

中图法分类号: 文献标识码:A 文章编号:1006-8961(年) -

论文引用格式: Shijie S, HuanSheng S. [孙士杰, 宋焕生. 基于地平面检测的 RGB-D 相机外参自动标定方法[J]. 中国图象图形学报, 2017

点云下地平面检测的 RGB-D 相机外参自动标定

孙士杰¹, 宋焕生¹, 张朝阳¹, 张文涛¹, 王璇¹

1. 长安大学, 陕西省, 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, 可以应用在三维场景重建、三维测量、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** 该方法首先将深度图转化为相机坐标系下的三维点云, 利用 MELSAC 方法自动检测三维点云中的平面, 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° , 平均俯仰角误差为 4.57° , 平均相机高度误差为 3.96cm. **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, 算法的具有较高地并行性, 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; 三维点云; 平面检测; 深度图

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

Sun ShiJie¹, Song HuanSheng¹, Zhang ChaoYang¹, Wang Xuan¹

1. Chang'an University, Shaanxi XiAn, 710064

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. It can be applied to 3D reconstruction, 3D measurement, robot gesture estimation, target detection and so on. The RGB-D camera (e.g. Kinect, PrimeSense, RealSense etc.) consists of two sensors: RGB sensor and depth sensor. The former sensor retrieves the RGB image, whereas the

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

收稿日期: ; **修回日期:**

第一作者简介: 孙士杰(1989-), 男, 长安大学博士研究生, 2014 年长安大学(在读)交通信息工程及控制博士学位, 主要研究方向为三维点云目标识别、机器视觉、跟踪、人工智能. E-mail: shijieSun@chd.edu.cn

通信作者: 孙士杰, 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(2017JQ6064)

latter retrieves depth image from the scene. In order to translate the 3D point cloud in the camera coordinate to the world coordinate, the extrinsic parameters of depth sensor need to be calibrated. The general calibration methods use the calibration objects (such as chessboard) to get the extrinsic parameter of the RGB-D color sensor which is regarded as the extrinsic parameter of the depth sensor approximately. These methods don't make full use of depth information, which make it difficult to simplify the calibration process. Besides, the ignorance of the difference between the depth sensor and color sensor can make some larger errors. In order to estimate the extrinsic parameter of depth sensor in the RGB-D camera accurately, some methods have been proposed by using the extrinsic parameters of depth sensor relative to the color sensor. However, these methods complicate the process of calibrating. In order to simplify the process of the calibration of the extrinsic parameter of depth sensor, it is necessary to make full use of the depth information. Results of methods based on the color image are the parameter of the color sensor, However, the majority applications of the RGB-D camera are based on depth sensor. It is also necessary to calibrate parameters of depth sensor directly. **Methods** 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 two conditions: 1) the angle between the z-axis of the camera and the ground plane is less than a specified threshold. 2) the z value of the ground plane in the world coordinate is larger than that of other points which are not in the ground plane. Besides we also 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 pointing to 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 by the following step. Firstly, 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 lots of planes. Secondly, planes in the 3D point cloud are detected by the MELSAC method. There are at most one ground plane in these detected planes. Thirdly, the spatial constraint rule between ground plane and camera is built. these detected planes are filtered by the spatial constraint rule, until finding the ground plane or iterating all the plane. The process will get stopped when it doesn't find a ground plane. Finally, by using the relation between the ground plane and the camera, points set are selected in order to calculate extrinsic parameter. **Result** The benchmark is the result of checkboard extrinsic calibration by processing the RGB image of RGB-D information which is retrieved from the PrimeSense camera. We record an 89.4s video which is used in the experiment. The videos contain two videos: the RGB video (the channel is 3) and the depth video (the channel is 3). There is a 7x7 checkboard in every frame of the RGB video, which is processed by the checkboard 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 degrees, the average pitch angle error is 4.57 degrees, and the average camera height error is 3.96cm. We also do the experiment of testing the robustness of the noise. In the experiment, we increase the variance of Gaussian noise in the depth frame and get the result of each variance Gaussian noise. With the increase of variance of Gaussian noise, the stability of calibration decreases. The result shows that our method performs well when the variance of Gaussian noise below 0.01. **Conclusion** Our proposed method makes full use of the depth information of the RGB-D camera, and simplifies the process of extrinsic calibration of the depth sensor greatly, which can be used in the actual application. For the sake of convenience, the source code is also published. This method can automatically detect the ground plane and doesn't need other calibration objects. From the experiments, we can conclude that the proposed method can calibrate each frame of recorded video accurately and it isn't 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 parallelly. Based on the parallel implementation, the proposed method will have real-time performance.

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^[1-3]、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三维场景重建和导航^[4,5]、目标识别和跟踪^[6,7]、三维测量等应用中. 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^[7]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.

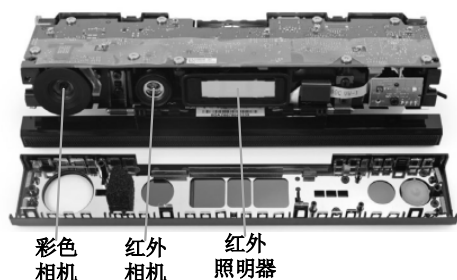


图1 Kinect的结构

Fig.1 Structure of the Kinect

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^[8,9]. 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

^[10], Munaro 等人^[6]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三维重建, 会有较大误差. Shibo 等人^[11]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^[8]将标定划分为三类: 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^[8]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三类: 第一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 ${}^D_C T$, 通过变换 ${}^D_C T = {}^D_C T {}^C_C T$, 得

到红外相机外参 ${}^D_D T$. 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. 第二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, 相对第一类标定方法而言, 并没有简化标定过程. 第三类使用深度图对外参相机进

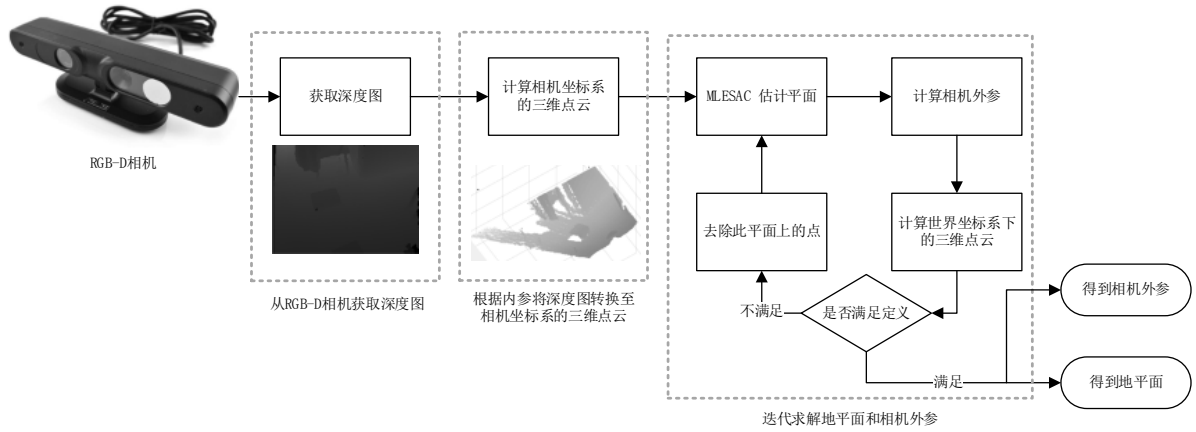


图 2 相机外参标定流程

Fig.2 Process of camera extrinsic parameter calibration

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。相机的外参为公式 (1)。

$${}^w_c T = \begin{bmatrix} {}^w_c R & 0 \\ {}^w_c t & 1 \end{bmatrix} \quad (1)$$

式中， ${}^w_c T \in \mathbb{R}^{4 \times 4}$ ， ${}^w_c R \in \mathbb{R}^{3 \times 3}$ 为相机的旋

转矩阵， ${}^w_c t \in \mathbb{R}^{1 \times 3}$ 为相机的平移矩阵， ${}^w_c(\cdot)$ 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三维点云；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三维坐标系，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。



图 3 相机坐标系

Fig.3 Camera Coordinate

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水平

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

收稿日期：； **修回日期：**

第一作者简介：孙士杰(1989-)，男，长安大学博士研究生，2014 年长安大学（在读）交通信息工程及控制博士学位，主要研究方向为三维点云目标识别、机器视觉、跟踪、人工智能。E-mail:shijieSun@chd.edu.cn

通信作者：孙士杰，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(2017JQ6064)

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;

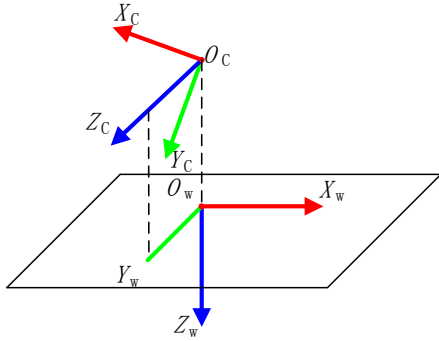


图 4 世界坐标系的位置

Fig.4 The position of world coordinate

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, 即相机的外参 ${}^w_c T$, 为 4×4 的矩阵, 有 12 个未知参数, 由公式(1). 本文首先计算世界坐标系到相机坐标系的变换矩阵 ${}^c_w T$, 然后利用公式(2)得到相机的外参.

$${}^w_c T = {}^c_w T^{-1} \quad (2)$$

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:

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

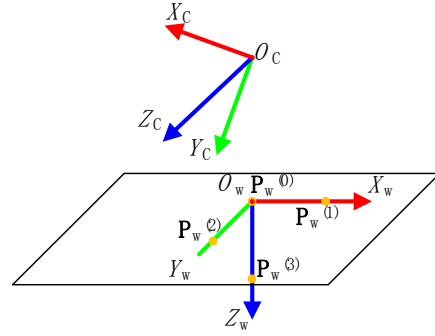


图 5 在世界坐标系下选中的点

Fig.5 Selected points in world coordinate

通过计算对应的相机坐标系下的点, 即

$P_c^{(0)}, P_c^{(1)}, P_c^{(2)}, P_c^{(3)}$, 便可求出 ${}^c_w T$, 见公式(4).

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

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: $\mathbf{n}_c = [a \ b \ c]$ 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 $P_w^{(0)}$ 在相机坐标系下的坐标为:

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

点 $P_w^{(3)} = [0 \ 0 \ 1]$ 在世界坐标系 Z 轴上, 由 $O_c, P_c^{(0)}, P_c^{(3)}$ 共线, 那么有:

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

因 Y_w 轴为 Z_c 轴在平面上的投影, 故

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

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

$$\mathbf{r}_{\mathbf{n}_y} = \frac{\mathbf{P}_{Z_c} - \mathbf{P}_C^{(0)}}{\|\mathbf{P}_{Z_c} - \mathbf{P}_C^{(0)}\|_2} \quad (8)$$

\mathbf{P}_{Z_c} 为相机坐标系的 $[0 \ 0 \ 1]$ 点到平面的投影.

通过解方程(9), 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 $\mathbf{P}_C^{(1)}$, 如公式

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

通过公式(5), (6), (7), (10), 求出从世界坐标系下选出的四个点在相机坐标系下的坐标, 然后利用公式(4)得到世界坐标点到相机坐标点的转换矩阵, 最后利用(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, 本文首先利用内参将深度图转化为相机坐标系下的三维点云, 利用 MLESAC 方法^[12]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三维点云

如图 6, 是相机的成像模型, 像平面的坐标轴分别用 X_I 和 Y_I 来表示, 相机坐标系下的点

$\mathbf{P}_C = (X_C, Y_C, Z_C)$ 在像平面上的成像点为

$\mathbf{p}_I = (u_I, v_I)$, 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 . 那么有公式(11)可将 (u, v, d_p) 转换为 (X_C, Y_C, Z_C) .

$$\begin{cases} f_x = \frac{f}{d_x} \\ f_y = \frac{f}{d_y} \\ Z_C = d_p * s \\ X_C = \frac{(u_I - c_x) \cdot Z_C}{f_x} \\ Y_C = \frac{(v_I - c_y) \cdot Z_C}{f_y} \end{cases} \quad (11)$$

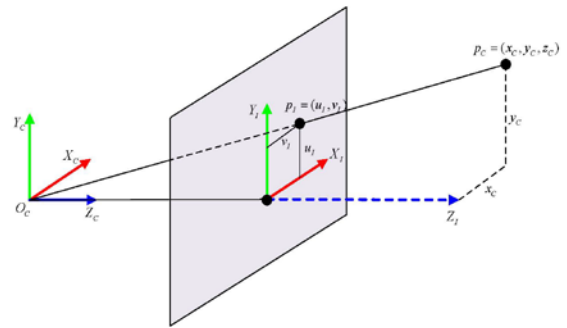


图 6 相机的成像模型

Fig.6 The model of camera

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:

$$\begin{cases} \theta < \frac{\pi}{4} \\ median P_{w_i}^Z > min P_{C_o}^Z + \varepsilon \end{cases} \quad (12)$$

式中, θ 为相机的 X_C 轴与平面的夹角,

$median$ 是取中值的操作, $P_{w_i}^Z$ 为拟合平面模型

的内点 Z 值集合, $P_{C_o}^Z$ 为拟合平面模型的外点 Z

值集合, ε 为设定的容差值. 公式(12)表示两个

条件：1. 相机相对于平面的倾斜角不超过 45° ；
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线，本方法的流程如图 7。

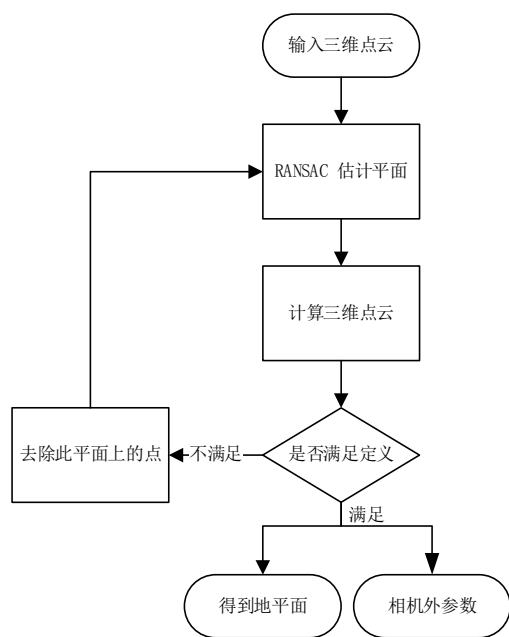


图 7 地平线估计的流程图

Fig.7 The process of ground plane estimation

首先输入相机坐标系下的三维点云，利用 MLESAC 方法，计算平面，并记录满足该平面的内点集合，利用此平面和节 2.2 的方法，计算出相机的外参数，结合相机的内参数，将相机坐标系下的三维点云转换至世界坐标系下的三维点云，判断此时是否满足条件(12)，若不满足，从点云中除去记录的内点集合，继续拟合平面，否则此平面为地平线，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，对本文所提出的算法进行验证。为了便于精度对比，本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(θ)，通过对比三个参数来衡量标定的精度。实验按照以下步骤进行：
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；

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的输入，估计出相机的内参；

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 $\theta_{\text{chessboard}}$)；

遍历深度视频的每帧，利用本文所述方法检测地平线，并估计出相机的外参，得到每帧深度视频的相机姿势(H_{plane} , ϕ_{plane} 和 θ_{plane})

计算每帧视频的相机高度差(ΔH)、侧倾角差($\Delta \theta$)、偏航角差($\Delta \phi$)，使用公式(13)。

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

在实验步骤 1 中，所采集的视频格式如表 1。

表 1 视频参数

Table 1 Parameters of video

	彩色视频	深度视频
视频大小	320x240	320x240
帧速率(帧/s)	30	30
时长(秒)	89.400	89.400
像素字节	24	16
视频格式	RGB24	Mono16

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，我们重复步骤 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 7×9 ，每个棋盘长和宽均为 40mm，利用棋盘标定相机外参的结果如图 8。



图 8 包含棋盘的彩色视频帧

Fig.8 Color video frame comprising chessboard

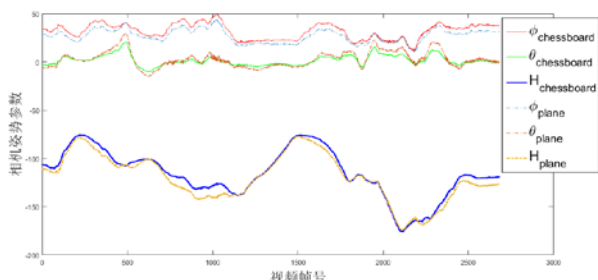


图 9 棋盘标定方法结果

Fig.9 The result of checkboard calibration method.

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; θ_{plane} , ϕ_{plane} , H_{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。

如图 10 为实验所得到的误差, 表 2 为最终所测得的结果。

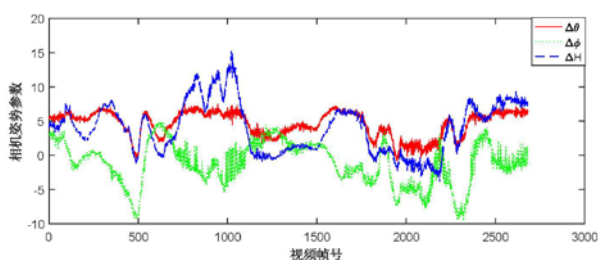


图 10 测试结果

Fig.10 Test result

表 2 相机外参误差统计表

Table 2 Table of extrinsic parameters error of camera

	θ	ϕ	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 240×320 , 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^[13]), 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。

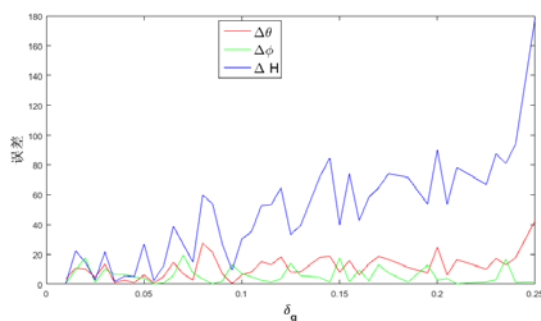


图 11 高斯噪声结果的影响

Fig.11 The effect of noise on the result

为了验证噪声对相机姿势的影响, 图 11 是在第 929 帧, 随着高斯噪声方差的增加, 相机姿势的误差变化。高斯噪声的均值为 0, 方差为 δ_g , 可见, 随着噪声方差的增大, 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。

四 结论与展望

本文所提出自动标定方法的原因是, 作者在进行基于 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, 可以在无行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, 本方法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。

本方法从三维点云中提取平面, 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相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]. 2017-07-04 [2017-08-29]. <https://en.wikipedia.org/wiki/Kinect>.
- [3] Freedman B, Shpunt A, Machline M, et al.

Depth mapping using projected patterns: US, US 8493496 B2[P]. 2013.

[4] Endres F, Hess J, Sturm J, et al. 3-D Mapping With an RGB-D Camera[J]. IEEE Transactions on Robotics, 2017, 30(1):177-187.

[5] Labbe M, Michaud F. Online global loop closure detection for large-scale multi-session graph-based SLAM[C]// International Conference on Intelligent Robots and Systems. Chicago, IL, USA: IEEE, 2014:2661-2666.

[6] Munaro M, Basso F, Menegatti E. OpenPTrack: Open source multi-camera calibration and people tracking for RGB-D camera networks[J]. Robotics & Autonomous Systems, 2016, 75:525-538.

[7] Munaro M, Menegatti E. Fast RGB-D people tracking for service robots[J]. Autonomous Robots, 2014, 37(3):227-242.

[8] Liao Q, Liu M, Tai L, et al. Extrinsic Calibration of 3D Range Finder and Camera without Auxiliary Object or Human Intervention[J]. 2017.

[9] Basso F, Menegatti E, Pretto A. Robust Intrinsic and Extrinsic Calibration of RGB-D Cameras[J]. 2017.

[10] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(11):1330-1334.

[11] Li S, Zhuo Q. A New Approach to Calibrate Range Image and Color Image From Kinect[J]. 2012, 2(24):252-255.

[12] Torr P H S, Zisserman A. MLESAC: A New Robust Estimator with Application to Estimating Image Geometry[J]. Computer Vision & Image Understanding, 2000, 78(1):138-156.

[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]. Computer Vision Iet, 2015, 9(6):967-977.

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

收稿日期：； **修回日期：**

第一作者简介：孙士杰(1989-)，男，长安大学博士研究生，2014 年长安大学（在读）交通信息工程及控制博士学位，主要研究方向为三维点云目标识别、机器视觉、跟踪、人工智能. E-mail: shijieSun@chd.edu.cn

通信作者：孙士杰, 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(2017JQ6064)