

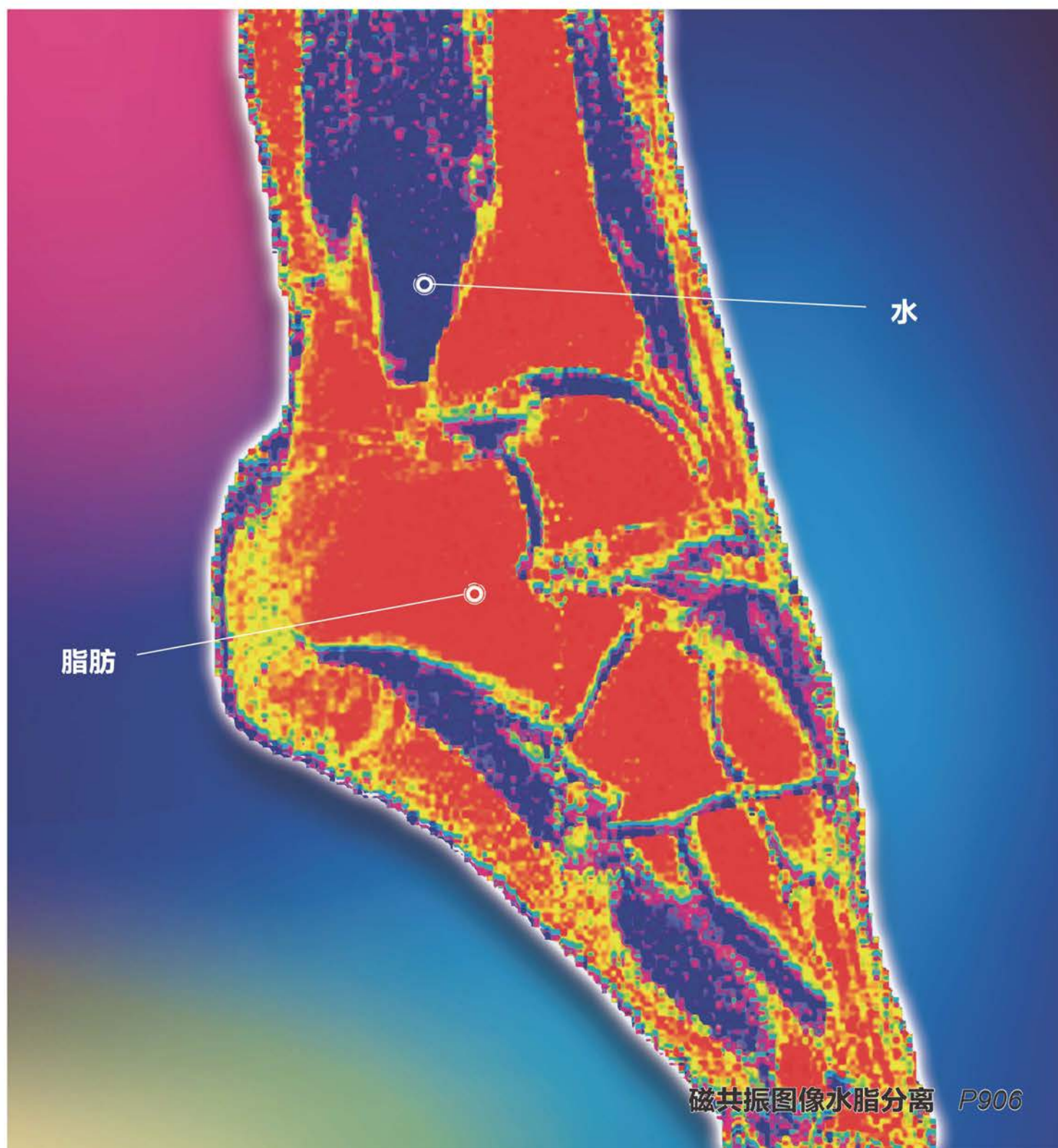


主办: 中国科学院遥感与数字地球研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所

# 中国图象图形学报

2018  
06  
VOL.23

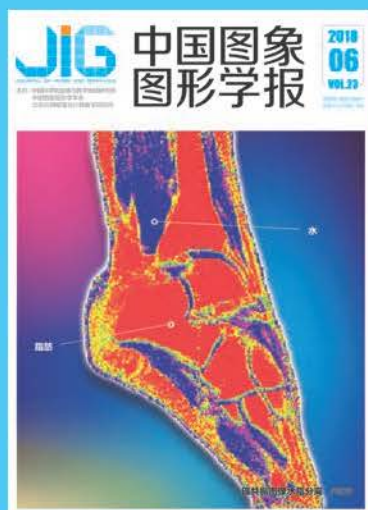
ISSN1006-8961  
CN11-3758/TB



磁共振图像水脂分离 P906

# 中国图象图形学报

刊名题字：宋健 月刊（1996年创刊）



第23卷第6期（总第266期）  
2018年6月16日

中国精品科技期刊  
中国国际影响力优秀学术期刊  
中国科技核心期刊  
中文核心期刊

## 版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

## Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院  
主办单位 中国科学院遥感与数字地球研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所

主 编 顾行发  
编辑出版《中国图象图形学报》编辑出版委员会  
邮政信箱 北京9718信箱  
邮 编 100101  
电子信箱 jig@radi.ac.cn  
电 话 010-64807995  
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号  
总 发 行 北京报刊发行局  
订 购 全国各地邮局  
海外发行 中国国际图书贸易集团有限公司  
(邮政信箱: 北京399信箱 邮编: 100048)  
印刷装订 北京科信印刷有限公司

## Journal of Image and Graphics

Title inscription: Song Jian Monthly, Started in 1996

Superintended by Chinese Academy of Sciences  
Sponsored by Institute of Remote Sensing and Digital Earth, CAS  
China Society of Image and Graphics  
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa  
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics  
P.O.Box 9718, Beijing, P.R.China  
Zip code 100101  
E-mail jig@radi.ac.cn  
Telephone 010-64807995  
Website www.cjig.cn

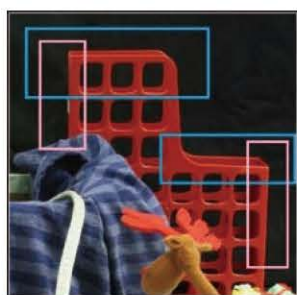
Distributed by Beijing Bureau for Distribution of Newspapers and Journals  
Domestic All Local Post Offices in China  
Overseas China International Book Trading Corporation  
(P.O.Box 399, Beijing 100048, P.R.China)  
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB  
ISSN 1006-8961  
CODEN ZTTXFZ  
国外发行代号 M1406  
国内邮发代号 82-831  
国内定价 60.00元





复合域的显著性目标检测方法(第0846页)



特征融合的双目半全局匹配算法及其并行加速实现(第0874)



结合拓扑持续性和热扩散理论的3维模型分割(第0887页)

## 综述

### 数字视频区域篡改的检测与定位

姚晔, 胡伟通, 任一, 翁韶伟 ..... 0779

## 图像处理和编码

### 联合局部约束的邻域嵌入人脸超分辨率重建

黄福珍, 周晨旭, 何林巍 ..... 0792

### 亮度一梯度联合约束的车牌图像超分辨率重建

孙京, 袁强强, 李冀玮, 周春平, 沈焕锋 ..... 0802

### 位级同步置乱扩散和像素级环形扩散图像加密算法

梁颖, 张绍武 ..... 0814

## 图像分析和识别

### 多特征融合与交替方向乘子法的行人再识别

齐美彬, 王慈淳, 蒋建国, 李倩 ..... 0827

### 区域建议网络的细粒度车型识别

杨娟, 曹浩宇, 汪荣贵, 薛丽霞, 胡敏 ..... 0837

### 复合域的显著性目标检测方法

崔丽群, 赵越, 胡志毅, 赵雨康 ..... 0846

### 背景吸收的马尔可夫显著性目标检测

蒋峰岭, 张海涛, 杨静, 孔斌 ..... 0857

## 图像理解和计算机视觉

### 点云下地平面检测的RGB-D相机外参自动标定

孙士杰, 宋焕生, 张朝阳, 张文涛, 王璇 ..... 0866

### 特征融合的双目半全局匹配算法及其并行加速实现

吕倪祺, 宋广华, 杨波威 ..... 0874

## 计算机图形学

### 结合拓扑持续性和热扩散理论的3维模型分割

杨军, 张鹏 ..... 0887

### 带参数的同次Ball曲线

李军成, 李兵, 易叶青 ..... 0896

## 医学图像处理

### 相位分区和局部多项式曲面拟合的相位解缠绕

程军营, 王常青, 冯衍秋, 陈武凡 ..... 0906

### 跨模态多标签生物医学图像分类建模识别

于玉海, 林鸿飞, 孟佳娜, 郭海, 赵哲焕 ..... 0917

## 遥感图像处理

### 面向SAR目标识别的深度卷积神经网络结构设计

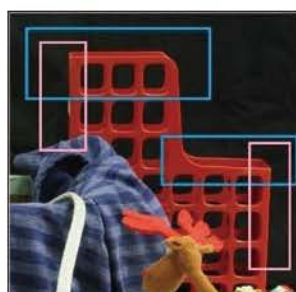
谷雨, 徐英 ..... 0928

# CONTENTS

## JOURNAL OF IMAGE AND GRAPHICS



Saliency object detection method based on complex domains (P0846)



Semi-global stereo matching algorithm based on feature fusion and its CUDA implementation(P0874)



Three-dimensional shape segmentation by combining topological persistence and heat diffusion theory(P0887)

### Medical Image Processing

Detection and localization of digital video regional tampering

Yao Ye, Hu Weitong, Ren Yizhi, Weng Shaowei ..... 0779

### Image Processing and Coding

Face super-resolution reconstruction method based on joint local constraint neighbor embedding

Huang Fuzhen, Zhou Chenxu, He Linwei ..... 0792

License plate image super-resolution based on intensity-gradient prior combination

Sun Jing, Yuan Qiangqiang, Li Jiwei, Zhou Chunping, Shen Huanfeng ..... 0802

Image encryption algorithm based on bit-level synchronous permutation diffusion and pixel-level annular diffusion

Liang Ying, Zhang Shaowu ..... 0814

### Image Analysis and Recognition

Person re-identification based on multi-feature fusion and alternating direction method of multipliers

Qi Meibin, Wang Cichun, Jiang Jianguo, Li Ji ..... 0827

Fine-grained car recognition method based on region proposal networks

Yang Juan, Cao Haoyu, Wang Ronggui, Xue Lixia, Hu Min ..... 0837

Saliency object detection method based on complex domains

Cui Liqun, Zhao Yue, Hu Zhiyi, Zhao Yukang ..... 0846

Image saliency detection based on background-absorbing Markov chain

Jiang Fengling, Zhang Haitao, Yang Jing, Kong Bin ..... 0857

### Image Understanding and Computer Vision

Automatic extrinsic calibration for RGB-D camera based on ground plane detection in point cloud

Sun Shijie, Song Huansheng, Zhang Chaoyang, Zhang Wentao, Wang Xuan ..... 0866

Semi-global stereo matching algorithm based on feature fusion and its CUDA implementation

Lyu Niqui, Song Guanghua, Yang Bowei ..... 0874

### Computer Graphics

Three-dimensional shape segmentation by combining topological persistence and heat diffusion theory

Yang Jun, Zhang Peng ..... 0887

Ball curve of the same degree with a parameter

Li Juncheng, Li Bing, Yi Yeqing ..... 0896

### Medical Image Processing

New phase-unwrapping method based on phase partition and local polynomial surface fitting

Cheng Junying, Wang Changqing, Feng Yanqiu, Chen Wufan ..... 0906

Classification modeling and recognition for cross modal and multi-label biomedical image

Yu Yuhai, Lin Hongfei, Meng Jiana, Guo Hai, Zhao Zhehuan ..... 0917

### Remote Sensing Image Processing

Architecture design of deep convolutional neural network for SAR target recognition

Gu Yu, Xu Ying ..... 0928

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2018)06-0866-08

论文引用格式: Sun S J, Song H S, Zhang C Y, Zhang W T, Wang X. Automatic extrinsic calibration for RGB-D camera based on ground plane detection in point cloud[J]. Journal of Image and Graphics, 2018, 23(6): 0866-0873. [孙士杰, 宋焕生, 张朝阳, 张文涛, 王璇. 点云下地平面检测的 RGB-D 相机外参自动标定[J]. 中国图象图形学报, 2018, 23(6): 0866-0873. ][DOI:10.11834/jig.170490]

## 点云下地平面检测的 RGB-D 相机外参自动标定

孙士杰, 宋焕生, 张朝阳, 张文涛, 王璇

长安大学, 西安 710064

**摘要:** 目的 RGB-D 相机的外参数可以被用来将相机坐标系下的点云转换到世界坐标系的点云, 可以应用在 3 维场景重建、3 维测量、机器人、目标检测等领域。一般的标定方法利用标定物(比如棋盘)对 RGB-D 彩色相机的外参标定, 但并未利用深度信息, 故很难简化标定过程, 因此, 若充分利用深度信息, 则极大地简化外参标定的流程。基于彩色图的标定方法, 其标定的对象是深度传感器, 然而, RGB-D 相机大部分则应用基于深度传感器, 而基于深度信息的标定方法则可以直接标定深度传感器的姿势。**方法** 首先将深度图转化为相机坐标系下的 3 维点云, 利用 MELSAC 方法自动检测 3 维点云中的平面, 根据地平面与世界坐标系的约束关系, 遍历并筛选平面, 直至得到地平面, 利用地平面与相机坐标系的空间关系, 最终计算出相机的外参数, 即相机坐标系内的点与世界坐标系内的点的转换矩阵。**结果** 实验以棋盘的外参标定方法为基准, 处理从 PrimeSense 相机所采集的 RGB-D 视频流, 结果表明, 外参标定平均侧倾角误差为  $-1.14^\circ$ , 平均俯仰角误差为  $4.57^\circ$ , 平均相机高度误差为 3.96 cm。**结论** 该方法通过自动检测地平面, 准确估计出相机的外参数, 具有很强的自动化, 此外, 算法具有较高地并行性, 进行并行优化后, 具有实时性, 可应用于自动估计机器人姿势。

**关键词:** RGB-D 相机; 自动标定; 3 维点云; 平面检测; 深度图

### Automatic extrinsic calibration for RGB-D camera based on ground plane detection in point cloud

Sun Shijie, Song Huansheng, Zhang Chaoyang, Zhang Wentao, Wang Xuan

Chang'an University, Xi'an 710064, China

**Abstract: Objective** The extrinsic parameter of the RGB-D camera can be used to convert point cloud in camera coordinate to that in world coordinate. This parameter can be applied to 3D reconstruction, 3D measurement, robot gesture estimation, and target detection, among others. The RGB-D camera (e.g., Kinect, PrimeSense, and RealSense) consists of two sensors: RGB sensor and depth sensor. The former sensor retrieves the RGB image, whereas the latter retrieves depth image from the scene. To translate the 3D point cloud in the camera coordinate to the world coordinate, the extrinsic parameters of depth sensor have to be calibrated. The general calibration methods use calibration objects (such as chessboard) to obtain

收稿日期: 2017-09-08; 修回日期: 2018-01-06; 预印本日期: 2018-01-13

基金项目: 国家自然科学基金项目(61572083); 中央高校基本科研业务费专项基金项目(310824171003); 陕西省自然科学基金项目(2017JQ6064)

第一作者简介: 孙士杰(1989—), 男, 长安大学交通信息工程及控制专业博士研究生, 主要研究方向为 3 维点云目标识别、机器视觉、跟踪、人工智能。E-mail: shijieSun@chd.edu.cn

**Supported by:** National Natural Science Foundation of China (61572083); Fundamental Research Funds for the Central Universities (310824171003); Natural Science Foundation of Shaanxi Province, China (2017JQ6064)



the extrinsic parameter of the RGB-D color sensor, which is regarded as the extrinsic parameter of the depth sensor approximately. These methods do not make full use of depth information, thereby causing difficulty in simplifying the calibration process. Moreover, lack of knowledge on the difference between the depth sensor and color sensor can cause large errors. Thus, to accurately estimate the extrinsic parameter of the depth sensor in the RGB-D camera, some methods have been proposed by using the extrinsic parameters of depth sensor relative to the color sensor. However, these methods complicate the calibration process. To simplify the calibration process of the extrinsic parameter of the depth sensor, the depth information should be fully utilized. Results of the methods are based on the color image with the parameter of the color sensor. However, the majority of applications on the RGB-D camera are based on the depth sensor. Moreover, parameters of the depth sensor should be directly calibrated. **Method** We build the spatial constraint relation between the ground plane and the camera, which can be used to select the ground plane from planes detected in the 3D point cloud. The ground plane should satisfy the following conditions: 1) The angle between the z axis of the camera and the ground plane is less than the specified threshold. 2) The z value of the ground plane in the world coordinate is larger than that of the other points, which are not in the ground plane. Moreover, we create the world coordinate based on the detected ground planes automatically. The origin point of the world coordinate is the projection of the origin point of the camera coordinate to the plane and the y axis of the world coordinate is the projection of the z axis of the camera coordinate to the plane. In addition, the direction of the z axis of the world coordinate is toward the origin point of the world coordinate from the origin point of the camera coordinate. We calibrate the extrinsic parameter of the RGB-D camera in the following steps. First, we reconstruct the 3D point cloud from the depth image retrieved from the depth sensor of the RGB-D camera. The reconstructed 3D point cloud is in the camera coordinate, whose subset forms a large number of planes. Second, planes in the 3D point cloud are detected by the MELSAC method. At most, one ground plane exists in the detected planes. Third, the spatial constraint rule between the ground plane and camera is built. The detected planes are filtered by the spatial constraint rule until the ground plane is found or all the planes are iterated. The process stops when a ground plane cannot be found. Finally, by using the relation between the ground plane and the camera, point sets are selected to calculate the extrinsic parameters. **Result** In the experiment, the benchmark is the result of checkerboard extrinsic calibration method only processing the RGB image of RGB-D information which is retrieved from the PrimeSense camera. We record an 89.4 s video and use it in the experiment. The videos contain two sub-videos: RGB video whose channel is 3 and depth video whose channel is 1. A 7x7 checkerboard is found in every frame of the RGB video, which is processed by the checkerboard-based method. The input of our proposed method is the frame of the depth video. The result shows that the average tilt angle error is  $-1.14^\circ$ , the average pitch angle error is  $4.57^\circ$ , and the average camera height error is 3.96 cm. An experiment to test the robustness of the noise is also performed. The variance of the Gaussian noise in the depth frame is increased, and the result of each variance Gaussian noise is obtained. The stability of calibration decreases with the increase in the variance of Gaussian noise. The result shows that our method performs effectively when the variance of the Gaussian noise is below 0.01. **Conclusion** Our proposed method fully utilizes the depth information of the RGB-D camera, and simplifies the process of extrinsic calibration of the depth sensor. Thus, our method can be used in actual application. For convenience, the source code is also published. This method can automatically detect the ground plane and does not require other calibration objects. Therefore, the proposed method can calibrate each frame of the recorded video accurately, and it is not sensitive to the noise in the depth image. In addition, the algorithm has high parallelism. The process of estimating planes in the 3D point cloud and filtering these planes can be implemented in a parallel manner. The proposed method will have real-time performance based on this parallel implementation.

**Key words:** RGB-D camera; automatic extrinsic calibration; 3D point cloud; plane detection; depth map

## 0 引言

RGB-D 相机, 比如 Kinect<sup>[13]</sup>、PrimeSense、Asus

Xtion Pro 等,是在传统的 RGB 相机的基础上增加了红外相机,图1是 Kinect 的结构图,可见,它内部包含:彩色相机、红外相机和红外照明器;彩色相机输出彩色图,红外相机输出深度图,红外照明器发射红外

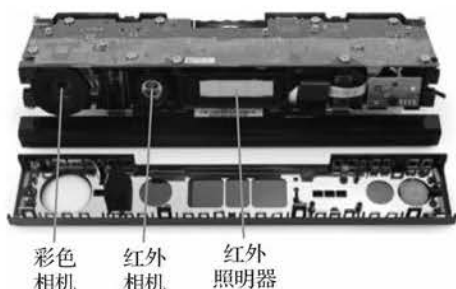


图1 Kinect 的结构

Fig. 1 Structure of the Kinect

光,用于深度图的计算.随着 Kinect 等低成本的 RGB-D 相机的出现,它们越来越多地被应用在 3 维场景重建和导航<sup>[45]</sup>、目标识别和跟踪<sup>[6-7]</sup>、3 维测量等应用中.这些应用往往需要对相机的外参进行标定.

例如在目标检测和识别应用<sup>[7]</sup>中,首先应标定出相机的外参,然后对场景中的行人进行跟踪.如果能够获得相机的外参,可以消除透视带来的负面影响,统一检测算法,简化识别过程,提高识别速度.

但是,目前针对 RGB-D 相机的标定,往往研究的是彩色相机相对于红外相机的外参标定<sup>[8-9]</sup>.针对 RGB-D 相机本身外参往往近似与彩色相机的标定,彩色相机标定一般采用棋盘标定的方法<sup>[10]</sup>,Munaro 等人<sup>[6]</sup>利用棋盘对多个相机外参进行标定,然后在标定的基础上进行行人检测.这种标定方法一方面没有充分利用 RGB-D 相机所提供的深度信息,另一方面标定的结果本质上是彩色相机的外参数,如果直接用于深度图 3 维重建,会有较大误差. Shibo 等人<sup>[11]</sup>设计一种专门针对 RGB-D 红外相机可以识别的标定板,该标定板上有一定间隔的孔,通过自动识别孔来对红外相机进行标定.虽然这是针对红外相机进行的标定,但是需要设计专用的标定物,增大了标定的难度. Liao 等人<sup>[8]</sup>将标定划分为 3 类:1)需要标定物的标定,2)需要人为干预的标定,3)全自动的标定;按照这个分类,本文提出的方法属于其中的第 3 类.

参照 Liao 等人<sup>[8]</sup>所提出的标定分类,本文将 RGB-D 相机外参标定的方法划分为 3 类:第 1 类标定出彩色相机的外参,然后利用彩色相机和红外相机的外参  $T_c^0$ ,通过变换  $T_0 = T_c^0 T_c$ ,得到红外相机外参  $T_0$ .该类方法的优点在于可以直接利用彩色图的标定方法,但是缺点是需要彩色相机相对于红外相机的外参数,而且没有利用深度图所提供的信

息,标定过程复杂.第 2 类标定方法是通过设计特定标定物,在红外相机所提供的深度图上进行特征检测,通过特征(比如棋盘角点)得到相机的外参数.这种类型的方法虽然是对红外相机进行的标定,但是,需要设计特定的标定物,相对第 1 类标定方法而言,并没有简化标定过程.第 3 类使用深度图对外参相机进行标定,通过检测深度图上的目标,利用目标的与世界坐标系之间关系,进行标定,这类标定方法直接利用深度信息,可以大大简化标定的过程,提高标定的效率.本文所提方法——基于地平面检测相机外参自动标定方法,以下简称地平面标定方法,属于第三类.地平面标定方法能够利用 RGB-D 相机的深度信息,直接对 RGB-D 相机的红外相机进行标定.

## 1 外参标定过程

RGB-D 相机包含彩色相机和红外相机,本文主要针对红外相机的外参进行标定,后文出现的外参标定默认指红外相机的外参标定.相机的外参为

$$T_c^w = \begin{bmatrix} R_c^w & 0 \\ t_c^w & 1 \end{bmatrix} \quad (1)$$

式中,  $T_c^w \in \mathbb{R}^{4 \times 4}$ ,  $R_c^w \in \mathbb{R}^{3 \times 3}$  为相机的旋转矩阵,  $t_c^w \in \mathbb{R}^{1 \times 3}$  为相机的平移矩阵,  $(\cdot)_c^w$  表示由相机坐标系转换到世界坐标系.

如图 2 是外参标定的主要流程.建立世界坐标系后,从 RGB-D 相机获得深度图像;利用红外相机的内参,将深度图转换称相机坐标系下的 3 维点云;最后利用迭代求解点云中的地平面,在得到地平面的同时计算出相机的外参.

### 1.1 相机坐标系和世界坐标系的建立

相机坐标是以红外相机为原点所建立 3 维坐标系,如图 3,红外相机为相机坐标系原点  $O_c$ ,  $X_c$  轴沿着 Kinect 的横向方向,  $Z_c$  轴垂直于 Kinect,指向拍摄的方向.

通常世界坐标可以任意建立,但是为了方便计算相机的外参,如图 4 所示,在建立世界坐标的时候,应满足下面要求:

- 1) 世界坐标系的原点为相机坐标系原点在地平面的投影点;
- 2) 世界坐标系的  $Y_w$  轴为相机坐标系  $Z_c$  轴在地

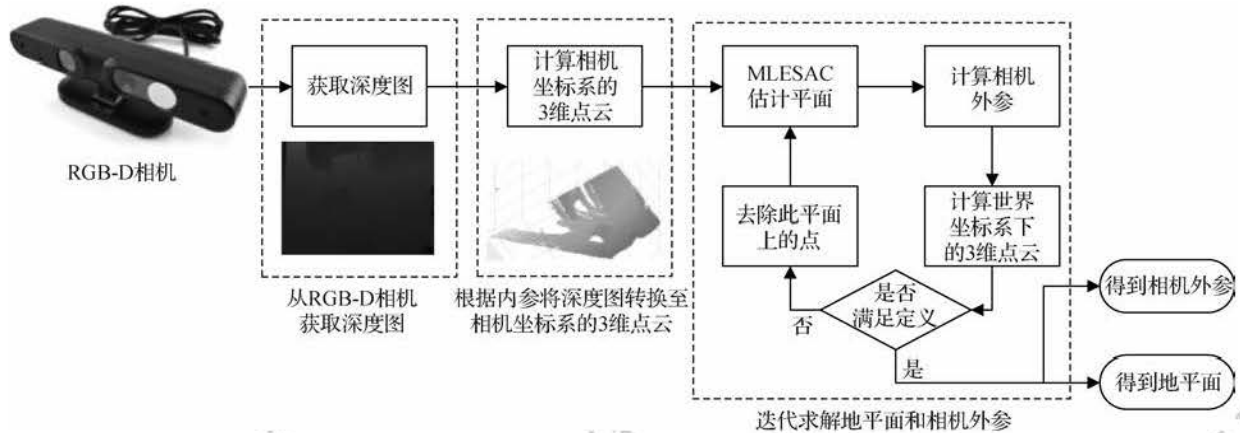


图2 相机外参标定流程

Fig. 2 Process of camera extrinsic parameter calibration



图3 相机坐标系

Fig. 3 Camera coordinate

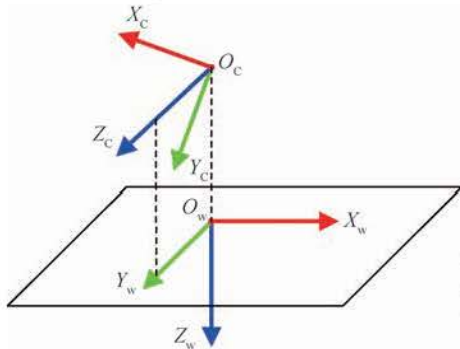


图4 世界坐标系的位置

Fig. 4 The position of world coordinate

面的投影;

3) 世界坐标系的  $Z_w$  轴垂直于地面向下;

4) 建立的为右手坐标系。

以这种方式建立世界坐标系, 可以较为方便地寻找相机坐标与世界坐标对应的点组, 简化标定过程。

## 1.2 相机外参计算

相机坐标系到世界坐标系的变换矩阵, 即相机

$$T_w^c = \begin{bmatrix} \frac{(P_c^{(1)} - (P_c^{(0)})^T)}{\|(P_c^{(1)} - (P_c^{(0)})^T)\|_2} & \frac{(P_c^{(2)} - (P_c^{(0)})^T)}{\|(P_c^{(2)} - (P_c^{(0)})^T)\|_2} & \frac{(P_c^{(3)} - (P_c^{(0)})^T)}{\|(P_c^{(3)} - (P_c^{(0)})^T)\|_2} & 0 \\ P_c^{(0)} & & & 1 \end{bmatrix} \quad (4)$$

的外参  $T_w^c$ , 为  $4 \times 4$  的矩阵, 有 12 个未知参数, 见式(1)。首先计算世界坐标系到相机坐标系的变换矩阵  $T_w^c$ , 得到相机的外参, 即

$$T_c^w = T_w^c^{-1} \quad (2)$$

记相机坐标系的点为  $P_c^{(i)}$ , 与之对应的世界坐标系的点为  $P_w^{(i)}$ , 如图 5, 在世界坐标系下选择 4 个特殊的点, 即

$$\begin{cases} P_w^{(0)} = [000] \\ P_w^{(1)} = [100] \\ P_w^{(2)} = [010] \\ P_w^{(3)} = [001] \end{cases} \quad (3)$$

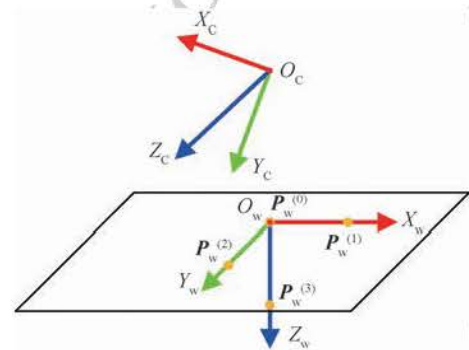


图5 在世界坐标系下选中的点

Fig. 5 Selected points in world coordinate

通过计算对应的相机坐标系下的点, 即  $P_c^{(0)}$ ,

便可求出  $T_w^c$ , 即



已知相机坐标系下的平面方程:  $ax + by + cz + d = 0$ , 其垂直于地面向下的法向量为:  $\mathbf{n}_c = [abc]$  且  $\|\mathbf{n}_c\|_2 = 1$ 。由世界坐标系原点为相机坐标系原点在平面上的投影, 可知世界坐标的原点  $\mathbf{P}_w^{(0)}$  在相机坐标系下的坐标为

$$\mathbf{P}_c^{(0)} = \frac{|d|}{\sqrt{a^2 + b^2 + c^2}} \mathbf{n}_c \quad (5)$$

点  $\mathbf{P}_w^{(3)} = [001]$  在世界坐标系  $Z$  轴上, 由  $O_c, \mathbf{P}_c^{(0)}, \mathbf{P}_c^{(3)}$  共线, 那么有

$$\mathbf{P}_c^{(3)} = \mathbf{P}_c^{(0)} + \mathbf{n}_c \quad (6)$$

因  $Y_w$  轴为  $Z_c$  轴在平面上的投影, 故

$$\mathbf{P}_c^{(2)} = \mathbf{P}_c^{(0)} + \mathbf{n}_y \quad (7)$$

式中,  $\mathbf{n}_y$  为  $\mathbf{P}_c^{(0)}\mathbf{P}_c^{(2)}$  的单位向量, 有

$$\mathbf{n}_y = \frac{\mathbf{P}_{2c} - \mathbf{P}_c^{(0)}}{\|\mathbf{P}_{2c} - \mathbf{P}_c^{(0)}\|_2} \quad (8)$$

$\mathbf{P}_{2c}$  为相机坐标系的  $[001]$  点到平面的投影。通过解方程, 可得投影点。

$$\begin{cases} ax + by + cz + d = 0 \\ \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \end{cases} \quad (9)$$

式中,  $a, b, c, d$  为平面方程的参数,  $(x_0, y_0, z_0)$  为平面上任一点。

当求出  $\mathbf{P}_c^{(0)}, \mathbf{P}_c^{(2)}, \mathbf{P}_c^{(3)}$  后, 通过向量叉乘可以求得  $\mathbf{P}_c^{(1)}$ , 即

$$\mathbf{P}_c^{(1)} = \frac{\mathbf{P}_c^{(2)} - \mathbf{P}_c^{(0)} \times \mathbf{P}_c^{(3)} - \mathbf{P}_c^{(0)}}{\|\mathbf{P}_c^{(2)} - \mathbf{P}_c^{(0)} \times \mathbf{P}_c^{(3)} - \mathbf{P}_c^{(0)}\|_2} \quad (10)$$

通过式(5)一式(7), 式(10), 求出从世界坐标系下选出的4个点在相机坐标系下的坐标, 然后利用式(4)得到世界坐标点到相机坐标点的转换矩阵, 最后利用式(2)的到红外相机的外参矩阵。

## 2 地平面估计过程

为了利于平面检测, 首先利用内参将深度图转化为相机坐标系下的3维点云, 利用 MLESAC (maximum likelihood estimation sample consensus) 方法<sup>[12]</sup>提取平面, 迭代遍历平面, 求出外参, 利用外参和点云来判断平面是否为地平面, 最终得到相机外参, 见图7。

### 2.1 将深度图转换为相机坐标系下的3维点云

如图6, 是相机的成像模型, 像平面的坐标轴

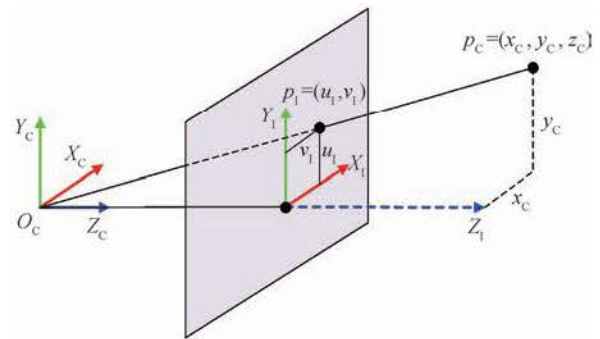


图6 相机的成像模型

Fig. 6 The model of camera

分别用  $X_1$  和  $Y_1$  来表示, 相机坐标系下的点  $P_c = (X_c, Y_c, Z_c)$  在像平面上的成像点为  $P_1 = (u_1, v_1)$ , 相机的焦距为  $f$ , 光轴与像平面交点坐标为  $c_x$  和  $c_y$ , 像素的长宽分别为:  $d_x$  和  $d_y$ , 像素值为  $d_p$ , 像素值在相机坐标系下对应的点的  $Z$  坐标值的比例为  $s$ 。那么可将  $(u, v, d_p)$  转换为  $(X_c, Y_c, Z_c)$ , 计算公式为

$$\begin{cases} f_x = \frac{f}{d_x} \\ f_y = \frac{f}{d_y} \\ Z_c = d_p * s \\ X_c = \frac{(u_1 - c_x) \cdot Z_c}{f_x} \\ Y_c = \frac{(v_1 - c_y) \cdot Z_c}{f_y} \end{cases} \quad (11)$$

### 2.2 地平面估计

地平面估计是后序相机外参计算的基础, 它的准确度, 决定了相机外参的准确度。在一个场景中, 会拟合出多个平面, 那么通过下面条件来判断平面是否为地平面

$$\begin{cases} \theta < \frac{\pi}{4} \\ \text{median } P_{w1}^2 > \min P_{c0}^2 + \varepsilon \end{cases} \quad (12)$$

式中,  $\theta$  为相机的  $X_c$  轴与平面的夹角,  $\text{median}$  是取中值的操作,  $P_{w1}^2$  为拟合平面模型的内点  $Z$  值集合,  $P_{c0}^2$  为拟合平面模型的外点  $Z$  值集合,  $\varepsilon$  为设定的容差值。式(12)表示两个条件: 1) 相机相对于平面的倾斜角不超过  $45^\circ$ ; 2) 根据平面计算出外参后, 将点云转换至世界坐标系下, 平面上的点集的  $Z$  值最大。

根据该定义, 本文通过筛选平面, 来计算符合条件的地平面, 本文方法的流程如图7所示。

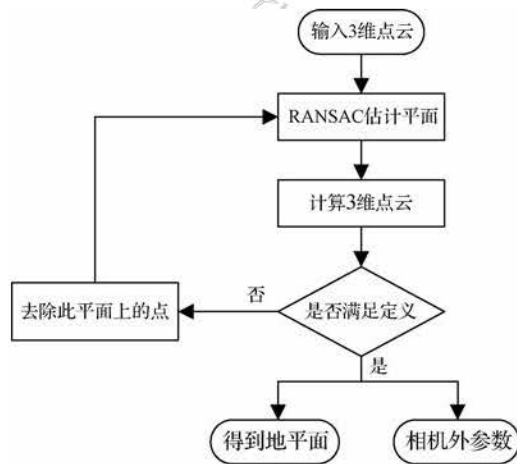


图7 地平面估计的流程图

Fig.7 The process of ground plane estimation

首先输入相机坐标系下的3维点云, 利用MLESAC方法, 计算平面, 并记录满足该平面的内点集合, 利用此平面和2.2节的方法, 计算出相机的外参数, 结合相机的内参数, 将相机坐标系下的3维点云转换为世界坐标系下的3维点云, 判断此时是否满足条件式(12), 若不满足, 从点云中除去记录的内点集合, 继续拟合平面, 否则此平面为地平面, 停止, 输出相机的外参。

### 3 实验

#### 3.1 实验过程

本实验使用PrimeSense相机采集视频数据, 利用Matlab平台进行仿真, 对本文算法进行验证。为了便于精度对比, 实验将彩色相机的棋盘标定结果作为基准数据。由于矩阵不宜进行比较, 实验过程中, 将外参转化为相机的高度( $H$ )、侧倾角( $\theta$ ), 通过对比3个参数来衡量标定的精度。实验按照以下步骤进行:

1) 使用PrimeSense相机采集视频数据(视频的每帧都有清晰的棋盘), 视频数据包含彩色视频和深度视频;

2) 从彩色视频中随机选择 $N(N > 20)$ 帧视频帧, 作为张正友相机标定方法<sup>[14]</sup>的输入, 估计出相机的内参;

3) 遍历彩色视频的每帧, 利用内参和当前帧所

检测的棋盘角点估计出相机外参, 得到每帧彩色视频的相机姿势( $H_{\text{chessboard}}$ ,  $\phi_{\text{chessboard}}$  和  $\theta_{\text{chessboard}}$ );

4) 遍历深度视频的每帧, 利用本文所述方法检测地平面, 并估计出相机的外参, 得到每帧深度视频的相机姿势( $H_{\text{plane}}$ ,  $\phi_{\text{plane}}$  和  $\theta_{\text{plane}}$ ), 计算每帧视频的相机高度差( $\Delta H$ )、侧倾角差( $\Delta\theta$ )、偏航角差( $\Delta\phi$ ), 计算公式为

$$\begin{cases} \Delta H = H_{\text{chessboard}} - H_{\text{plane}} \\ \Delta\phi = \phi_{\text{chessboard}} - \phi_{\text{plane}} \\ \Delta\theta = \theta_{\text{chessboard}} - \theta_{\text{plane}} \end{cases} \quad (13)$$

在实验步骤1)中, 所采集的视频格式如表1所示。

表1 视频参数

Table 1 Parameters of video

	彩色视频	深度视频
视频大小/像素	320 × 240	320 × 240
帧速率/(帧/s)	30	30
时长/s	89.400	89.400
像素字节/byte	24	16
视频格式	RGB24	Monol6

视频的的每帧都含有棋盘, 如图8, 在实验过程中, 使用棋盘估计相机外参作为基准数据, 由于在利用张正友标定方法计算相机内参的时候, 不同的棋盘, 对应着不同的内参, 因此, 重复步骤1)~3)共86次, 每幅视频帧可以估计出86种相机外参, 以其中值作为基准数据。

采用的棋盘的大小是 $7 \times 9$ , 每个棋盘长和宽均为40 mm, 利用棋盘标定相机外参的结果如图8所示。



图8 包含棋盘的彩色视频帧

Fig.8 Color video frame comprising chessboard

图9中,  $\theta_{\text{chessboard}}$ ,  $\phi_{\text{chessboard}}$ ,  $H_{\text{chessboard}}$  分别为棋盘方法所测得的相机俯仰角、侧倾角、高度的中位



数;  $\theta_{\text{plane}}$ ,  $\phi_{\text{plane}}$ ,  $H_{\text{plane}}$  为本文所提出的方法测得的相机俯仰角、侧倾角和高度的中位数。

图 10 为实验所得到的误差, 表 2 为最终所测得的结果。

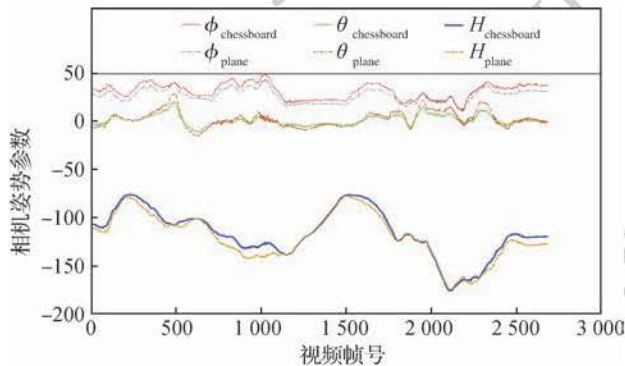


图 9 棋盘标定方法结果

Fig. 9 The result of checkboard calibration method

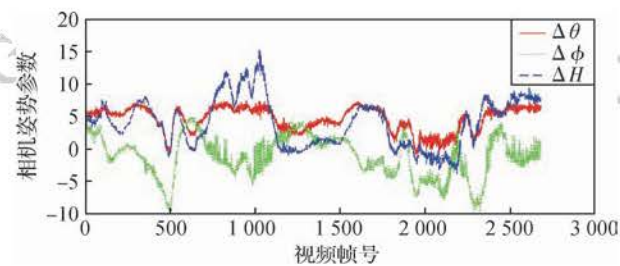


图 10 测试结果

Fig. 10 Test result

表 2 相机外参误差统计表

Table 2 Table of extrinsic parameters error of camera

	$\theta$	$\phi$	$H$
最大	7.48	4.89	15.22
最小	-0.75	-9.62	-1.14
平均	15.22	-3.86	3.96

### 3.2 误差分析

在实验中, 由于没有高精度的仪器来衡量相机的外参, 所以使用棋盘标定方法测出来的相机外参作为基准数据来衡量本方法的精度。影响本实验精度的因素有以下几点:

1) 角点检测的量化误差。由于使用的视频帧是  $240 \times 320$  像素, 因此在计算相机内参和外参的时候, 因角点检测的量化误差较为严重的;

2) 棋盘方法测量的 RGB-D 相机的彩色传感器的外参, 而本方法测量则是 RGB-D 相机的红外相

机的外参;

3) RGB-D 相机扫描场景数据的噪声, 会影响地平面的检测, 进而影响相机外参的精度;

4) 地平面检测中 MLESAC 迭代停止的参数也会影响本实验的结果。

### 3.3 场景噪声对地平面检测的影响分析

RGB-D 相机有诸多噪声来源(如: 温度、环境光入射角度和光强、纹理等<sup>[13]</sup>), 本文采用的 MLESAC 可以处理微小噪声, 但不能够处理噪声过多或者数据缺失过多的场景:

1) 强阳光会造成地平面噪声点过多, 导致平面参数估计不准确, 相机的高度、侧倾角产生轻微影响;

2) 地面的反射率过高(比如: 地面上放置一个镜子), 会造成该区域的数据缺失, 若缺失数据过多, 则会引起无法检测地平面, 导致无法估计出相机的外参。

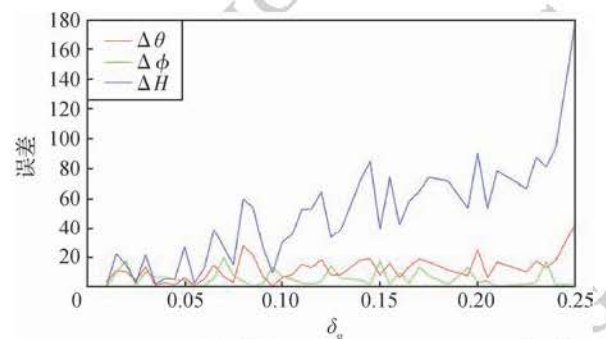


图 11 高斯噪声结果的影响

Fig. 11 The effect of noise on the result

为了验证噪音对相机姿势的影响, 图 11 是在第 929 帧, 随着高斯噪音方差的增加, 相机姿势的误差变化。高斯噪音的均值为 0, 方差为  $\delta_\epsilon$ , 可见, 随着噪音方差的增大, 高度的误差增加, 俯仰角和侧倾角的稳定性降低, 此外, 当方差大于 0.25, 则无法估计出正确的平面。

## 4 结 论

本文自动标定方法的原因是, 在进行基于 RGB-D 相机的行人检测时, 由于环境震动影响, RGB-D 相机会偏移原来的位置, 外参数会发生很大变化, 直接影响检测精度; 利用此种自动外参标定方法, 可以在无人出现时自动进行外参的重新矫正, 从



而解决上述提到的问题。

本文方法可适用于3D相机(散斑相机, TOF相机, 双目相机)的外参自动调整, 具有很高的可并行性, 并行实现后, 具有实时性, 可以用于移动机器人上的3D相机的外参自动标定。

本文方法从3维点云中提取平面, 利用平面与世界坐标的位置关系, 得到相机外参, 通过外参判断当前的平面是否为地平面, 如果不为地平面, 则继续使用下一平面计算外参, 直到寻找到地平面, 得到外参, 红外相机的外参数。本文给出了地平面的条件, 从而能够保证所建立的世界坐标系的正确性, 创新性地将平面检测和相机外参估计结合在一起, 从而达到自动外参标定的目的。本方法不需要额外的标定物, 同时是针对红外相机的标定, 标定结果可信, 并且提供本文方法的源码(<https://github.com/shijieS/RGBDCameraExtrinsicCalibration.git>)。

目前方法的局限性是需要场景中存在地平面, 若检测不到地平面, 则不能正常地进行标定。未来需要解决两个问题: 1) 提高标定的精度, 利用场景中的行人等目标, 进行精细标定; 2) 多个相机的外参标定, 根据多个相机拍摄的不同区域进行自动标定, 若两个相机无公共区域, 则通过设计简单的标定物, 根据标定物的几何尺寸进行自动标定。

## 参考文献 (References)

- [1] Microsoft. Kinect for X-box 360[EB/OL]. 2017-01-01 [2017-10-12]. <http://www.xbox.com/en-US/xbox-one/accessories/kinect>.
- [2] Wikipedia. Kinect[EB/OL]. 2018-01-12 [2017-08-29]. <https://en.wikipedia.org/wiki/Kinect>.
- [3] Barak F, Alexander S, Meir M, et al. Depth mapping using projected patterns: US, US8493496B2[P]. 2013-07-23.
- [4] Endres F, Hess J, Sturm J, et al. 3-D mapping with an RGB-D camera[J]. IEEE Transactions on Robotics, 2014, 30(1): 177-187. [DOI: 10.1109/TRO.2013.2279412]
- [5] Labbé M, Michaud F. Online global loop closure detection for large-scale multi-session graph-based SLAM[C]//Proceedings of 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Chicago, IL, USA: IEEE, 2014: 2661-2666. [DOI: 10.1109/IROS.2014.6942926]
- [6] Munaro M, Basso F, Menegatti E. OpenPTTrack: Open source multi-camera calibration and people tracking for RGB-D camera networks[J]. Robotics and Autonomous Systems, 2016, 75: 525-538. [DOI: 10.1016/j.robot.2015.10.004]
- [7] Munaro M, Menegatti E. Fast RGB-D people tracking for service robots[J]. Autonomous Robots, 2014, 37(3): 227-242. [DOI: 10.1007/s10514-014-9385-0]
- [8] Liao Q H, Liu M, Tai L, et al. Extrinsic calibration of 3D range finder and camera without auxiliary object or human intervention[J]. arXiv:1703.04391, 2017.
- [9] Basso F, Menegatti E, Pretto A. Robust intrinsic and extrinsic calibration of RGB-D cameras[J]. arXiv:1701.05748, 2017.
- [10] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334. [DOI: 10.1109/34.888718]
- [11] Li S B, Zhuo Q. A new approach to calibrate range image and color image from Kinect[C]//Proceedings of the 4th International Conference on Intelligent Human-Machine Systems and Cybernetics. Nanchang, Jiangxi, China: IEEE, 2012, 2(24): 252-255. [DOI: 10.1109/IHMSC.2012.156]
- [12] Torr P H S, Zisserman A. MLESAC: a new robust estimator with application to estimating image geometry[J]. Computer Vision and Image Understanding, 2000, 78(1): 138-156. [DOI: 10.1006/cviu.1999.0832]
- [13] Belhedi A, Bartoli A, Bourgeois S, et al. Noise modelling in time-of-flight sensors with application to depth noise removal and uncertainty estimation in three-dimensional measurement[J]. IET Computer Vision, 2015, 9(6): 967-977. [DOI: 10.1049/iet-cvi.2014.0135]
- [14] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on pattern analysis and machine intelligence. 2000, 22(11): 1330-4. [DOI: 10.1109/34.888718]