

## 基于三个 Kinect 的个性化人体重建

乔思航, 程志全, 陈寅, 党岗  
(国防科学技术大学计算机学院, 长沙 410073)



**摘要:** 个性化人体重建是在虚拟场景中建立用户个性化的三维几何模型。使用三台微软新推出的 Kinect 深度摄像机完成个性化人体的重建, 工作流程如下: 首先, 使用厚度不计的褶皱纸张, 完成三台 Kinect 的几何标定; 其次, 三台 Kinect 同时从三个视角自动扫描待建静态人体对象。借助 Kinect 的摄像机运动控制功能, 每个摄像机自动地从上向下获取人体对象的深度点云; 再次, 借助于相机的几何标定矩阵, 自动完成三个点云的注册; 最后, 重建出模型表面。实验结果表明, 能重建出高逼真度的个性化三维人体模型。与同类工作相比, 具有较好的易用性, 无需转盘等设备, 更容易搭建系统; 此外, 获取时间 3 秒左右, 用户无需长时间保持静止状态, 友好性更强。

**关键词:** 人体建模; Kinect; 几何标定; 表面重建

中图分类号: TP391.9 文献标识码: A 文章编号: 1004-731X (2013) 10-2408-04

### Personalized Full-body Reconstruction Using Three Kinect

QIAO Si-hang, CHENG Zhi-quan, CHEN Yin, DANG Gang

(Department of Computer, National University of Defense Technology, Changsha 410073, China)

**Abstract:** Personalized full-body reconstruction builds 3D geometry shapes, approximating the personalized users, in the virtual scenes. A system could reconstruct personalized full-body shapes by using three Microsoft Kinect depth cameras. The system works in four steps: *Firstly, these three Kinect are geometrically registered by using a collapsed paper. Secondly, the static subject is captured by each camera, which automatically rotates its orientation by the embedded motion driver. Thirdly, three point clouds captured by the cameras are registered together using the former geometrical registration information. Finally, the surface is automatically reconstructed.* The experiments on various human subjects demonstrate that the system could build realistic personalized full-body models. Compared to the related work, the system has such advantages: easy configuration without turnover devices, nice user friendship without user long-time static pose holding.

**Key words:** full-body reconstruction; Kinect; geometrical calibration; surface reconstruction

### 引言

个性化人体重建是指在虚拟场景中建立与用户外形相近的三维几何模型。它能够建立起用户的虚拟人模型, 将用户参与程度提升到一个新阶段, 能推动虚拟人在动画、

影视制作和虚拟现实等领域的迅猛发展。

本文使用低成本的设备<sup>[1]</sup>完成个性化人体的捕获与建模任务, 实现了一种基于三个 Kinect 的个性化人体建模系统, 该系统的主要特色包括:

- 1) 低成本: 该系统使用低价位的 Kinect, 2013 年的市场价格在 1300 元左右;
- 2) 易用性: 该系统无需转盘等设备, 容易搭建, 此外用折叠的纸张自动完成相机的几何标定;
- 3) 友好性: 不再受严苛的实验室捕获条件的限制, 应用场所走出了实验室。更重要的是, 用户无需保持长时间的静止状态, 3 秒钟左右即可完成数据的捕获;

收稿日期: 2013-05-31 修回日期: 2013-07-12

基金项目: 国家自然科学基金 (61103084)

作者简介: 乔思航(1990-), 男, 陕西三原人, 硕士生, 研究方向为计算机图形学; 程志全(1977-), 男, 黑龙江五常人, 博士, 研究方向为计算机图形学等; 陈寅(1986-), 男, 江苏扬州人, 博士生, 研究方向为计算机图形学等; 党岗(1973-), 男, 陕西韩城人, 博士生, 副研, 研究方向为虚拟现实技术等。

http: www.china-simulation.com

• 2408 •

4) 功能性: 生成的人体外形与用户的真实数据非常逼近, 所有测试用例的实验结果身高误差不得超过 1 厘米, 胸围误差不得超过 2 厘米。

## 1 相关工作

传统三维建模扫描技术通常使用结构光扫描仪或者激光扫描仪获得高精度数据, 用以生成高真实感的三维虚拟人个性化模型<sup>[2]</sup>。但是, 这类三维扫描设备大都极其昂贵。例如, Cyberware 人体扫描系统<sup>[3]</sup>市面报价为 24 万元, 普通用户很难承受如此价格。此外, Cyberware 等设备均要求用户保持一定时间的静止状态, 如 Cyberware 的捕获时间为 17 秒。众所周知, 用户是很难在长时间内保持完全静止。

Kinect<sup>[1]</sup>是微软公司在 2010 年推出的一款新型摄像机, 能够捕获深度点云数据。它具有非接触式、成本低(不依赖特殊设备)和智能化等优点, 为三维人体模型的生成奠定了硬件基础。但是, Kinect 扫描得到的原始数据是不完整、充满噪声且没有注册在一起的。在后续开发中, 微软进一步推出了面向 Kinect 的 Kinect Fusion 系统<sup>[4]</sup>。该系统利用一台围绕目标物体移动的 Kinect 实时捕获物体三维模型的信息, 通过整合来自 Kinect 深度摄像机采集到的连续数据流, 重建出一个物体的三维模型。尤其是, 随着 Kinect 的移动, 被重复扫描到的位置精度会逐渐加强。但是, 使用一台 Kinect, 快速获取保持静态的人体是非常具有挑战性的。实验表明, 一台手持移动的 Kinect 使用 Kinect Fusion 完成扫描至少需要 40 秒以上的时间。生理学的最新研究结果<sup>[5,6]</sup>表明, 3 秒是一个能够自然保持“假死”静止的时间段。因此, 使用一台 Kinect 完成静态人体的扫描是不可能完成的任务。

Tong et al.<sup>[7]</sup>借助了三台 Kinect 完成人体的捕获。但是, 该系统需要大约 30 秒的数据采集时间, 而且, 文献<sup>[7]</sup>的系统还需要转盘等设备。这严重限制了重建系统的适用范围。

本课题组的前期工作<sup>[5]</sup>使用了两台 Kinect 完成了 3 秒内人体的捕获。但由于人体的自遮挡问题, 两台 Kinect 所捕获的数据之间有显著的空洞, 会导致重建结果失真。此外, 由于 Kinect 距离人体偏远, 重建出的模型质量精度还不够高。为此, 本文使用了三台 Kinect, 搭建了更鲁棒更易用的个性化人体捕获系统。

## 2 捕获系统

本文的方法在搭建好实验环境后, 利用一张褶皱的纸张可以完成 Kinect 相机的几何标定, 依赖 Kinect Fusion 技术和 Kinect 摄像机的运动控制功能可以完成点云的自动捕

获, 然后利用相机几何标定中获得的三台 Kinect 相对位置信息可以将人体三个侧面点云信息注册成完整的点云模型, 最后对点云模型进行表面重建即可完成人体模型三维重建的目的, 如图 1 所示。

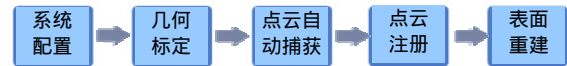


图 1 捕获系统流程图

### 2.1 系统配置与几何标定

当三台 Kinect 深度摄像机放置好并对准方向后, 用户可以任意姿态站在三台 Kinect 中间, 距离三台 Kinect 1.5~3 米, Kinect 的高度为用户身高的一半左右, 如图 2 所示, 这样的简单布置后, 实验环境搭建完成。数据采集过程中, 用户只需要保持姿势静止 3 秒, 三台 Kinect 就会自动自上而下转动视角, 使 Kinect 垂直方向视角从 43°扩大到 97°, 完成整个人体的数据采集。

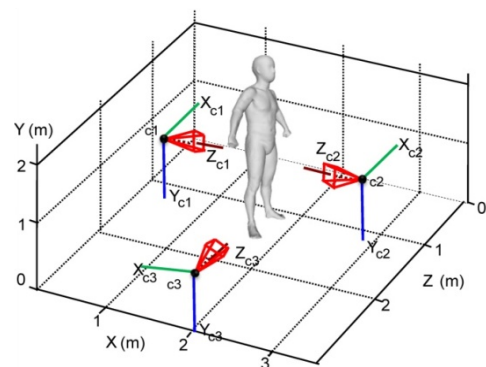


图 2 实验场景示意图

几何相机标定的目标是: 对于任何一个三维空间内的点, 通过三台 Kinect 相机捕捉到的位置信息可以计算出在一个统一坐标系下该点的三维坐标。与计算机图形学领域传统使用的多相机标定方法不同的是, 本文充分利用了多 Kinect 采集到的深度数据和几何信息来进行相机标定。本文使用三个 Kinect 采集模型信息, 但是标定过程采用两两

标定的方法, 选定一个 Kinect 作为参考, 另外两台 Kinect 分别与这台 Kinect 标定, 从而将三台 Kinect 的坐标系统一起来, 实现拼接三台 Kinect 采集到的三维模型的目的。

首先, 让两台 Kinect 捕捉同一目标, 本文使用一张揉成一团再打开的有褶皱的 A4 纸作为标定过程的目标物体。如图 3 所示。选择褶皱的纸张的目的在于, 纸张厚度很薄可以忽略不计, 使得两台标定的 Kinect 观察纸张的两面可以对齐的很好, 而且纸上有很多随机的褶皱, 丰富的几何细节使得标定计算更加准确。几何标定使用的算法是 Cheng et al. 2013<sup>[8]</sup>提出的 SuperMatching 算法, 通过该算法可以将两台 Kinect 从两面扫描到的目标纸张上丰富的特征提取出来并求出特征点的对应关系, 利用这些对应的特征点匹配关系, 可以计算出两幅标定数据的旋转角度即旋转矩阵, 即得到了一台 Kinect 相对于另一台的标定矩阵。



图 3 用做标定对象的褶皱的 A4 纸

## 2.2 点云的自动捕获

传统方法, 如 Chen et al.<sup>[7]</sup> 采用上中下三个角度采集三组数据, 再将其拼成一副完整的三维点云数据, 这种方法复杂, 不能自动完成, 需要后期处理。本文借助 Kinect 摄像机运动控制功能, 自动的上下移动 Kinect 视角, 依赖 Kinect Fusion 技术自动建立用户模型<sup>[8]</sup>, 可以在 3 秒内完成单 Kinect 三维数据的捕捉。目前的 Kinect Fusion 技术, 能够接受来自移动中的 Kinect 深度摄像机捕获的数据并实时地创建高质量三维模型。

## 2.3 点云注册

点云注册的目的在于将三台 Kinect 从不同角度扫描到的用户三维模型的三个侧面数据正确的拼接在一起, 形成完整的人体三维模型。在 2.1 中通过[9]的 SuperMatching 算法已经获得了 Kinect 两两之间的标定矩阵, 依赖于这两个标定矩阵, 可以分别把另两个 Kinect 扫描到的两个侧面人体点云模型信息变换到基准 Kinect 采集到的人体点云模型上, 从而成功的将三组人体不同角度的点云信息注册到

一起, 形成完整的用户三维点云模型。

## 2.4 表面重建

得到用户三维点云数据后, 我们采用泊松重建的方法惊醒模型表面的重构。泊松重建可以全局地考量所有采样点, 不需要分割或者融合, 该算法有较好的鲁棒性。本文利用 MeshLab 平台提供的 Poisson Reconstruction 功能, 重建出的用户三维人体模型非常完整, 没有明显缺陷。

## 3 实验

本文使用 VS2010 下的 C++语言实现该系统, Kinect 数据采集是在 Kinect SDK1.7 版本下的 Kinect Fusion 基础上做优化后进行的, 使用的是 Kinect SDK1.7 的驱动程序。本文系统在 ThinkPad W520(Intel Core i7 CPU @2.20GHz (8 CPUs)处理器, 4GB 内存, Quadro 1000M 2GB 显存显卡)笔记本上测试运行。

图 4 展示了我们重建出的两个用户的三维模型, 左边的是 Kinect 采集的 RGB 图像中间的 Kinect Fusion 自动捕获的一个角度的点云, 右边的是重建出的完整虚拟人三维模型。在时间开销上对每个十几兆左右大小的模型数据而言, 算法各部分处理平均时间见表 1, 其中采集数据时间 3 秒远低于[7]的 30 秒, 数据处理时间 16 秒低于[7]的 30 秒。

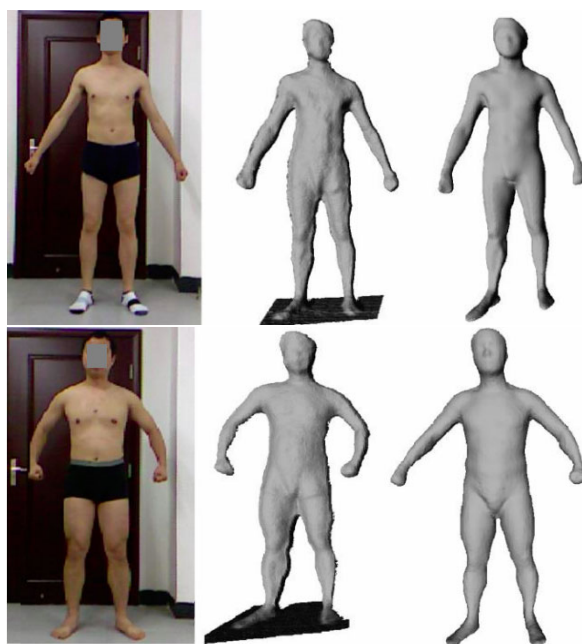


图 4 虚拟人重建结果

表 1 算法各部分时间统计 (秒)

相机标定	采集数据	点云注册	表面重建
1	3	6	10

<http://www.china-simulation.com>



为了评价重建的虚拟人模型精确度,我们选取了七个人体上的生理特征分量分别度量重建出的虚拟人模型与对应用户的实际身体数据之间的误差,表 2 展示了以厘米为单位的误差情况,可以看出本文采用的方法重建出的结果较好的重现了用户的身体各部分特征,具有很好的真实感。图 5 展示了本文重建出的结果与本课题组之前工作的对比,相对于[6](左侧)中身体模型侧面明显的空洞和缺失,本文重建出的人体模型(右侧)完整度和真实感有明显提高。

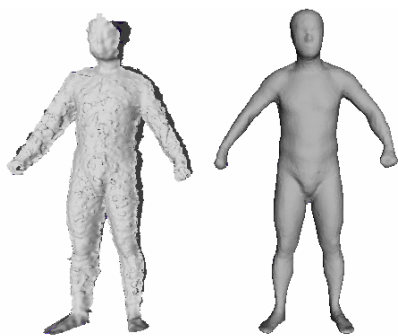


图 5 重建结果与[6]对比

表 2 用户身体不同部位平均误差(厘米)						
身高	背部高度	肩宽	臂长	腿长	腰围	臀围
1.0	1.3	1.8	2.8	1.6	3.5	2.5

本文提出的系统需要数据采集环境简单,仅需三台 Kinect 设备并且操作简单易行,设备造价便宜,每台 Kinect 设备价格仅需千余元,数据采集仅需 3 秒,数据处理在二十秒可以完成,比[7]速度明显加快。本文重建出的用户结果在体型等整体效果来看表现教好,明显优于[6]的效果,但是对于脸部,发型,手指等细节部分处理效果较差。

4 结论

虚拟人三维模型广泛应用在三维电影、虚拟现实、虚拟手术、游戏和训练模拟中,能够提供具有高度真实感的用户体验。本文提出了一种基于三个 Kinect 设备对任意用户生成三维人体模型的系统,系统采用三台 Kinect 从不同

角度扫描人体,通过几何标定将三组数据拼接在一起形成完整的虚拟人点云模型,最后通过表面重建生成人体模型。该系统已经在我们的实验环境下实验证明可行,实验结果表明本文提出的系统简单易用,造价便宜,需要空间小,采集速度快(3s),后期处理较快并且重建效果较好。

但是,本文的方法也存在一些不足,重建出的模型精度不高;诸如脸部表情,发型,手指等许多细节被丢弃;后期处理时间仍然较长,如果能做到实时性的建立虚拟人三维模型,将使得系统有更广泛的应用。

参考文献:

[1] Microsoft. Kinect [EB/OL]. (2012) [2013-7-1].  
<http://www.xbox.com/kinect>.

[2] Allen B, Curless B, Popović Z. The space of human body shapes: reconstruction and parameterization from range scans [J]. ACM Transactions on Graphics (TOG) (S0730-0301), 2003, 22(3): 587-594.

[3] Cyberware. The Cyberware Whole Body Color 3D Scanner. [EB/OL]. (2012) [2013-7-1].  
<http://www.cyberware.com/products/scanners/wbx.html>.

[4] 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. USA: ACM, 2011: 559-568.

[5] D N Stern. The present moment: In psychotherapy and everyday life [M], Norton Co, New York, USA, 2004. Chen Y, Cheng Z Q. Personalized avatar capture using two Kinects in a moment [C]// SIGGRAPH Asia 2012 Posters. USA: ACM, 2012: 3.

[6] Chen Y, Cheng Z Q. Personalized avatar capture using two Kinects in a moment[C]//SIGGRAPH Asia 2012 Posters. ACM, 2012: 3.

[7] Tong J, Zhou J, Liu L, et al. Scanning 3d full human bodies using kinects [J]. IEEE Transactions on Visualization and Computer Graphics (S1077-2626), 2012, 18(4): 643-650.

[8] Izadi Shahram. 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. USA: ACM, 2011.

[9] Z-Q Cheng, Y Chen, R R Martin, Y La. Supermatching: Feature Matching using Supersymmetric Geometric Constraints [J]. IEEE Transactions on Visualization and Computer Graphics (S1077-2626), 2013.