提高系统鲁棒性和正确率。由于 RANSAC 得到的球面模型与实际球体的可视半球面拟合，计算结果中  $z$  坐标偏差与半径误差相关，即半径越大， $z$  坐标越大。所以对于半径已知的球体，可以根据半径计算结果调整  $z$  坐标，以减小误差，即：

$$z = z_c - (r_c - r_{\text{real}}) \quad (7)$$

式中： $z_c, r_c$ —RANSAC 计算得到的球心  $z$  坐标和球体半径； $r_{\text{real}}$ —球体实际半径。

## 3 实验结果和分析

### 3.1 球体图像获取和三维重建

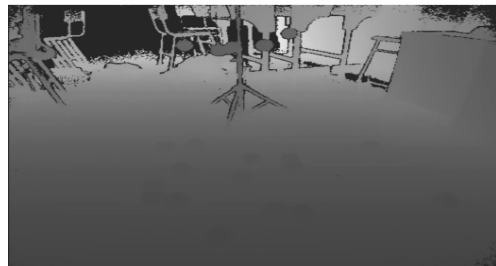
实验使用普通工控机作为软件系统运行平台，工控机采用 Intel(R) Core(TM) i3-4030U 处理器，主频 1.90 GHz。利用基于 Kinect V2 的 3D 重建系统，实时地获取场景图像如图 3 所示。

融合图 3 的彩色图像和深度图像生成具有色彩信息的三维点云数据如图 4 所示，完成场景的三维重建。

由图 4 可见：场景中有若干球体，分别置于地面和悬挂于支架上。



(a) 彩色图像



(b) 深度图像

图 3 Kinect 获取的图像

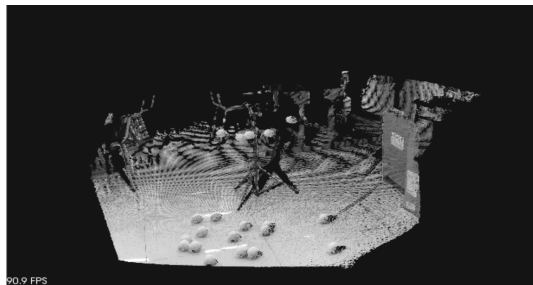


图 4 球体三维重建点云图

### 3.2 背景分割

本研究使用色彩阈值分割方法和距离阈值分割方法，对场景的三维点云进行背景分割。经过背景分割，无用的背景信息完全去除了，得到了目标球体区域的三维点云。

### 3.3 点云簇分割

为了快速地对众多目标球体进行分割，本研究使用欧几里得距离方法将单个球体从整体中分割出来，结果如图 5 所示。

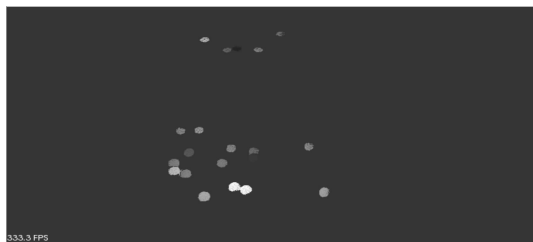


图 5 球体点云欧氏距离分割结果

由图 5 可知：各球体的点云很好地分割开来，球体点云边缘的离群噪声点也被去除，但对相接触的球体

分割效果不稳定。

### 3.4 基于法线的条件滤波

为了更好地分割相接触的球体,本研究使用基于法线的滤波方法,切断两球体点云的粘连部分,法线阈值分割结果如图6所示。

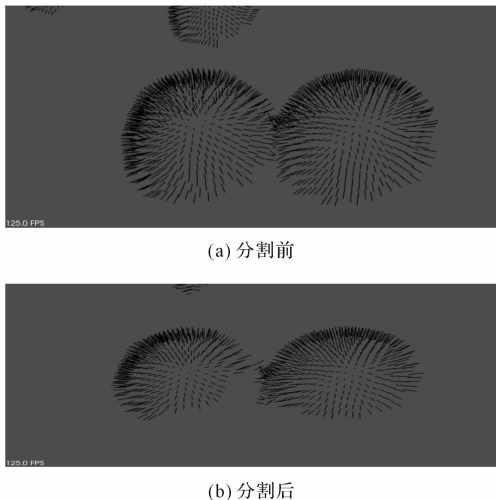


图6 法线阈值分割结果

由图6可知:球面上的各点的法线计算结果与球面一致,可见分割算法可以有效地切除两个球体粘连的点云。

基于欧几里得分割算法得到了更好的分割效果,如图7所示。

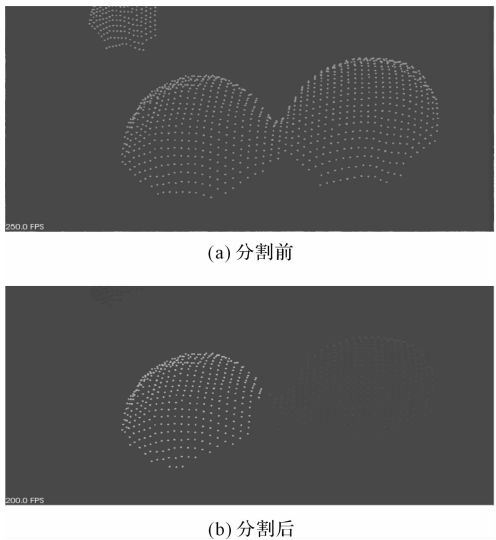


图7 欧几里得距离条件下的法线阈值分割结果

由图7可知:该算法与基于法线阈值分割的算法相比有更好的分割效果。

### 3.5 球体参数计算及误差分析

完成球体点云的分割之后,即可用RANSAC算法计算每个球体的参数。本文对处于不同深度的半径

30 mm标准球体进行了定位和尺寸测量,分析了不同条件下球体定位和尺寸测量的正确率、球心坐标误差和半径误差。实验结果如表(1~2)所示。

表1 球体半径计算结果平均误差(单位:mm)

光照条件	深度 $z(\Delta)$			
	0.5 m ~ 0.8 m	0.8 m ~ 1.0 m	1.0 m ~ 1.3 m	1.3 m ~ 1.5 m
自然光	3.08	3.29	3.85	4.82
室内日光灯	2.75	3.41	3.51	5.15
加光源	2.64	3.18	3.42	4.68

表2 球心坐标计算结果平均误差(单位:mm)

深度值	坐标 $(\Delta)$			
	$x$	$y$	$z$ 未优化	$z$ 优化后
0.5 m ~ 0.8 m	0.93	2.29	1.67	1.12
0.8 m ~ 1.0 m	1.47	3.63	4.10	1.57
1.0 m ~ 1.3 m	1.38	1.74	3.38	1.27
1.3 m ~ 1.5 m	1.17	1.70	2.58	1.38

由表1可知:本文方法在不同光照、不同深度条件下均有很好的表现,半径平均误差在4 mm左右。由表2可知:球心坐标计算误差小于4.1 mm之间,但是考虑到kinect V2深度相机的精度为4 mm,本文方法计算误差在可接受范围内。

限于篇幅本文仅给出自然光条件下的结果,其他光照条件下的偏差相差不大。由实验可知:在不同距离上,球心坐标误差很小。综合实验结果可知:本文所提出的方法在0.5 m ~ 1.3 m范围内球体检测识别的正确率约为99.5%,1.3 m ~ 1.5 m范围的正确率约为98%,需要说明的是“优化”对应2.5节公式(7)所对应的基于预知的球体半径信息对 $z$ 坐标采用的优化方法,深度值(即为 $z$ 坐标值)以及球心坐标均以深度相机坐标系为参考。

### 3.6 算法效率分析

通过实验可知:本文方法具有很高的实时性和准确性,仅从相机获取一对彩色图像和深度图像即可完成球体的定位和尺寸测量,精度高;检测10个30 mm半径的球体仅需0.6 s左右,检测20个30 mm球体需要1.4 s左右。

## 4 结束语

针对球体识别和定位问题,本文采用一种利用RGB-D相机进行三维重建的方法,对RGB-D相机的标定、深度图像和彩色图像的配准、彩色三维点云的生成等进行了研究。

该检测方法可以在多种光照、距离条件完成对多个球体的定位与尺寸测量,具有很高的实时性、准确

性、鲁棒性,仅从相机获取一帧图像即可在 1.4 s 内完成对 20 个球体的定位和半径测量,正确率 98% 以上,半径平均误差仅 4 mm,球心坐标误差仅 2 mm 左右。

本文提出的方法不但适用于球体和类球体的识别和定位,对于其他物体的 3D 重建和识别也具有借鉴价值。

### 参考文献 (References):

- [1] LASKEY M, LEE J, CHUCK C, et al. Robot grasping in clutter: Using a hierarchy of supervisors for learning from demonstrations[C]. IEEE International Conference on Automation Science and Engineering, Fort Worth: IEEE,2016.
- [2] SHARMA D J, KAUR J. Robotic apple harvesting using computer vision based on shape & colour analysis and object positioning[J]. **Digital Signal Processing**, 2014, 6 (7): 226-233.
- [3] KHOSHROO A, AREFI A, KHODAEI J. Detection of red tomato on plants using image processing techniques[J]. **Agricultural Communications**, 2014, 2(4): 9-15.
- [4] SERAFIN J, GRISETTI G. Using extended measurements and scene merging for efficient and robust point cloud registration[J]. **Robotics & Autonomous Systems**, 2017, 92 (6): 91-106.
- [5] IZADI S, KIM D, HILLIGES O, et al. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera[C]. Proceedings of the 24th annual ACM Symposium on User Interface Software and Technology, Santa Barbara: ACM, 2011.
- [6] NEWCOMBE R A, FOX D, SEITZ S M. DynamicFusion: reconstruction and tracking of non-rigid scenes in real-time [C]. 2015 IEEE Conference on Computer Vision Pattern Recognition, Boston: IEEE, 2015.
- [7] FLEISHMAN S, DRORI I, COHENOR D. Bilateral mesh denoising[J]. **Pro Siggraph**, 2003, 22(3): 950-953.
- [8] ZHENG Y, FU H, AU K C, et al. Bilateral normal filtering for mesh denoising[J]. **IEEE Trans Vis Comput Graph**, 2011, 17(10): 1521-1530.
- [9] 麦春艳,郑立华,孙红,等.基于 RGB-D 相机的果树三维重构与果实识别定位[J].农业机械学报,2015(s1): 35-40.
- [10] SVEIER A, KLEPPE A L, TINGELSTAD L, et al. Object detection in point clouds using conformal geometric algebra [J]. **Advances in Applied Clifford Algebras**, 2017, 27 (2): 1961-1976.
- [11] BAZAZIAN D, CASAS J R, RUIZ-HIDALGO J. Fast and robust edge extraction in unorganized point clouds[C]. International Conference on Digital Image Computing: Techniques and Applications, Adelaide: IEEE, 2016.
- [12] LIMBERGER F A, OLIVEIRA M M. Real-time detection of planar regions in unorganized point clouds[J]. **Pattern Recognition**, 2015, 48(6): 2043-2053.
- [13] ZHANG Z Y. A flexible new technique for camera calibration[J]. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2000, 22(11): 1330-1334.
- [14] KANG D J, LEE W H. Automatic circle pattern extraction and camera calibration using fast adaptive binarization and plane homography [J]. **International Journal of Precision Engineering & Manufacturing**, 2010, 11 (1): 13-21.
- [15] SCHNABEL, R, WAHL R, KLEIN, R. Efficient RANSAC for point-cloud shape detection[J]. **Computer Graphics Forum**, 2010, 26(2): 214-226.

[编辑:李 辉]

### 本文引用格式:

冯春,花省,王树磊,等.基于 RGB-D 相机的球体三维重建及定位方法研究[J].机电工程,2019,36(5): 549-554.

FENG Chun, HUA Xing, WANG Shu-lei, et al. 3D reconstruction and positioning for spheres based on RGB-D camera[J]. Journal of Mechanical & Electrical Engineering, 2019, 36(5): 549-554.

《机电工程》杂志: <http://www.meem.com.cn>