

基于 Kinect 的 3D 室内模型创建

康辉 陈晨 胡文博 吴可 金肖雅

摘要 三维重建是计算机视觉的一个重要目标,可以帮助人们快速精确地将日常生活中的物体数字化,并有着广泛的应用前景。本文提出了一种基于 Kinect 的成本低廉、快速精准且操作简便的三维重建方法。首先对 Kinect 彩色摄像机进行标定,获取摄像机的内参数矩阵,并将深度摄像机与彩色摄像机对齐;使用了一种新的基于联合双边滤波的帧间滤波算法,对原始深度图像进行该滤波处理;通过 PCL 获取点云,并利用 ICP 算法完成点云精确配准,得到拼接的三维点云图。实验结果表明,本文方法能够快速精准地创建 3D 室内模型。

关键词 Kinect; 三维重建; 帧间滤波; PCL; ICP 算法配准

3D-Object Modeling with Kinect in Indoor Condition

Kang Hui Chen Chen Hu Wenbo Wu Ke Jin Xiaoya

Abstract 3D Reconstruction is a vital target of Computer Vision, which can help people quickly and accurately to digital objects in daily life and which also has broad application foreground. This paper proposes a low cost, fast accurate and easy operation method of 3D Reconstruction, which is based on Kinect. First of all, the calibration of device color video camera to obtain the camera parameter matrix, and align the depth camera and color camera. Secondly, using a new kind of inter-frame filtering algorithm, which is based on the joint of bilateral filtering, to do the filter processing on the original depth image. Thirdly, through PCL accessing point cloud, and using ICP algorithm to complete point cloud registration precision, and in the end getting the splicing 3D point clouds. The experimental results show that the method can accurately create 3D indoor model quickly.

Key words Kinect; 3D reconstruction; Inter-frame filter; PCL; ICP registration

0 引言

伴随着三维技术的发展,各种三维技术已经静静走进人们的日常生产、生活和工作。目前三维重建技术应用已经相当广泛,诸如航天、造船、考古、工业测量、电子商务、逆向工程、计算机动画、显微摄像学等领域。鉴于三维重建广泛的应用前景,本文提出了一种基于 Kinect 的成本低廉、快速精准且操作简便的三维重建方法。

Kinect 由红外线发射器、红外摄像机和彩色摄像机组成,能同时获得深度和彩色图像,目前已有研究者将其应用于三维重建。基于 Kinect 的三维重建方法并不是一种独立的方法,而是与结构光法类似。

Whelan 等人^[1]针对 KinectFusion 方法能映射的重建范围局限于一个固定大小和方位的立方体中,不适用于 Kinect 移动一段较长距离的缺点研究了一种改进算法。

Zhang 等人^[2]为解决动态场景下的三

维重建问题提出了一种算法,仅用一个深度摄像机拍摄,使用 ICP 算法计算摄像机的姿态,利用过滤器和形态学操作处理追踪失败的点来达到分割前景和背景图像的目的。所得前景和背景图像分别用于前景和背景重建。

宋诗超^[3]对三维人体建模和测量方法进行了研究,用深度双阈值法对 Kinect 获得的深度图像进行深度信息提取,引入曲率阈值,根据曲率与该阈值的大小关系选择最小距离法或角度偏差法来合并点云,用三基准点对齐法对点云进行配准,用 Delaunay 算法进行表面重建。

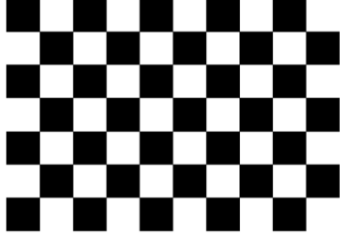
张翠红^[4]在利用 Kinect 进行室内场景的三维重建时,先提取出计算速度快、稳定性高的 SURF 特征,再用 ICP 算法进行配准,使用牛顿迭代法使得误差函数最小化。

目前基于 Kinect 的三维重建方法的精度仍然相对较低。而对深度图像的滤波算法和配准算法是影响三维重建效果的主要因

素。因此，如何提高深度图像滤波算法和配准算法的精度，就成为基于 Kinect 的三维重建研究的难点之一。

1 Kinect 相机标定^{[5][6][7]}

本文中选用黑白块交替排列的棋盘，这不仅可以保证待测物体的对称性，而且还可以自然地使用亚像素定位函数。棋盘格如下：



标定时，利用摄像机从不同方位对标定板进行拍摄，本文以 25 幅图像采集进行标定。

标定的实现是在 Matlab 环境下运行 Jean-Yves Bouguet 编写的定标程序得到的。我们利用 Camera Calibration Toolbox，首先运行 calib（或 calib_gui）命令，对 Kinect 相机拍摄的 25 幅图像进行定标，并将定标文件结果存储为 Calib_Results_Kinect.mat。

标定时采集的图像如下：

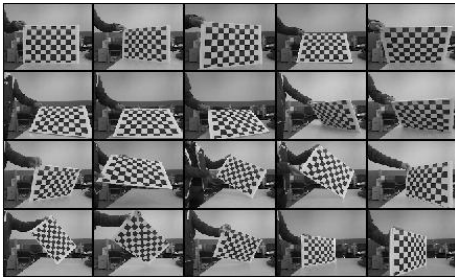


图 1.1 相机标定

通过对 Kinect 相机定标，可以得到 Kinect 相机的内参数矩阵。本文实验中得到的参数如下所示：

$$K = \begin{bmatrix} 512.65059 & 0 & 314.26860 \\ 0 & 514.67398 & 223.71073 \\ 0 & 0 & 1 \end{bmatrix}$$

2 Kinect 深度图像的滤波

由于 Kinect 获取的深度图像含有大量的噪声，所以必须对每一次获取的深度图像进行实时去噪处理。本文对几种不同的滤波方法进行对比分析，使用了一种新的基于联合双边滤波的帧间滤波算法，对原始深度图像进行该滤波处理。

2.1 高斯滤波

高斯滤波是一种线性的平滑滤波器，适用于消除高斯噪声，被广泛应用于图像去噪处理。从原理上来说，高斯滤波是一种改进的加权平均过程。根据像素点空间之间的相关性（即距离越近，相关性越大），与当前像素点距离越近的像素点的权值越大。其滤波核如公式 (2.1.1) 所示：

$$w(i, j) = \exp\left(-\frac{(i-x)^2 + (j-y)^2}{2\sigma^2}\right)$$

$w(i, j)$ 为滤波器在点 (i, j) 处的权值，为高斯函数的标准差，决定了高斯滤波的性能。

2.2 双边滤波

双边滤波是高斯滤波的一种改进算法，相对于高斯滤波最重要的一个特性是在滤波的同时保持图像边缘。双边滤波在考虑像素间空间相关性的同时，也考虑了像素值的相似程度。对于双边滤波，权值系数由空间域权值 w_s 和图像灰度域权值 w_r 组成，即

$$w = w_s \times w_r$$

其中：

$$w_s = \exp\left(-\frac{(i-x)^2 + (j-y)^2}{2\sigma_s^2}\right)$$

$$w_r = \exp\left(-\frac{(I(i, j) - I(x, y))^2}{2\sigma_r^2}\right)$$

σ_s 、 σ_r 为基于高斯函数的标准差，决定了双边滤波器的性能，限定了像素点的相对空间位置、灰度变化范围。

2.3 联合双边滤波^[14]

联合双边滤波器是 Petschningg 等在 2004 年提出一种双边滤波器改进算法，双边滤波器的权值都是由一幅待处理图像得到的，如果能提供另外一幅信息丰富的图像

计算得出权值效果会更好。算法首先对同一场景拍摄两幅照片，其中一幅打开了闪光灯，另一幅没有；接着将闪光灯打开拍摄的图像高频部分求出权值，作为引导和校正图像；最后合成未开闪光灯拍摄的图像低频部分，Petschningg 等的实验证明其算法比双边滤波效果更好。

Kinect 镜头同时采集 RGB 彩色图像和深度图像，完全可以利用彩色图像中完整的边缘信息对深度图像中缺失的部分进行补全，根据联合双边滤波的思想，本文将彩色图像作为引导图像，对深度图像进行校正补全，改进的公式如下：

$$w = w_s \times w_g$$

其中：

$$w_s = \exp\left(-\frac{(i-x)^2 + (j-y)^2}{2\sigma_s^2}\right)$$

$$w_g = \exp\left(-\frac{(G(i,j)-G(x,y))^2}{2\sigma_g^2}\right)$$

$G(i,j)$ 为彩色图像灰度化后在像素 (i,j) 处的灰度值。

在本文中，选取领域范围为 11×11 ， $\sigma_s = 4$ ， $\sigma_r = 4$ 对采集到的深度图进行联合双边滤波。

2.4 基于联合双边滤波的帧间滤波

为了提高深度图像时间和空间上的稳定性，减少噪声，本文采用如下方法：

(1) 在同一视角下对拍摄的 N 帧经过联合双边滤波后图像序列 D_i 中的每个相同像素点 (x,y) 的深度值取帧间滤波得到图像 $D(x,y)$ ：

$$D(x,y) = \underset{i=0,1,2,\dots,N}{middle} D_i(x,y)$$

(2) 将 N 帧图像中每个相同的像素点的深度值分别与帧间图像作差，当差小于一个阈值 T 且不为 65535，即认为改深度值为可靠点。当 N 帧图像中的相同像素可靠点数目 m 大于一个阈值 n 时，对这 m 个可靠点作均值得到结果：

$$D'(x,y) = \underset{\substack{m > n \\ |D_i(x,y) - D(x,y)| < T}}{avg} D_i(x,y)$$

本文选取阈值 T 为 20，可靠点阈值 n 为 9。本文的实验结果如下图所示：

图 2.1 原彩色图

图 2.2 原深度图

图 2.3 双边滤波后深度图

图 2.4 本文滤波算法滤波后深度图



图 2.1



图 2.2



图 2.3



图 2.4

由效果图可以看出，双边滤波对原始深度图像的缺失点弥补的作用不明显，而本文采用的算法对深度缺失点有较好的弥补作用。

3 点云配准

通过前文所作的工作，我们已经利用 Kinect 摄像机获得了深度图像，并且对深度图像进行了滤波处理，结合 RGB 图像得到了三维点云数据。下面就介绍三维点云数据配准的原理，然后利用 ICP 精确配准算法对获得的点云数据进行配准，并给出实验结果和分析。

3.1 配准原理

根据多个视角的深度图像可获得多个点云，但每个点云的数据都是基于各自摄像机坐标系的。这就需要对点云进行坐标变换^[8]，将各个点云转换到统一的世界坐标系下。本文针对刚性物体进行研究，不涉及尺度变换，故各点云的坐标变换可简单视为旋转变换和平移变换的合成。求解变换参数的过程即为配准。在配准阶段，需利用前一帧和当前帧的点云数据，并以前一点云的变换参数作为初始值，根据 ICP 算法计算得出变换参数。

ICP (Iterative Closest Point) 算法^[9]是配准领域的经典算法，最早是由 Besl 等人提出的。该算法以最小二乘法为基础，搜索最近点得到对应点对，以点对的点到点欧式距离之和作为目标误差函数。后续研究

者研究了不同的搜索对应点的策略，搜索目标由最近点演变为对应点，故 ICP 算法有时也称为迭代对应点算法。

3.2 利用 PCL 实现 ICP 算法

PCL 点云库^{[10][11][12][13]}能够对点云数据进行获取、滤波、分割、配准、可视化等处理。PCL 点云库封装了一系列配准算法，极大的方便了点云配准的学习和研究过程。下面是本文在 Intel(R) Core(TM) i5-3230M CPU 2.60GHz，双核四线程 CPU，4GB 内存的 Windows 7 x32 Ultimate 系统上利用 PCL 点云库中的算法以下采样体素大小 r （单位：米）和迭代次数 n 为自变量，时间 t （单位：秒）为因变量，分别对点云数据进行配准的效果图：

图3.1 配准前

图3.2 $r = 0.01$ $n = 10000$

图3.3 $r = 0.01$ $n = 1000$

图3.4 $r = 0.03$ $n = 10000$

图3.5 $r = 0.05$ $n = 10000$

图3.6 $r = 0.07$ $n = 10000$

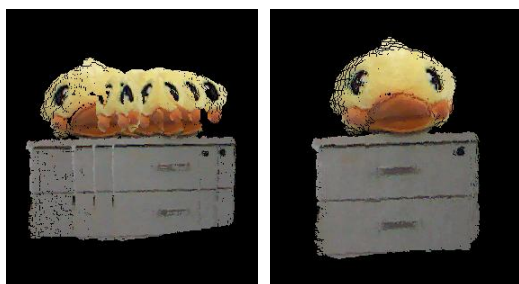


图3.1



图3.2



图3.3

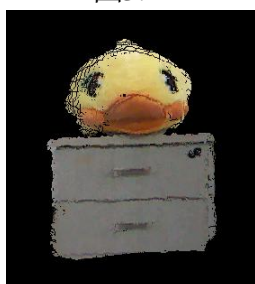


图3.4



图3.5



图3.6

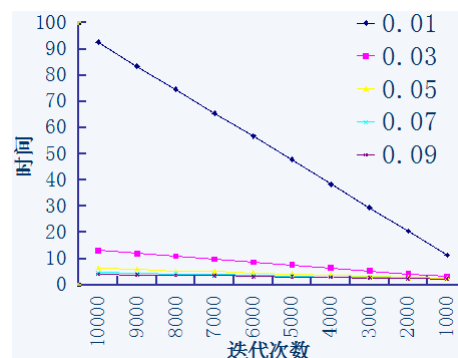


图3.7 r 、 n 、 t 关系图

由效果图可以看出配准的精确度和用时由下采样体素大小 r 和迭代次数 n 决定，可以看出当 r 较小 n 较大时配准效果较好，当 r 较大，即使 n 较大配准精度也较差。为了使配准精度和用时都在合理的范围内，我们最终采取 $r = 0.01, n = 1000$ 得到了较好的重建模型，而且该方法能够做到实时，自动化程度和鲁棒性比较高。

4 总结与展望

本文虽然较好的实现了 3D 室内模型的创建，但是还存在以下不足，还需进一步研究与改进：

1. 本文使用的滤波算法虽然效果较好，但是对于大片区域深度缺失的深度图像而言仍然无法较好的弥补缺失点。
2. 本文没有采取基于 GPU 编写的 ICP 算法，因此在实时性上还不足。

目前，Kinect 应用技术的开发还只是初见端倪，但已经表现出强大的生命力和多样性发展趋势。随着 Kinect 技术的进一步成熟和充分挖掘，相信下一代 Kinect 将会配备更强大的处理芯片，具有更高的分辨率和传输速率，能够实时捕捉人体细节，为人机交互带来质的突破，必将在更多的领域得到广泛地应用。

参考文献:

- [1] Whelan T, Johannsson H, Kaess M, et al. Robust real-time visual odometry for dense RGB-D mapping[C]//Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE 2013: 5724-5731.
- [2] Zhang J, Li Y, Li J, et al. Dynamic scenes reconstruction based on foreground and background splitting[C]//Wireless Communications&Signal Processing (WCSP),2013 InternationalConference on. IEEE, 2013: 1-4
- [3] 宋诗超.基于 Kinect 的三维人体建模与测量的研究[D].上海:东华大学, 2013
- [4] 张翠红.基于 Kinect 深度相机的室内三维场景重构[D].大连:大连理工大学, 2013.
- [5] 邱茂林, 马颂德, 李毅.计算机视觉中摄像机定标综述.自动化学报.2000,26(1): 43 — 55
- [6] R. I. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press. 2003
- [7] Z. Zhang. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000, 22(11): 1330-1334
- [8] Gelfand N,Mitra N,Eurographics Symposium Pottmann H,etal.Robust Global Registration. on Geometry Processing. 2005. 197-206
- [9] Besl P J, McKay H. D. A method for registration of 3-D shapes.Transactions on Pattern Analysis and Machine Intelligence, 1992,IEEE14(2):239-256
- [10] 朱德海.点云库 PCL 学习教程.北京:北京航空航天大学出版社, 2012
- [11] <http://www.pointclouds.org>
- [12] Point cloud library develop team.Point cloud library documentation[EB/OL].(2012)
- [13] <http://pointclouds.org/documentation/>
- [14] Johannes Kopf, Michael F. Cohen, Dani Lischinski, Matt Uyttendaele. Joint Bilateral Upsampling[J].ACM Transactions on Graphics. 2007, v.26 n.3
- [15] 基于 Kinect 的三维重建_范哲.西安电子科技大学.2013
- [16] 基于 Kinect 深度信息的实时三维重建和滤波算法研究_陈晓明,上海交通大学.2013