

双视系统的室内三维场景重建研究

陈颖 赵来旺 詹洪陈 丁尧

南京大学金陵学院 南京 210089

(1803901252@qq.com)

摘要 在早期三维场景重建中,受硬件条件的限制,无法很好地对场景进行三维重建。随着硬件的更新迭代,目前可利用结构光和双视系统进行更加高效的三维场景重建。以 ZED 双目相机和捷宝 AD-10 电动云台构建硬件平台,通过双目相机获得场景的点云信息,基于全局 SBGM(Semi-Global Block Matching)算法的立体匹配和 RGB-D 图(Red Green Blue Depth Map)的生成来进行单一场景点云重建,并通过 ORB 特征匹配和 ICP(Iterative Closest Point)点云配准融合,实现对室内场景的全景三维重建。实验对比了双目立体视觉的场景重建在过远/过近目标、低纹理特征目标、玻璃等材质目标上的优势,同时在点云的三维场景重建中,提出通过稀疏化来优化点云信息,对比单一采集与多次采集的重建效果。经过实验,该系统在折中采用次数的前提下,可以兼顾场景重建细节与显示效果,并对不同场景目标的三维重建具有实际的工程借鉴意义和应用价值。

关键词: 三维场景重建;双目视觉;全景场景重建;PCL;ICP

中图分类号 TP391

Study on Reconstruction of Indoor 3D Scene Based on Binocular Vision

CHEN Ying, ZHAO Lai-wang, ZHAN Hong-chen and DING Yao

Jinling College, Nanjing University, Nanjing 210089, China

Abstract In the early three-dimensional scene reconstruction, due to hardware constraints, it can not be very good for the three-dimensional reconstruction of the scene. With the hardware update iteration, at present, the structured light and dual vision system is used for for more efficient three-dimensional scene reconstruction. The hardware platform is built with zed binocular camera and jabao Ad-10 electric cloud platform, and the point cloud information of the scene is obtained through binocular camera. Based on the stereo matching of the global sbgm (semi global block matching) algorithm and the red green blue depth map generation, the single scene cloud reconstruction is carried out, and through the orb feature matching and ICP (iterative closure Point) point cloud registration and fusion can realize panoramic 3D reconstruction of indoor scene. The experiment compares the advantages of binocular stereo vision scene reconstruction in far/near targets, low texture feature targets, glass and other material targets. At the same time, in the three-dimensional scene reconstruction of point cloud, this paper proposes to optimize the point cloud information through sparseness, and compares the reconstruction effect of single acquisition and multiple acquisition. After the experiment, the system can give consideration to the details of scene reconstruction and the display effect under the premise of the eclectic number of times, and it can be used for reference and application value for 3D reconstruction of different scene targets.

Keywords Binocular vision, 3D scene reconstruction, Panoramic scene reconstruction, PCL, ICP

近年来,随着视觉技术的发展,通过图像进行三维场景重建是图像技术应用的热点。目前三维场景重建多分为两个方向:双目立体相机视觉技术和主动式红外结构光技术。双目立体视觉模仿人类的眼睛通过左右试图的视差信息来进行深度信息的测算,这种系统结构简单,成本低,易实现,且更接近于人眼的视觉效果^[1]。主动式红外结构光技术的代表为微软开发的 Kinect,它可直接通过接受反射红外光的时延来计算对应场景的深度信息。目前利用 Kinect 实现三维重建的技术十分成熟,业界有很多相关的开源项目,如 RGBD SLAM (Simultaneous Localization and Mapping),该项目结合 Kinect 与 ROS(Robot Operating System),包含回环检测,能做到室内定位及地图重建^[2]。

本文利用 ZED 双目相机和电动云台构建双视视觉系统,

设计了一种基于双目的全景立体成像系统,在不增加硬件资源的前提下,通过该系统可实现对室内场景的三维全景重建。

1 双视系统的三维场景重建建模分析

1.1 双目立体视觉模型分析

对于双目立体视觉系统,可以通过图 1 所示的相似三角形的原理来计算现实环境中某点到相机基线上的距离。如图 1 所示,利用两个同构的单目相机搭建一个双目系统,对于空间中一点 P ,在左右两个视场中的投影分别记为 p 和 p' ,对应的横坐标分别为 X_R 和 X_L ,左右相机的焦距均为 f ,两个相机间的基线长度为 B ,空间中点 P 到基线的距离为 z_p ,利用三角形相似性质,可以得出 z 的计算公式如式(1)所示^[3-4]。

$$z_p = b * f / |X_R - X_T| \quad (1)$$

基金项目:江苏省高校自然科学基金(BK20151299)

This work was supported by the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province, China(BK20151299).

通信作者:丁尧(ammdistin@nju.edu.cn)

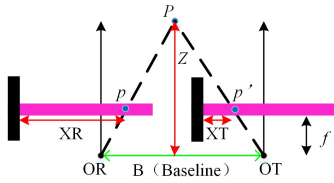


图1 双目立体视觉系统示意图

Fig. 1 Schematic diagram of binocular stereo vision system

若对空间中所有的点都进行式(1)所示的计算,并与图像中的像素一一对应,便可以生成图像的 RGB-Depth 数据。

如图 2 所示,在得到空间中点 P 的深度信息后,通过相机内外参数可以计算出点 P 在世界坐标系中的坐标 $P(X_{PW}, Y_{PW}, Z_{PW})$,完成深度数据到点云数据的转换^[5]。

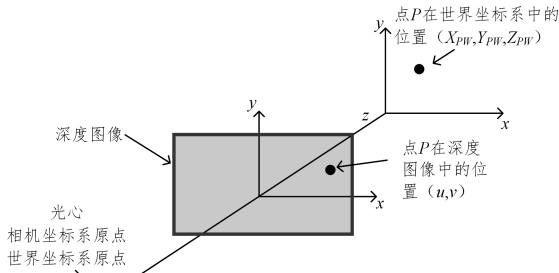


图2 深度图到点云模型图

Fig. 2 Depth map to point cloud model

可以将相机的成像近似成小孔成像,其数学模型如式(2)所示。

$$z_p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} [R \ T] \begin{bmatrix} X_{PW} \\ Y_{PW} \\ Z_{PW} \\ 1 \end{bmatrix} \quad (2)$$

式(2)中, z_p 为点 P 的深度信息,由式(1)计算得出, (u, v) 为点 P 在图像中的像素坐标, (u_0, v_0) 为图像的中心坐标, f 为相机镜头焦距, dx, dy 为像素尺寸。上述 5 个参数为相机内参, (X_{PW}, Y_{PW}, Z_{PW}) 为点 P 在世界坐标系下待计算的坐标, R, T 为相机坐标系到世界坐标系的转换映射关系^[6]。为便于计算,往往将相机坐标系和世界坐标系重合,因此式(2)可以简化为式(3)。

$$z_p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_{PW} \\ Y_{PW} \\ Z_{PW} \\ 1 \end{bmatrix} \quad (3)$$

通过式(3)可以计算出点 P 在世界坐标系下的坐标,如式(4)所示,从而得到点 P 的点云数据。

$$\begin{cases} X_{PW} = z_c * (u - u_0) * dx / f \\ Y_{PW} = z_c * (v - v_0) * dy / f \\ Z_{PW} = z_c \end{cases} \quad (4)$$

1.2 基于 ORB 和 ICP 的运动估计

ORB 特征点是在 FAST 角点的基础上添加描述子 BRIEF 的改进型 FAST 角点,提取 ORB 特征时,要先找出图像中的“FAST 角点”,再为其添加有着旋转不变特性的 BRIEF 描述子。先通过 FAST 角点检测快速的标定 FAST 角点,然后对原始的 FAST 特征点进行 Harris 响应值计算,最终选取固定数量的具有最大响应值的特征点作为角点^[7]。

提取好特征点后,需要对视觉系统中某一时间段内前后帧之间的特征点进行匹配,得出前帧和后帧之间的匹配点,并利用 ICP 算法计算匹配点直接的空间变换,估算误差。相邻前后两帧之间存在两组匹配的特征点分别为 $P = \vec{p}_1, \vec{p}_2, \dots, \vec{p}_n$ 和 $P' = \vec{p}'_1, \vec{p}'_2, \dots, \vec{p}'_n$,对于两个匹配点 \vec{p}_i 和 \vec{p}'_i ,都存在一个欧氏变换,如式(5)所示^[8]。

$$\vec{p}_i = R \vec{p}'_i + \vec{T} \quad (5)$$

式(5)中的欧氏变换矩阵 RT 可以近似理解为对相机运动的估算,对 RT 的计算可以通过迭代最近点 ICP 来解决。

2 实现双视系统的三维重建

双目立体视觉系统整体结构如图 3 所示。

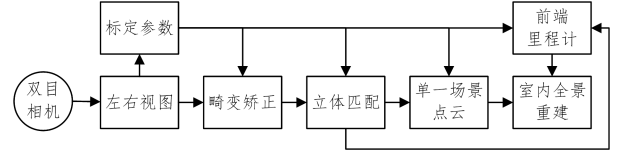


图3 室内全景重建结构图

Fig. 3 Interior panorama reconstruction structure drawing

2.1 单一场景下的三维场景重建与比较

本文利用 Stereo_Match 模块来实现 SGBM 立体匹配算法,单一视角下的全景场景重建如图 4 所示。

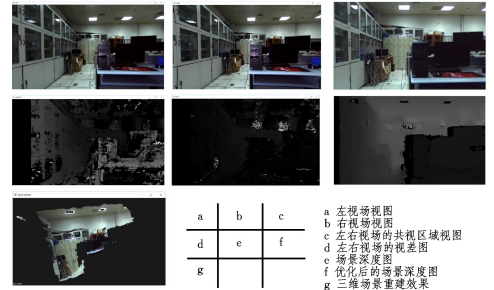


图4 单一视角下双目系统三维场景重建效果

Fig. 4 3D scene reconstruction of binocular system from a single perspective

在完成单一视角下的双目系统三维场景重建的基础上,本文又针对场景进行场景重建研究,结果如图 5 所示。



图5 不同环境目标的三维场景重建效果对比

Fig. 5 Comparison of 3D scene reconstruction effects of different environmental targets

2.2 室内三维全景重建效果

利用点云进行三维场景重建时,首先需要采集点云数据。受环境光和物体表面等因素影响,数据中会产生噪声,因此需要对生成的点云数据进行滤波。利用 PCL 的 Statistical OutlierRemoval 库函数,可以生成高斯分布的点云滤波器。滤波前后的三维场景如图 6 所示。

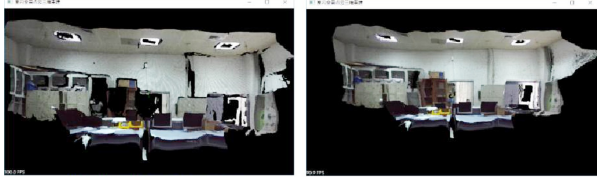


图6 滤波前后三维场景重建效果比较

Fig. 6 Comparison of 3D scene reconstruction before and after filtering

为了对全景信息进行三维场景重建,本文利用相机的位姿信息将多个相机坐标系下的点云转换到世界坐标系下来完成全景重建。首先通过 ORB 特征提取及匹配,利用 ICP SVD 方法来计算变换矩阵,并将其转换为平移向量+旋转四元数^[9]。然后以第一个点云为基准,把后续单一场景点云通过对应的位姿信息进行欧氏变换进而转换到以第一个点云为基准的世界坐标系上去,来实现三维场景的全景重建。实验室全景三维场景重建效果(俯视)如图 7 所示。



图7 实验室全景三维场景重建效果俯视图

Fig. 7 Top view of laboratory panoramic 3D scene reconstruction effect

3 基于稀疏重建的系统优化

PCL 中提供了 VoxelGrid 体素采样函数来实现点云数据稀疏化。通过体素采样,可以提取点云的关键结构,降低点云密度。场景稀疏化前后的对比如图 8 所示。

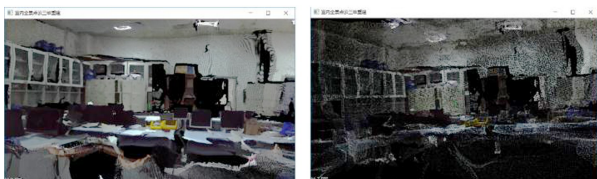


图8 稠密点云三维重建与稀疏点云三维重建对比图

Fig. 8 Comparison of dense point cloud 3D reconstruction and sparse point cloud 3D reconstruction

对比点云稀疏前后的三维重建效果,可以看出稀疏后的点云会损失部分细节。采集大量的图像,利用滤波与点云稀疏化将冗余点去掉。两种方法的重建效果如图 9 所示。

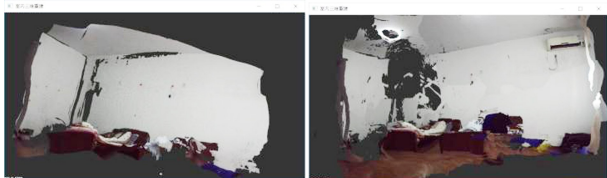


图9 一次采集与多次采集的重建效果比较

Fig. 9 Comparison of effects between one acquisition and multiple acquisitions

对两者的分析如表 1 所列。

表1 两种采集方式的重建结果对比分析

Table 1 Comparative analysis of reconstruction results of two acquisition methods

	效果	时间/s
采集 1 次	细节少,但点云冗余度低,显示效果较好。	8
采集 8 次	细节较为完善,但点云冗余度较高,显示效果不太好,运算时间长。	112

结束语 为了更贴合人眼的生物感观,本文利用双目相机构建双目立体视觉系统,来对场景进行三维场景信息采集与全景重建。研究表明,该系统 and 三维全景重建方法在地图重建、目标三维建模和沉浸式漫游等领域具有很好的应用价值和工程参考意义。

参考文献

- [1] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments[J]. The International Journal of Robotics Research, 2012, 31(5): 647-663.
- [2] HUANG Y, HUANG J, XIAO C S, et al. Research on Ship Trajectory Tracking Algorithm based on Binocular stereo Vision [J]. Computer Science, 2017(1): 308-313.
- [3] YU J W, WEI C. Indoor Mobile Measurement System based on SLAM and its application [J]. Bulletin of Surveying and Mapping, 2016(6): 146-147.
- [4] IZADI S, KIM D, HILLIGES O, et al. Kinect Fusion: 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. ACM, 2011: 559-568.
- [5] ZHANG C, LI Z, CHENG Y, et al. Meshstereo: A global stereo model with mesh alignment regularization for view interpolation [C]// Proceedings of the IEEE International Conference on Computer Vision. 2015: 2057-2065.
- [6] MIAO H C, WANG J H, ZHANG Q Y. Concave-convex Manufacturing Feature Recognition based on single view 3D Reconstruction [J]. Computer Science, 2019(7): 280-285.
- [7] HU Z Y, WU F C. Calibration method based on active vision camera [J]. Journal of Computer Science, 2002, 25(11): 1149-1156.
- [8] ZHOU S P, GENG G H, LI K, et al. A multi-view geometric 3D reconstruction method based on AKAZE Algorithm [J]. Computer Science, 2018(S2): 180-184, 207.
- [9] ZHAO F Q, JIA Y T. Point Cloud Registration Method based on Geometric Attributes and improved ICP [J]. Information Technology, 2019(4): 33-38.



CHEN Ying, born in 1999, undergraduate. Her main research interests include digital image processing and computer vision.



DING Yao, born in 1981, senior engineer, is a member of China Computer Federation. His main research interests include digital image processing, computer vision and machine learning.