

单 Kinect+回转台的全视角三维重建

李健¹ 李丰¹ 何斌² 杜希瑞¹

(1. 陕西科技大学电气与信息工程学院, 西安, 710021; 2. 同济大学电气与信息工程学院, 上海, 201804)

摘要: 为了解决当前全视角三维扫描系统价格昂贵操作复杂的问题, 提出利用 1 台 Kinect 和 1 个回转台来构建全视角三维模型的方法。研究涉及点云预处理、点云配准、全局误差修正以及色差修正等技术。首先使用 Kinect 采集数据并预处理, 结合转台约束利用图像几何特征进行粗配准, 随后使用迭代最近点 (Iterative closest point, ICP) 算法实现点云的精确配准。对于累积误差导致的闭环问题以及不同角度拍摄引起的色差问题, 通过全局误差修正与色差修正算法处理, 提升重建结果的精度。实验结果表明: 该方法可以实现三维物体的全视角重建, 并在精度上优于微软的 KinectFusion 方法。

关键词: 全视角; 三维重建; 转台; Kinect

中图分类号: TP391 **文献标志码:** A

Single Kinect and Rotating Platform for Full-View 3D Reconstruction

Li Jian¹, Li Feng¹, He Bin², Du Xirui¹

(1. Department of Electrical and Information Engineering, Shaanxi University of Science and Technology, Xi'an, 710021, China;
2. Department of Electrical and Information Engineering, Tongji University, Shanghai, 201804, China)

Abstract: Traditional full-view 3D reconstruction systems are expensive and complex. To solve the above problems, a new algorithm for reconstructing an omnidirectional 3D scanning system is proposed based on a single Kinect and a rotating platform. The algorithm involves point cloud preprocessing, registration, minimizing global error between coordinate frames and color rectification. Firstly, RGB-D data are acquired by a Kinect sensor and preprocessed, and then a matching method is developed under the swivel table constraint for coarse registration. Further, the iterative closest point (ICP) algorithm is utilized for fine registration. Finally, for the loop closure caused by cumulative error and the color difference induced by different shooting angles, a global error correction and a color correction algorithm are conducted to improve the accuracy of the reconstruction results. Experimental results show that the reconstruction method can achieve full-view reconstruction of 3D objects and is superior to the KinectFusion method of Microsoft in accuracy.

Key words: full-view; 3D reconstruction; rotating platform; Kinect

引 言

三维重建是计算机中处理、分析和操作模型的基础, 同时也是虚拟现实、增强现实中的关键技术。

Kinect深度相机由于其成本低廉、操作简单和携带轻便的特性使其拥有广泛的应用场景,在近几年备受关注。

使用Kinect进行三维重建,已有大量的研究工作^[1-3]。文献[1]对在计算机视觉领域内使用Kinect作为传感器的算法与应用进行了综述。文献[2]介绍了Kinect在三维重建方面的研究工作。文献[3]针对Kinect V2设计了一种深度数据去噪算法,提高了重建质量。文献[4]基于Kinect结合旋转平台进行重建,得到了较好的重建效果,但其旋转平台和相机内参需要事先标定并进行离线的全局优化操作。文献[5]使用粗配准加精配准的策略对文物进行三维重建,但没有进行全局误差修正,相邻帧间的误差积累导致重建模型精度不高。文献[6]同时利用彩色信息和深度信息实现了室内场景的稠密三维建模,该算法虽能得到带纹理的三维模型,但得到的模型比较粗糙,且不适用于重建少纹理结构的物体。文献[7]使用4台Kinect构建了一套快速三维人体重建系统,系统内置了一个通用人体三维模型,将一个标准模型逐步迭代变形到扫描的人体点云上,该方法专门用于人体的三维扫描,通用性不够强。

目前, KinectFusion^[8-9]是使用Kinect在较低成本下采集物体三维模型的最常用方案。KinectFusion通过手持Kinect传感器围绕物体连续拍摄,可对具有纹理或弱纹理的物体形成光滑的稠密表面。因其使用固定尺寸的容积模型,在每个网格存储带符号的距离(Truncated signed distance function, TSDF)值及其权重,目前只能在GPU上才能实现实时的重建效果,且KinectFusion通过手持操作,对操作人员的要求较高。

本文使用Kinect作为传感器、回转台作为重建平台,为SURF (Speeded up robust features) 特征^[10]匹配点的筛选提供额外约束,提升了两帧之间的粗配准精度。经过迭代最近点(Iterative closest point, ICP)^[11]算法精配准后,得到的重建模型累积误差较小,通过平均误差分配方法^[12]便可基本消除。随后使用泊松曲面重建及色差修正使重建结果更加精细化。实验部分通过与标准模型和KinectFusion方法重建结果的对比,验证了本文算法在重建精度上的优势。

1 全视角三维测量系统组成及算法框架

本文基于Kinect+回转台的全视角三维测量系统如图1所示。将待测物体置于转台上, Kinect置于物体前方,测量时转台旋转, Kinect间隔一定时间采集一帧图像,待物体旋转一周后采集结束。本文方法相邻帧间的旋转平移计算不依赖于转台的旋转角度求解,故对转台旋转速度无特殊要求。

为了实现对被测目标的全视角三维测量,建立如图2所示的算法框架,下面对各个主要步骤做一简要说明:

(1)点云预处理:对Kinect获得的彩色图像和深度图像进行有效区域分割、噪声去除以及三角网格化。

(2)基于图像特征的粗配准和ICP精配准:由于ICP算法对初始值的要求较高,对相邻两帧的RGB图像求取SURF特征并匹配,然后利用转台约束对SURF匹配点进行提纯,根据提纯后的匹配点计算两帧之间的旋转平移矩阵,作为ICP配准的初始值,在此基础上对相邻帧点云进行ICP精配准。

(3)全局误差修正与色差修正:相邻帧点云两两配准后,由于误差累积使得第一帧和最后一帧点云不能很好地重合在一起,需要将所有帧进行全局误差修正以使其达到更高的配准精度,通常称之为闭环优化。随后使用泊松重建^[13]对点云进行融合以剔除重建模型中的多余点。利用梯度连续性对不同

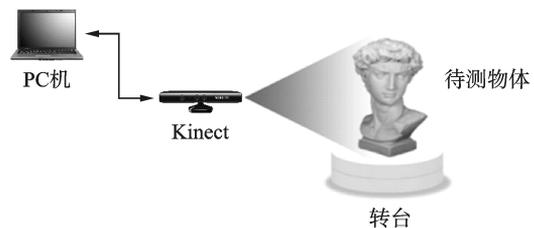


图1 测量系统组成

Fig.1 Measuring system components

角度拍摄的点云进行色差修正,最终得出一个较理想的重建结果。



图2 全视角三维重建算法框架

Fig.2 Full-view 3D reconstruction algorithm framework

2 算法详细描述

2.1 点云的预处理

点云的预处理主要包含3方面的内容:有效区域分割、噪声去除以及三角化。

(1)有效区域分割:有效区域分割分为物体有效区域分割与深度阈值过滤。物体有效区域分割,即只选择包含待重建物体的矩形区域进行后续处理。随后使用深度阈值过滤的方法进行深度有效区域选择,即认为一定深度范围内的点云为有效数据,将不处于有效范围内的点云直接删除,实验证实这一方法简单有效,并能有效缩减点云数据文件的大小,减少重建时间。

(2)噪声去除:Kinect获取的点云数据,尤其是边缘处的不稳定噪声点,会使后续的三维模型重建精度受到严重影响。使用常见的去噪方法不能很好地消除这些噪声。本文使用式(1)所示的数学形态学模板对点云进行腐蚀操作,然后再使用双边滤波^[14]方法对点云进行平滑处理。

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (1)$$

(3)三角化:通过区域分割和去噪操作,可以得到质量较高的点云,但因色差修正算法需要在三角化的基础上才能执行,因此这里对其进行三角网格化。由于Kinect获得的点云为具有图像结构的有序点云,所以采用三角剖分方法^[15]生成网格曲面,相比于现有的贪婪三角形等算法,三角剖分算法效率更高,应用于Kinect采集到的点云时效果更好。

2.2 相邻帧间配准

2.2.1 使用回转台约束的SURF粗配准

由于相邻两帧待配准点云间的初始位置对配准的精度和效率影响较大,为了避免陷入局部最优,这里在使用ICP方法进行两帧配准之前需要先进行粗配准。由于Kinect可以获得深度图像对应的RGB图像,因此学术界常常使用图像特征匹配的方法,如使用SIFT^[16](Scale-invariant feature transform)、SURF等进行点云的粗配准。但图像特征检测的方法存在两个缺陷:一方面,当重建目标是弱纹理物体时,特征检测方法常常失效;另一方面,当重建物体存在重复的纹理结构时,误匹配严重,即使采用随机抽样一致(Random sample consensus, RANSAC)算法提纯,也往往不能找到足够多正确的匹配点对。

本文利用转台解决相邻帧之间图像SURF匹配点的提纯问题。SURF算法的基本流程可以分为3大部分:特征点提取、特征点描述和特征点匹配。本文根据转台约束,物体仅围绕一个轴转动,误匹配

可以通过限制匹配点间的坐标来减少,即匹配的两个点的图像坐标的 y 值(竖直方向的坐标)之差必须在一定的范围内,大于此范围的就是无效匹配。图3(a)中从左连接到右下方向的连线表示错误的匹配点对,使用此方法剔除无效匹配后的结果如图3(b)所示。

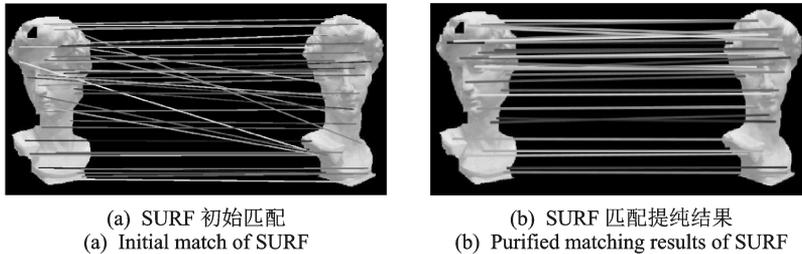


图3 转台约束下的SURF匹配点提纯

Fig.3 SURF matching point purifying under turntable

根据相邻帧间的匹配点,粗配准就是要找到一个旋转平移变换矩阵,使得匹配点对之间的欧式距离最短。这里采用优化代价函数的方式来实现这一目的,设计的代价函数如式(2)所示。

$$\min_{R,T} \sum \|R \cdot P + T - P'\|_2^2 + \lambda_{\text{rot}} E_{\text{rot}}(R) \quad (2)$$

式中: P 为当前帧已匹配图像特征点对应到点云中的三维坐标; P' 为上一帧中对应匹配特征点的三维坐标。其中第一项表示当前帧的匹配点在经过旋转变换 R 与平移变换 T 后与上一帧中匹配点的欧式距离,为了确保 R 是一个旋转矩阵,这里加入代价函数第二项,用来表征 R 必须满足的6个条件,即 R 的每一列必须是单位向量,并且各列之间是相互正交的,表示为

$$E_{\text{rot}}(R) = (c_1 \cdot c_2)^2 + (c_1 \cdot c_3)^2 + (c_2 \cdot c_3)^2 + (c_1 \cdot c_1^{-1})^2 + (c_2 \cdot c_2^{-1})^2 + (c_3 \cdot c_3^{-1})^2 \quad (3)$$

式中 c_1, c_2, c_3 为矩阵 R 的3个列向量。式(2)中 λ_{rot} 根据经验取值85。

2.2.2 ICP精配准

在粗配准所提供的初始匹配的基础上,精配准能够进一步改进和完善配准结果。精配准分为基于KD-Tree的关键点邻域搜索和ICP计算变换关系两个步骤。

(1)基于KD-Tree的关键点邻域搜索

粗配准步骤结束后,可以直接通过ICP算法计算点云的变换关系,但在已经得到较为精确的对应关系的前提下,可以通过KD-Tree搜索的方式选择对应点邻域来计算相对位置关系以节省计算成本。同时使用KD-Tree搜索的方式也可以选择已匹配点对 R 半径内的点云来计算最终的变换关系矩阵。

在具体的搜索过程中,将特征点作为查询点输入,给定查询点和查询距离的阈值(即查询半径 R),从点云找出所有与查询点距离小于阈值的点。如图4所示,以空间中一点 P 为圆心,给定的查询距离 R 为半径画球体,球体内点与 P 的距离小于半径 R 的点即为所要查找的邻域点。

(2)ICP计算变换关系

通过 R 半径搜索得到的对应区域是整体模型的一部分。只计算该部分点云的变换关系能够节约计算成本,将计算得到的 R, T 矩阵应用到原始模型上即可得到最终配准结果。

ICP算法是一种经典的基于自由形态曲面的配准方法,以四元素组为基础,为待配准目标点云中的三维空间点 $p_i = (x_i, y_i, z_i)$ 在参考点

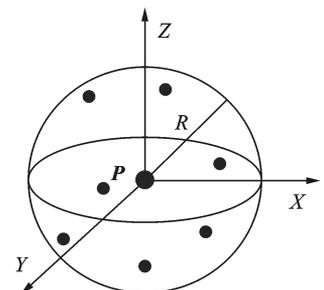


图4 R 半径搜索示意图

Fig.4 Sketch map of radius search

云中搜索对应的最近点 $q_i = (x_i, y_i, z_i)$, 三维点云的配准问题可转化为求解使式(4)中目标函数最小时的 R 和 T 的问题。

$$\operatorname{argmin}(R, T) = \sum_{i=1}^m \|Rp_i + T - q_i\| \quad (4)$$

式中: m 为点集中对应点对的个数; $\|Rp_i + T - q_i\|$ 表示 p_i 点经过 R, T 变换后到 q_i 点的欧氏距离。

2.3 全局误差修正与点云融合

2.3.1 全局误差修正

在完成了相邻帧间的两两配准后,可以得到一个坐标统一的全视角模型,通常情况下由于误差累积,第一帧和最后一帧的点云都不能很好地重合,需要对这个累积误差进行修正,以得到更为理想的重建结果。这里采用文献[12]中提出的闭环约束下的全局误差修正方法,以相邻两帧间的配准精度作为权重系数,将整体的误差平均分配到各个局部配准上,以实现全局误差的修正,其主要思想如下:

(1) 闭环约束,即所有的变换矩阵依次乘积的结果应该是单位矩阵

$$\hat{P}_{i,i+1} \cdot \hat{P}_{i+1,i+2} \cdot \dots \cdot \hat{P}_{n,1} \cdot \hat{P}_{1,2} \cdot \dots \cdot \hat{P}_{i-1,i} = I \quad (5)$$

式中 $\hat{P}_{i,i+1}$ 表示第 i 帧到 $i+1$ 帧的旋转平移变换矩阵。

(2) 正则项,即优化后的变换矩阵 $\hat{P}_{i,j}$ 和优化前的变换矩阵 $P_{i,j}$ 之差应尽量小

$$\operatorname{argmin}(\hat{P}_{i,j}) \sum \omega_{ij}^2 \|\hat{P}_{i,j} - P_{i,j}\|_2^2 \quad (6)$$

式中权系数 $\omega_{i,j} = 1/\operatorname{Dist}(P_{i,j}(C_j), C_i)$, 表示点云 C_j 经过变换 $P_{i,j}$ 后与点云 C_i 上最近点之间的平均距离的倒数,这个权系数越大,说明经过变换后的配准程度越高,在代价函数中这两帧的贡献较大。优化过程中这两帧的变换矩阵相较于其余帧更倾向于保持不变。求解在式(5)约束下式(6)的最小二乘解,得到全局优化后相邻帧间的旋转平移变换。

2.3.2 点云融合与色差修正

由于相邻帧间有很大一部分重合区域,故配准后的点云中存在许多重复的无用点,就需要对点云进行简化,即点云融合。这里使用泊松重建方法^[13]来进行点云融合。泊松曲面重建方法通过求解泊松方程来获得点云描述的曲面所代表的隐性方程。其求解采用隐性拟合的方式,重建结果为一封闭的曲面,具有较高的细节特征,并且能有效地屏蔽噪声的影响,鲁棒性高。

由于模型是由不同帧拼接起来的,而不同帧之间由于光照等因素的影响使物体上相同点的颜色不尽相同,这里就需要对不同帧之间的色差进行修正。使用文献[17]中提出的算法来进行色差修正,主要思想是首先计算出点云中的颜色梯度,然后利用梯度连续性求解泊松方程,最终得出一个颜色分布均匀的表面。

3 实验及结果分析

3.1 实验条件

搭建如图5所示的实验系统用以验证上述重建方法。测量系统由1台PC机,1台Kinect V2传感器,1个旋转平台组成。测量时被测物置于转台上,转台以约1 r/min的速度旋转,Kinect每隔4 s采集一帧图像(包



图5 测量系统

Fig.5 Measuring system

括彩色图和深度图),采集到的数据经由USB 3.0高速接口传输至PC机,PC机在接收到数据后同时完成图像预处理并将数据暂时保存在硬盘上。

测量在自然光照条件下进行,待测物体距离Kinect 2 m左右。PC机配置为i7-4720HQ处理器,NVIDIA GTX 960M显卡,机身内存8 GB,支持USB 3.0接口。使用采集程序连续拍摄约1 min,得到了连续的15帧点云数据,如图6所示。

3.2 点云预处理

点云预处理结果如图7所示,图7(a)为Kinect采集到某一帧的原始深度图,图7(b)表示对其进行有效区域分割之后的点云,可以看到,背景已被有效去除。图7(c)为去噪之后的结果,进而进行双边滤波的结果如图7(d)所示,对双边滤波后的点云进行三角化后的结果如图7(e)所示。

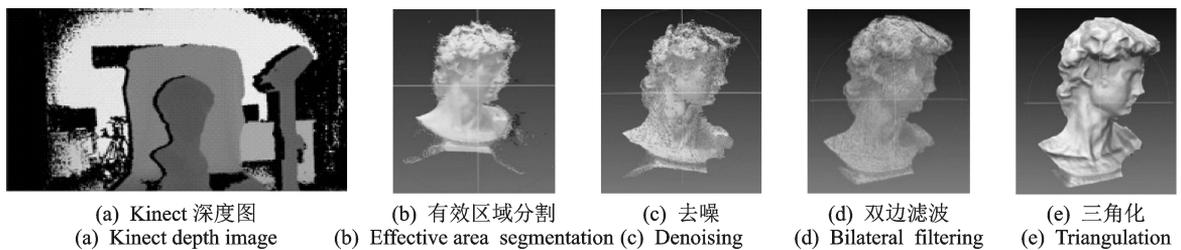


图7 点云预处理

Fig.7 Point cloud preprocessing

3.3 两帧配准

相邻帧之间的配准由SURF粗配准+ICP精配准进行,图8(a,b)中两种不同颜色分别表示相邻的两帧点云。图8(a)为SURF粗匹配的结果,图8(b)为ICP精配准后的结果。对比图8(a)与图8(b)可以看到经过粗配准两帧点云已基本重合,精配准进一步提高了配准的准确度。

3.4 全视角三维重建

相邻帧间配准后,便可以得到一个坐标统一的全视角模型,图9(a)为将各帧点云拼接起来的结果。由于误差累积的原因,第一帧和最后一帧的点云没能很好地重合,如图9(b)所示,在大卫像的侧面(红色矩形框部分)出现了明显的缝隙。使用全局误差修正方法对此误差进行修正,修正后的点云如图9(c)所示,可以看到此缝隙已观察不到,整体配准效果较理想。通过泊松重建融合了点云后的模型如图9(d)所示,赋予颜色后的效果如图9(e)所示,可以看出由于模型由不同



图6 采集到的点云数据

Fig.6 The captured point cloud data

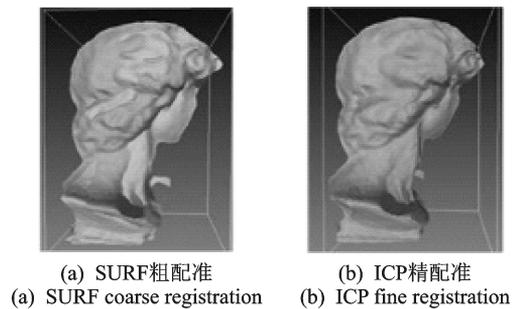


图8 SURF粗配准与ICP精配准
Fig.8 SURF coarse registration and ICP fine registration

帧拼接,但是不同帧之间由于光照等因素的影响,使得物体上相同点在不同帧上的颜色不尽相同,这里就需要累积之后的色差进行修正。色差修正后的模型如图 9(f)所示,整体颜色过度平滑,局部的亮斑或暗斑基本消除。

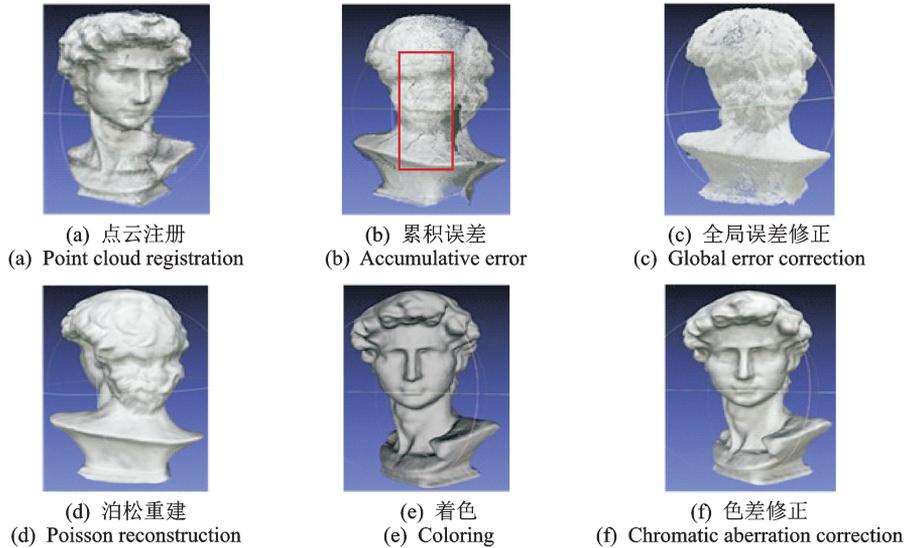


图 9 全视角三维重建

Fig.9 Full-view 3D reconstruction

3.5 误差对比

为了说明本文全视角重建方法的精度,比较了本文方法与标准模板以及 KinectFusion 之间的误差。首先将本文方法重建的模型与标准模板进行配准,然后测量标准模板与测得模型上最近点的距离,并用伪彩色图像来表示这个距离的大小,结果如图 10 所示。图 10(a)为石膏像标准模板在 4 个不同视角下的图片,图 10(b—d)为本文重建方法得出的结果及其与标准模板间误差的伪彩色表示图,图 10(e—g)为 KinectFusion 重建方法得出的结果及其与标准模板间误差的伪彩色表示图。与标准模板比较本文的平均误差为 4.1 mm,而 KinectFusion 重建结果的平均误差为 4.6 mm,并且在具有多帧点云的重叠数据区域 KinectFusion 也存在着较大的重建误差,如石膏像的头部。另外在石膏像的后脑部分的重建效果上,本文的重建方法相较于 KinectFusion 方法拥有更多的细节,更接近于真实数据。分析其原因,是因为 KinectFusion 使用 TSDF 值将深度估计值和测量值进行融合,导致部分信息被平滑掉了,而本文方法由于采用了 SURF+ICP 的配准方法,拥有更精确的配准结果,故能够保留更多的细节信息。但由于 Kinect 自身分辨率较低,仅使用 Kinect 在重建细节上难以达到标准模型的精细程度,例如头部头发的重建效果。

4 结束语

本文研究了基于 Kinect 和转台的全视角三维测量方法,包括 Kinect 点云数据预处理、转台约束下的 SURF 特征提纯、点云配准、累积误差修正以及色差修正等。文中利用 C# 和 Matlab 分别编写了采集程序和建模程序,实现了所提算法,并且以大卫像为测量对象验证了算法的可行性。通过与标准模型和 KinectFusion 方法进行误差比较,表明所提算法重建结果理想、实验过程简单、不需要事先对转台系统进行标定。但由于 Kinect 自身分辨率较低,仅使用 Kinect 深度数据在重建细节上还不能达到标准模型

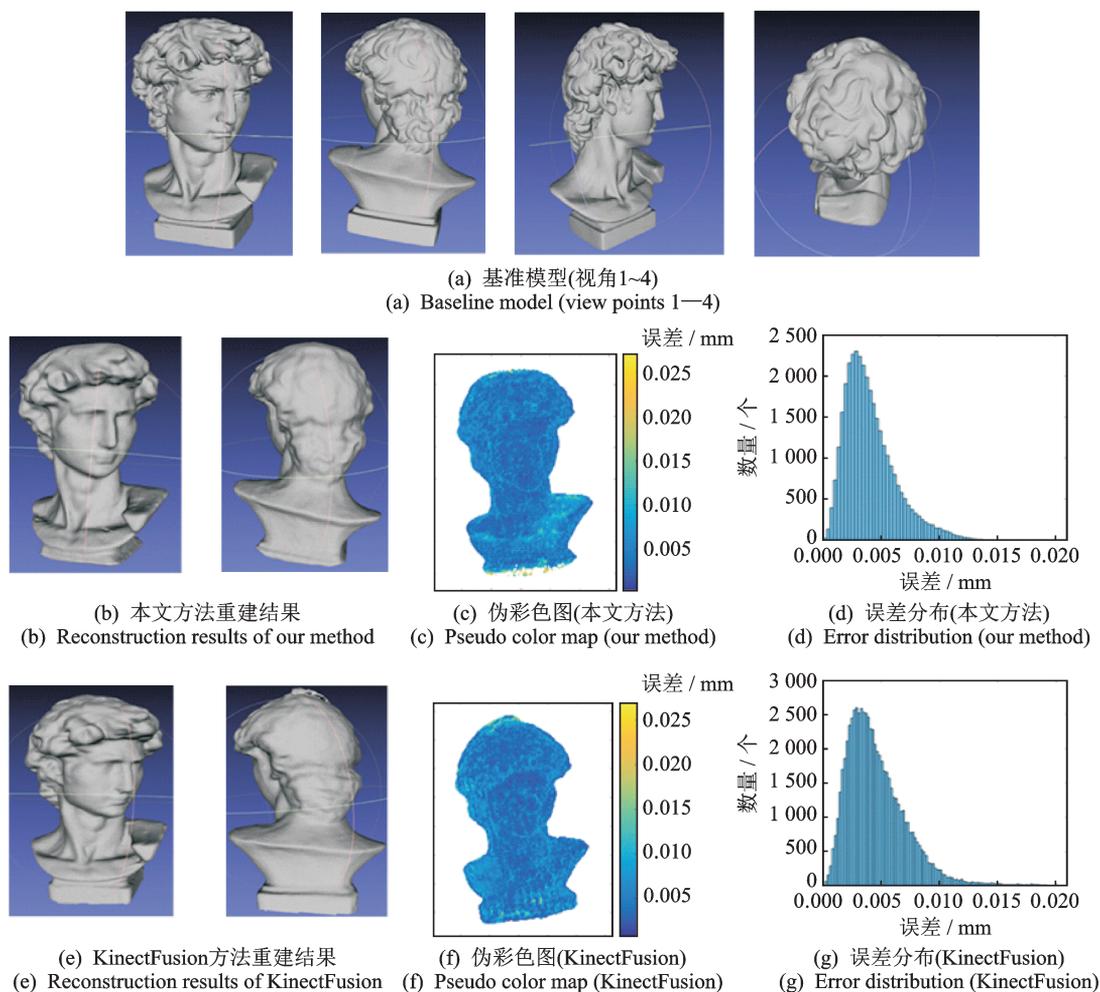


图10 本文方法与KinectFusion方法重建结果比较

Fig.10 Comparison of our method with the KinectFusion method

的精细程度,下一步工作将考虑 Kinect与明暗恢复形状、光度立体等基于光照约束的重建方法融合以进一步提升重建质量。

参考文献:

- [1] Han Jungong, Shao Ling, Xu Dong, et al. Enhanced computer vision with microsoft Kinect sensor: A review[J]. *IEEE Transactions on Cybernetics*, 2013,43(5): 1318-1334.
- [2] Li Shirui, Tao Kelu, Wang Siyuan, et al. 3D reconstruction by Kinect sensor: A brief review[J]. *Computer Aided Drafting, Design and Manufacturing*, 2014,24(1): 1-11.
- [3] 李诗锐,李琪,李海洋,等. 基于 Kinect V2的实时精确三维重建系统[J]. *软件学报*, 2016,27(10): 2519-2529.
Li Shirui, Li Qi, Li Haiyang, et al. Real-time accurate 3D reconstruction based on Kinect V2[J]. *Journal of Software*, 2016,27(10): 2519-2529.
- [4] Liu Xin, Xu Huarong, Hu Zhanyi. GPU based fast 3D-object modeling with Kinect[J]. *Acta Automatica Sinica*, 2012, 38(8): 1288-1297.
- [5] 孙家泽,陈皓,耿国华. 三维文物点云模型配准优化算法[J]. *计算机辅助设计与图形学学报*, 2016,28(7): 1068-1074.

- Sun Jiaze, Chen Hao, Geng Guohua. Registration optimization algorithm for 3D cultural relics point clouds model[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(7): 1068-1074.
- [6] Henry P, Krainin M, Herbst E, et al. RGB-D mapping: Using depth camera for dense 3D modeling of indoor environments[J]. The International Journal of Robotics Research, 2012, 31(5): 647-663.
- [7] 周瑾,潘建江,童晶,等. 使用Kinect快速重建三维人体[J]. 计算机辅助设计与图形学学报, 2013, 25(6): 873-879.
Zhou Jin, Pan Jianjiang, Tong Jing, et al. Fast 3D human body reconstruction from Kinects[J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(6): 873-879.
- [8] 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: 559-568.
- [9] Newcombe R, Izadi S, Hilliges O, et al. KinectFusion: Real-time dense surface mapping and tracking[C]//The IEEE International Symposium on Mixed and Augmented Reality. Basel, Switzerland:IEEE, 2011: 127-136.
- [10] Bay H, Tuytelaars T, Van G L. SURF: Speeded up robust features[C]//Computer Vision-ECCV. Graz, Austria: Springer, 2006: 404-417.
- [11] Paul J B, Neil D M. A method for registration of 3D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [12] Sharp G C, Lee S W, Wehe D K. Multi-view registration of 3D scenes by minimizing error between coordinate frames[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(8): 1037-1050.
- [13] Kazhdan M, Hoppe H. Screened poisson surface reconstruction[J]. ACM Trans Graph, 2013, 32(3): 29-1-29-13.
- [14] Tomasi C, Manduchi R. Bilateral filtering for gray and color images[C]//Sixth International Conference on Computer Vision. Bombay, India: IEEE Computer Society, 1998: 839-846.
- [15] Lee D T, Schachter B J. Two algorithms for constructing a delaunay triangulation[J]. International Journal of Computer & Information Sciences, 1980, 9(3): 219-242.
- [16] Lowe D G. Object recognition from local scale-invariant features[C]//Proceedings of the Seventh IEEE International Conference on Computer Vision. [S.l.]: IEEE, 1999, 2: 1150-1157.
- [17] Chuang M, Luo L J, Brown B J, et al. Estimating the Laplace-Beltrami operator by restricting 3D functions[J]. Computer Graphics Forum, 2009, 28(5): 1475-1484.

作者简介:



李健(1975-),男,教授,研究方向:图形图像处理、数据挖掘与机器学习和网络与信息安全等,E-mail: lijianjs@sust.edu.cn。



李丰(1994-),男,硕士研究生,研究方向:三维重建和构图。



何斌(1975-),男,教授,研究方向:新型机器人动力学及智能控制、多种传感器技术与数据网络集成应用和图像检测与识别技术及应用等。



杜希瑞(1991-),男,硕士,研究方向:计算机图形学与三维重建。

(编辑:张黄群)