

Research on Kinect-Based Human Reconstruction and Dimensional Measurement

YANG Yifei¹, CHEN Qijun²

(1. Electronic and information Engineering, Tongji University, Shanghai, China;

2. Electronic and information Engineering, Tongji University, Shanghai, China)

Abstract: In this paper, we implement three-dimensional Human reconstruction by using a depth camera developed by Microsoft, called Kinect, in complex indoor environment. With the help of the corresponding software development kit (SDK), we hold the Kinect in hand and scan the whole human body through moving the camera slowly. In order to remove the interference of background, we utilize the player mask in raw Kinect depth information. Then, we can isolate the human pixels from the complex background pixels. Also, the camera position and posture are obtained by adjacent point cloud registration. Then we get the matrix to integrate different point cloud under different visual angle. On this basis, we can compute the key points of human body. Here, the engineering reference of Human Standard Size plays a vital role in reducing the point cloud retrieval scope and improving the retrieval precision. Meantime, we set a reasonable sampling frequency based on the accuracy of Kinect depth value, by which, greatly decreasing computing complexity and increasing precision to some extent.

Key words: Kinect; Three-dimensional reconstruction; dimensional measurement; point cloud registration

基于 Kinect 的人体三维建模与尺寸测量技术的研究

杨一飞¹, 陈启军²

(1. 同济大学电子与信息工程学院, 上海, 中国, 201804;

2. 同济大学电子与信息工程学院, 上海, 中国, 201804)

摘要: 本文借助微软开发的 Kinect 深度摄像头, 以及与之相对应的软件开发工具包 (SDK), 实现在复杂的室内环境中对人体进行三维建模并通过对点云的计算得到相应的尺寸值。采用手持 Kinect 的方法对人体进行扫描, 获得人体各个角度的深度信息。为了移除背景的干扰, 本文运用了 Kinect 原始深度数据中的人体掩模标志位, 来分割出人体像素点。并且, 通过相邻点云的配准得到摄像机位置信息和姿态, 从而可以将不同视角下点云合成在统一的立体空间内。在此基础上计算人体的关键尺寸点, 这里借助人体的标准尺寸的工程学参考来降低点云检索范围, 提高检索精度。同时, 基于 Kinect 的深度信息精度设置合理的采样频率, 大大减小了计算开销, 并且精度也有所提高。

关键词: Kinect; 三维建模; 尺寸测量; 点云配准

1 引言

随着计算机技术的发展,人们对于服装的合身性要求越来越高了。传统的手工测量人体尺寸的方法已经处在被淘汰的边缘,不仅是由于手工测量的繁琐与不方便性,同时也考虑到手工测量精度偏低,往往无法达到现阶段人们需求的精度级别。非接触式测量开始慢慢进入人们的视野,其应用前景非常广泛。建立人体三维模型通常有两种方法:一种是通过主流的 Maya、3DMax 等三维软件进行构建,这种方法是建立在标准化的人体尺寸基础上,局限性很大;另一种就是利用三维扫描技术,通过采集不同视角的人体点云三维坐标来进行建模。现阶段比较成熟的非接触式测量方法主要有基于激光扫描仪来实现的,也有基于红外线和结构光的,以现代光学为基础,融合电子学、计算机图像学、信息处理、计算机视觉等科学技术为一体,精度和测量速度都优于传统方法。但是在现实生活中它们都无法被广泛运用,其主要原因就是诸如此类的三维测量装置动辄需要几十万甚至上百万美元,这是一般的厂商所无法承担起的。目前投入商业应用的人体扫描系统有 Cyberware、TC2 image twin、Vitronics、Inspeck 等^[2]。

本文需要解决的问题是在复杂的室内环境中对人体进行实时的三维建模,并且在三维模型的基础上测量出关键尺寸数据。而我们无法承担现有商用扫描仪巨大的成本因素,要寻求一种既经济又能满足精度需求的三维建模方法。

2 Kinect 人体扫描实现三维重建

2.1 Kinect 简介

Kinect 是由微软开发的一款体感摄像头,最初只要 Xbox 版本,后来推出 For Windows 版本,为 PC 端开发体感应用带来了很大的便捷。Kinect 摄像头有三个镜头组成,中间的镜头是 RGB 彩色摄影机,用来采集彩色图像。左右两边镜头则分别为红外线发射器和红外线 CMOS 摄影机所构成的 3D 结构光深度感应器,用来采集深度数据(场景中物体到摄像头的距离)。彩色摄像头最大支持 1920*1080 分辨率成像,红外摄像头最大支持 640*480 成像。Kinect 还搭配了追焦技术,底座马达会随着对焦物体移动跟着转动。Kinect 也内建阵列式麦克风,由四个麦克风同时收音,比对后消除杂音,并通过其

采集声音进行语音识别和声源定位。

目前针对 Kinect 的驱动主要有三家,分别是微软官方的 Kinect SDK for Windows, 最早开始支援 Kinect 的第三方驱动 OpenNI (目前已经更新到 2.X), 以及 Linux 平台上的 Libfreenect。

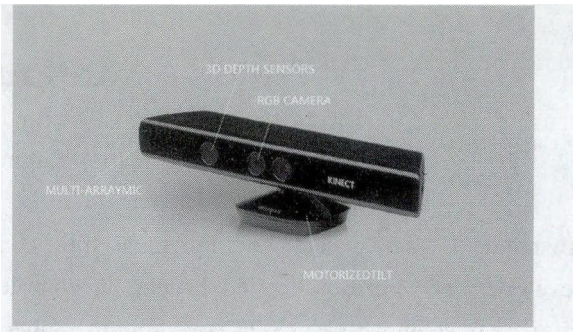


图 1 Kinect 深度传感器

2.2 摄像机位置配准

由于本文采用的是手持 Kinect 进行人体全方位扫描的方法来实现三维重建,当 Kinect 移动的时候,我们必须时刻记录下传感器的位置以及姿势信息。通过 Kinect 每一帧图像的深度信息,以及不同图像帧之间的关联,可以将不同角度采集的数据转换成摄像头坐标之间的转移矩阵。

2.2.1 深度信息转换

我们直接通过 SDK 从 Kinect 得到的深度数据是不能直接用于后期处理的,需要将原始深度帧数据转换为以米为单位的浮点数据,紧接着对该数据进行优化,通过获取摄像头的坐标信息,将这些浮点数据转化为和 Kinect 摄像头朝向一致的点云数据。

可以直接调用 SDK 中的 Reconstruction 类下面的成员函数 DepthToDepthFloatFrame()来完成相应工作。需要注意的是,我们在这个转换处理中就应该将除去人体以外的像素全部过滤掉,这样做可以大大减小后续的计算开销。考虑到从 Kinect 中得到的是 16bit 的深度数据流,前 13 位表示深度值,后三位用来存放 Player Index,简单的说就是 Kinect 中的深度数据信息里已经保存了人体识别的标志位,我们可以通过判断标志位是否为零来区分人体像素点与环境像素点。

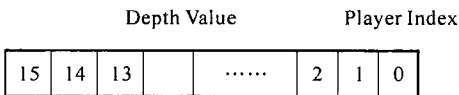


图 2 深度数据帧结构

在深度数据帧中,我们要获取后 3 位的 Player Index,主要代码如下所示:

```

Parallel.For(
    0, Height,
    y =>
    {
        int index = y * this.Width;
        for(int x = 0; x < this.Width; ++x, ++index)
        {
            if(DepthImagePixels[index].PlayerIndex == 0)
                DepthImagePixels[index].Depth = 0;
        }
    });

```

图3 从背景中分离人体像素

代码中的 `DepthImagePixels` 变量就是从深度帧数据中获取的深度信息数组, `Index` 是数组的索引, 通过遍历整个深度数据帧来找到人体像素点。这里考虑到性能问题, 采用了 `Parallel` 并行计算, 可以提高 CPU 计算速度。

2.2.2 摄像头姿态跟踪

采用迭代最近点算法 (ICP) 来实现点云之间的刚性配准是最常用的做法。先简单介绍迭代最近点算法: 根据某种几何特性对数据进行匹配, 并设这些匹配点为假想的对应点, 然后根据这种对应关系求解运动参数。再利用这些参数对数据进行变换。并利用同一几何特征, 确定新的对应关系, 重复上述过程^[7]。

假定三维空间中的两个点, $p_i = (x_i, y_i, z_i)$, $q_i = (x_i, y_i, z_i)$, 它们的欧式距离表示为:

$$d(p_i, q_i) = \|p_i, q_i\|$$

$$= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

三维点云匹配问题的目的是找到 P 和 Q 变化的矩阵 R 和 T , 对于 $q_i = Rp_i + T + N_i, i=1, 2, \dots, N$, 利用最小二乘法求解最优解使得:

$$E = \sum_{i=1}^N |(Rp_i + T) - q_i|^2$$

最小时的 R 和 T 。

这时我们就可以得到全局的摄像头的姿态信息, 包括摄像头的位置和朝向, 通过使用交互型的配准算法在摄像头移动时不断获取其姿态, 这样系统始终知道当前摄像头相对于起始摄像头的相对姿态。如果场景中的追踪被中断, 那么需要将摄像头的位置和上一次的摄像头位置对齐才能继续进行追踪。

2.3 深度信息融合

将从已知姿态摄像头产生的深度影像数据融合为代表摄像头视野范围内的人体立方体。这种对深度数据的融合是逐帧, 连续进行的, 同时通过平滑算法进行了去噪, 也处理了某些场景内的动态变

化, 比如人体一个细微的移动。随着 Kinect 的缓慢移动, 从不同市场角度观察人体表面, 人体表面的点云数据会持续优化, 最后得到完整的人体模型。

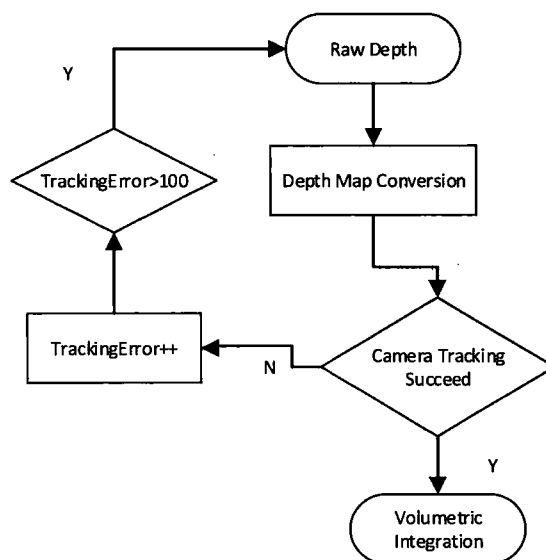


图4 整体流程图

3 人体点云处理与计算

在通过 Kinect 得到人体的点云数据之后, 如何准确地计算人体相关尺寸成为本文所要解决的另一个关键问题。

3.1 预处理

为了方便下面的计算, 我们需要调整点云的朝向, 让人体的正面始终朝向 Kinect 的方向。在这个方位, 我们可以有效利用人体的左右对称性来帮助我们计算相关尺寸^[3]。

- 定点云的最高和最低两端, 确定足部一端;
- 以一定的步长截取 5 个足部截面;
- 提取截面基线, 找到截面上距离基线的最近和最远点, 从最近点到最远点的矢量方向即为人体正面朝向。

3.2 计算人体特征点

本文计算的人体尺寸为制作上衣一般需要的尺寸, 如下表所示:

表1 需要采集的人体尺寸

尺寸名称	尺寸介绍
肩宽	被测者手臂自然下垂, 测量左右肩峰点之间的水

	平弧长。
胸围	被测者直立，正常呼吸，用软尺经肩胛骨、腋窝和乳头测量的最大水平围长。
腰围	被测者直立，正常呼吸，腹部放松，胯骨上端与肋骨下缘之间自然腰际线的水平围长。
臀围	被测者直立，在臀部最丰满处测量的臀部水平围长。
背腰长	用软尺测量自第七颈椎点沿脊柱曲线至腰际线的曲线长度。
臂长	被测者直立，手臂自然下垂，用软尺测量自肩峰点至尺骨茎突点所得的直线距离。
颈根围	用软尺经第七颈椎点、颈根外侧点及颈窝点测量的颈根部围长。

经过分析可以发现，所有的尺寸都是基于几个特征点而得到的。故先求得所需的特征点，再由特征点处的相应截面与点云相交点的曲线拟合，即可以求出所需的围长信息。

若每一个尺寸特征点的求取都需要扫描整个人体点云的话，效率会很低，本文采取根据《中国成年人人体尺寸》和《工作空间人体尺寸》研究分析得到的尺寸位置与身高的比例关系，大大降低了检索范围。

表2 人体各静态尺寸与身高H之间的线性关系

项目名称	线性关系
身高	1.00
颈部	0.84
肩部	0.82
腋窝	0.75
胸围	0.72
腰围	0.61
中臀围	0.53
胯部	0.47

3.2.1 腋窝点识别

参照表2的尺寸位置与身高之间的线性关系，确定人体腋窝所在的高度范围。考虑到我们在建模过程中的参数设定，体像素密度大约为4 mm，所以为了识别的准确度，我们采取上下浮动一个特定阈值来扩大检索范围，这个值 $\delta = 0.04$ 。

我们采取的步长为体像素密度的一半2mm，在检索范围内上下各遍历20次，找到左右两个凹陷点，即为我们所需要的腋窝点^[2]。

3.2.2 跨点识别

跨点的识别相对来说更简单，首先也是根据尺寸位置与身高的线性关系，确定跨点所在的高度范围，同样设定 $\delta = 0.04$ 的上下浮动范围。从上往下作Y轴的截面，步长为2 mm，取当前截面的所有点的X值的平均值， $\bar{x} = \frac{1}{x} \sum x_i$ ，若 \bar{x} 与当前截面所有点的欧式距离大于某一阈值，则称该点为所需的跨点。

3.3 人体部位划分

在得到腋窝点和跨点之后，可以将人体粗略地划分为三个部分：上肢，下肢，躯干。因为我们所要求的尺寸除了臂长外都处于躯干部分，所以进行人体部位划分之后可以很好地排除非躯干点对围度尺寸计算的影响^[6]。

3.4 关键尺寸提取

借助Kinect SDK可以得到人体20个关键骨骼点的三维坐标。本文只需要肩膀点的坐标和手腕点的坐标，在被测者直立、两臂自然下垂的时候，取前10帧图像的肩膀、手腕坐标进行平均，再计算两平均值之间的欧式距离得到臂长。

其余尺寸的值都是在点云数据中得到。这里简单介绍其中的三个尺寸数据的测量方法。

首先是胸围，我们只需从上到下搜索胸部所在范围，找到每个截面下Z的最小值，如果Z的最小值会随着搜索的进展越来越小，直到下一次截面所得Z值的最小值大于上一次检索值，则返回上一次的Z值所在截面为胸围所在截面^[5]。

其次是腰围，腰围的基准值要参考X值的大小随着搜索的变化，与胸围的计算方法一致。

最后是背腰长，在我们得到腰围的高度和颈根围的后颈点之后，通过搜索Y值在两者之间并且X值在一定范围内的点所拟合成的曲线即为背腰长。

4 总结与展望

4.1 取得的成果

- 本文在对人体进行三维建模的时候，并没有对环境条件进行约束，而是通过Kinect SDK得到的深度信息的Player Mask来分割出人体像素。使得本文提出的方法可以在

复杂的环境中进行。

- 有效利用微软对于人体骨骼点的识别算法,采用分段式来计算相应的尺寸。先在人体两臂自然下垂的站立姿态下计算得到臂长,再缓慢展开双臂到一定角度触发三维建模的指令,移动 Kinect 并开始整合不同视角下的深度信息。
- 对人体各部位进行划分,有效提高精度的同时减少了 CPU 的计算量。

4.2 需要改进的地方

本文所实现的方法还需要进一步的完善。需要提高摄像机姿态追踪的效率;本文目前只是实现了其中 7 个尺寸数据的自动提取,还需进一步实现其他数据的计算方法;手持相机的方法相对来说比较麻烦,在条件允许的情况下最好可以通过制作一个机械装置来实现手持移动这部分的功能。

参考文献

- [1] Dragomir Anguelov, PraveenSrinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers. SCAPE: Shape Completion and Animation of People[J]. ACM Transactions on Graphics, 2005.Vol.24(3): 408-416.
- [2] 宋诗超,禹素萍,许武军. 基于 Kinect 的三维人体扫描、重建及测量技术的研究[J]. 天津工业大学学报, 2012, Vol.31(5):34-41.
- [3] 林德静,孙晓东. 基于三维扫描的人体尺寸提取技术[J]. 北京服装学院学报, 2005, Vol.25(3): 36-41.
- [4] 田庆国,葛宝臻,杜朴,郁道银,吕且妮. 基于激光三维扫描的人体特征尺寸测量[J]. 光学精密工程, 2007, Vol.15(1): 84-88.
- [5] 李伟,张宝光. 应用于工业设计的人体尺寸数据的分析处理[J]. 中国科技论文在线.
- [6] 程达丽,李雪飞. 基于点云数据的人体特征尺寸提取[M]. 北京服装学院, 2011.
- [7] 周春艳,李勇,邹峥嵘. 三维点云 ICP 算法改进研究[J]. 计算机技术与发展, 2011, Vol.21(8).