

基于深度信息的快速身份识别方法^①

赵东杰, 汪同庆

(重庆大学 光电技术及系统教育部重点实验室, 重庆 400044)

摘要: 采用传统视频信号进行身份识别时, 易受遮挡、复杂背景等因素干扰的问题, 本文提出一种利用 Kinect 深度信息进行身份快速鉴别的方法. 首先利用微软 Kinect 设备获取人体俯视图(深度图像), 然后根据深度信息提取以下特征: (1)身高, (2)肩宽, (3)深度直方图, 根据人体生理结构的差异达到判别人身份的目的. 实验结果表明, 该方法计算简单, 具有较高的识别精度和较强的鲁棒性.

关键词: Kinect; 深度图像; 深度直方图; 身份识别

Rapid Person Identification Method Based on Depth Information

ZHAO Dong-Jie, WANG Tong-Qing

(Key Laboratory of Optoelectronic Technology & Systems Ministry Technology of Education, Chongqing University, Chongqing 400044, China)

Abstract: Traditional person identification methods have a low tolerance for occluded situation and complicated background. We focus on an image from an overhead camera. we utilize depth information for the person identification. We apply three features to the identification method: (1)the height, (2)shoulder width, (3)depth histogram. Experimental result shows our method has higher accuracy, strong robust-ness and good real-time with the existing methods of person identification.

Key words: Kinect; depth information; depth histogram; person identification

近年来, 人的身份识别已成为计算机视觉领域中一个重要的研究方向. 最直观的方法是采用人脸识别^[1], 步态识别也是一个研究热点, 国内外很多研究机构都做了深入的研究. 但是人脸识别和步态识别等传统生物识别技术对光照有较高的要求, 弱光或无光环境下不能正常工作, 对于多人情况下发生的遮挡现象也没有好的解决方法. 本文把微软 Kinect^{[2][3][4]}摄像机垂直地面放置, 获取目标的俯视深度信息, 解决了传统方法中的光照及遮挡问题, 提高了识别的稳定性和鲁棒性.

1 人体深度信息的获取

1.1 Kinect 简介

Kinect 是微软公司于 2010 年 6 月新推出的一款

Xbox360 体感设备, 该设备可以获取物体的 RGB 信息和深度信息, 并且深度信息与光照无关, 即便是在无光的情况下, 仍然能够得到清晰的深度图像. 其有效采集范围为 0.8m-4m. 深度图像的像素值表征的是目标点所在平面与相机所在平面之间的垂直距离.

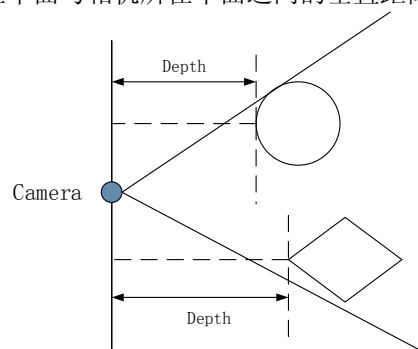


图 1 深度图像像素值的实际意义

^① 基金项目:中央高校基本科研基金(10611201312014)

收稿时间: 2014-03-03; 收到修改稿时间: 2014-04-11

深度图像的像素值只是一个相对数值，并不代表实际的距离，根据二阶高斯模型拟合^[5]得到的深度信息与实际距离关系曲线与实际测量得到的关系曲线非常接近，假设深度图像某一点的深度值为 d_{raw} ，则可有下式求出该点的实际深度值：

$$d = K \tan(d_{raw}/1284.5 + 1.1863) - O \quad (1)$$

其中， $K = 0.1236$ ， $O = 0.037$

根据实际距离深度值，可求得像素位置为 (i,j,d) 的点的实际位置 (x,y,z) ，其转换公式为：

$$\begin{cases} x = (i - \frac{w}{2}) * (d - 10) * s * \frac{w}{h} \\ y = (j - \frac{h}{2}) * (d - 10) * s \\ z = d \end{cases} \quad (2)$$

其中， $s=0.0021$ 。

1.2 获取人体深度信息

将 Kinect 垂直于地面固定在天花板上，以获取目标的俯视图，Kinect 的安装示意图如图 2 所示：

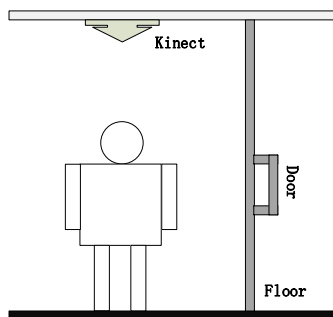


图2(a) 正面图

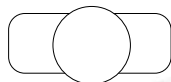


图2(b) 俯视图

图 2 Kinect 的安装示意图

Kinect 能够同时获得目标的彩色信息与深度信息，本文中深度图像分辨率采用 240×320 像素。目标 RGB 图及相对应的深度图如图 3 所示：



图 3 深度图像与彩色图像

2 特征提取

2.1 特征分析

由 Kinect 得到的俯视深度图像能够反映人的身体结构的差异，我们提取以下特征：(1)身高，人的身高是我们判别人身份的重要依据；(2)肩宽，人肩宽是易计算又稳定的特征；(3)深度直方图，深度图像中，人体头部、肩膀及地面三块区域的深度值具有明显的跳变现象，深度直方图能够反映深度图像中像素点的聚集规律。

2.2 特征描述

2.2.1 身高计算

身高指的是从头顶点到地面的垂直距离，可以利用头部到 Kinect 相机的平均距离来间接表征身高大小，也就是人头部区域的平均像素值。

首先确定头部区域^{[6][7]}，Matas 等^[8]提出的最大稳定极值区域(MSER)算法是公认的较好的区域检测算法，这里定义的最大稳定极值区域是区域面积变化率随着阈值的变化保持稳定的区域：

1) 根据 Kinect 离人头部的距离设置有效的阈值范围，这里设置为 $(30,110)$ 。

2) 逐一比较图像像素值与给定阈值，小于或等于给定阈值的点记为白色，大于给定阈值的点记为黑色，并求出

$$K = \frac{A_n - A_{n-1}}{A_n} \times 100\% \quad (3)$$

其中， A_n, A_{n-1} 为取相应阈值时白色区域的面积， K 为面积变化率。

3) K 若在取第 n 个阈值时变化较大，就把取第 $n-1$ 个阈值时的白色区域提取出来，直到提取出所有符合条件的白色区域，得到的白色区域就是最大稳定极值区域，如图 4 所示：



图 4 最大稳定极值区域

4) 在得到的最大稳定极值区域中，首先根据面积大小将面积异常白色区域置为黑色，人的头部近似圆形或椭圆形，故可以根据圆形度^[9]剔除非头部的白

色区域, 并将其设置为黑色, 如图 5 所示. 圆形度公式为:

$$C = \frac{4\pi \times A}{P^2} \quad (4)$$

其中, P 为区域周长, A 为区域面积.

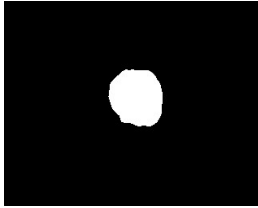


图 5 头部区域

5) 得到人头部区域之后, 就可求出头部区域像素的平均值, 如下式:

$$M = \frac{1}{N} \sum_{i,j} f(x_i, y_j) \quad (5)$$

其中, N 为头部点的个数, (x_i, y_j) 为白色区域点的坐标, f(x_i, y_j) 为原图像中对应点的像素值, M 为头部区域像素的平均值, 可以用 M 值来表征人的身高.

2.2.2 肩宽计算

Kinect 的水平视场角为 57°, 知道了深度值, 根据三角几何知识, 就可以计算出物体的实际宽度, 计算肩宽的示意图如图 6 所示:

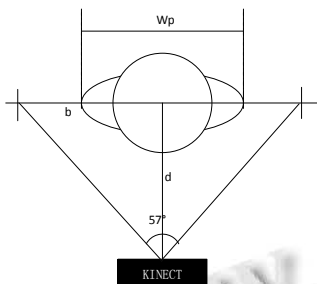


图 6 肩宽计算示意图

计算过程如下:

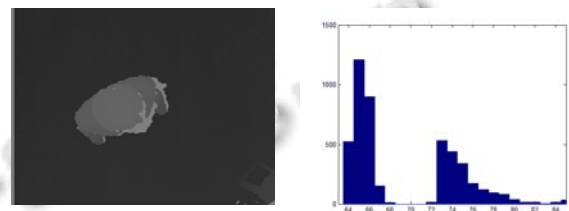
$$\begin{aligned} b &= d \tan(28.5^\circ) \\ \frac{W_p}{320} &= \frac{W_r}{2b} \\ W_r &= \frac{2bW_p}{320} \end{aligned} \quad (6)$$

其中, W_p 为人体在深度图像中的像素宽度, W_r 为肩膀的实际宽度, d 为深度距离. 人体的肩膀宽度是稳定的, 即便是人在 Kinect 视场中旋转, 其肩膀在深度图像中的像素宽度也不会发生变化, 所以公式(6)可以用

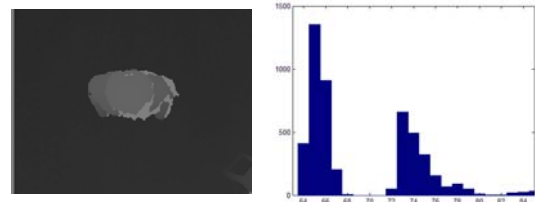
来计算 Kinect 视场中不同朝向的人的肩宽.

2.2.3 深度直方图

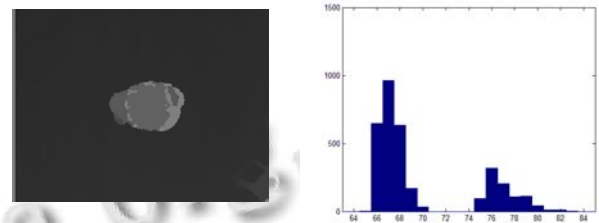
深度直方图中, 由左向右, 表示目标距离从近到远的变化, 峰值代表目标物体. 由图 7(a)、(b)可以看出, 同一人在 Kinect 视场中不同位置拍摄的深度图像, 其深度直方图中柱状图分布、峰值变化等规律是一致的. 图 7(c)所示为另一人的深度直方图, 可以看出, 其与(a)、(b)所示直方图有明显的差异.



(a) 深度图像及其直方图



(b) 深度图像及其直方图



(c) 深度图像及其直方图

图 7 深度直方图

综上所述, 根据人体的俯视深度图像提取如下特征^[10]:1)身高 H; 2)肩宽 W; 3)深度直方图, 每个样本的深度直方图特征 X_i=[x₀, x₁...x_i] j=0,1,2, ...255. 不同的人的深度图像中所占的像素点数目不同, 故而需要做归一化处理, 处理方法如下:

$$x_i' = \frac{x_i}{\sum x_j} \quad (7)$$

归一化处理之后得到新的特征向量 X_i'=[x₀', x₁', ...x_j'], j=0,1,2, ...255.

在这里, 用于身份识别的特征向量为 D=[X', H, W], 维数为 258, 由于特征向量 D 中参数的量纲不同, 为了消除量纲的影响, 需要对所得数据进行标准化处

理, 选取 Z-Score 标准化方法, 即

$$D' = \frac{D_i - \mu}{\sigma}, i = 0, 1, \dots, 257 \quad (8)$$

其中, μ 为样本中相应项的均值, σ 为方差, 此时得到的特征向量 D' 中参数无量纲, 本实验采用支持向量机(SVM)进行特征训练与识别.

2.3 实验结果与分析

实验平台为 VS2010, 结合 OpenCV 和 LIBSVM 完成数据采集、处理及样本训练识别工作, 深度图像分辨率为 320*240, 本实验以 12 位男同学为研究对象, 最高者 180cm, 最低者 162cm, 其中有 5 人体型差别不大. 样本的身高、肩宽及深度直方图特征分别用 H、W 和 Hist 代替. 下表所示为测试结果:

表 1 分类识别结果

	识别率	
	中心位置	位置偏移
H	33.3%	33.3%
W	16.7%	16.7%
Hist	75%	58.3%
H+W	41.7%	33.3%
H+Hist	83.3%	58.3%
W+Hist	83.3%	58.3%
H+W+Hist	91.7%	58.3%

表中的中心位置指的是人居于 Kinect 视场中心, 图像中仅有头部和肩部, 位置偏移指的是人体偏离 Kinect 视场中心的位置. 从表中分类结果可以看出, 身高特征最为稳定, 这是因为 Kinect 采集的深度信息是目标点所在位置到摄像机所在平面的垂直距离, 所以人体位置发生变化时, 获取的身高信息没有变化. 而深度直方图特征识别率最高, 因为深度直方图蕴含的信息最丰富, 能直观反映像素点出现的频率及聚集规律, 两大像素点聚集区域, 分别代表头部和肩部, 聚集像素点的数目则表示人体头部和肩部的区域面积. 但是当人的位置偏离中心视场较远或者站姿歪曲时, 将会引入大量的噪声, 导致识别精度严重下降. 从实验中可以看出, 体型相似的人, 可以通过特征组合的方式做进一步的鉴别, 而且识别效果会比单一特征更好.

3 结语

本文提出一种基于 Kinect 深度信息的快速身份识

别方法, 利用最大稳定极值区域算法求出身高, 结合 Kinect 视场几何关系求出肩膀宽度, 算法直观、易懂, 计算简单, 能够满足实时性的要求. 实验表明, 该方法识别精度高、鲁棒性强, 完全不受背景及光线等因素的影响.

参考文献

- 1 王跃明, 潘纲, 吴朝晖. 三维人脸识别研究综述. 计算机辅助设计与图形学学报, 2008, 20(7): 819-829.
- 2 Introducing Kinect for Xbox360. <http://www.xbox.com/en-US/kinect>. [2012-08-10]
- 3 吴国斌, 李斌, 阎骥洲. KINECT 人机交互开发实践. 北京: 人民邮电出版社, 2013.
- 4 余涛. Kinect 应用开发实践—用最自然的方式与机器对话. 北京: 机械工业出版社, 2013.
- 5 Stowers J, Hayes M. Altitude control of a quadrotor helicopter using depth map from Microsoft Kinect sensor. 2011 IEEE International Conference on Mechatronics(ICM). IEEE, 2011. 358-362.
- 6 王健, 王孝通, 徐晓刚, 等. 基于梯度的随机 Hough 快速圆检测方法. 计算机应用研究, 2006, 23(8): 64-165.
- 7 Luciano S, Arras KO. People detection in RGB-D data. Intelligent Robots and Systems, IROS 2011. 3838-3843.
- 8 Matas J, Chum O, Urban M. Robust wide baseline stereo from maximally stable extremal regions. Image and Vision Computing, 2010, 22(10): 761-767.
- 9 冯维, 胡涛, 朱欣焰. 基于 Kinect 的深度数据融合方法. 计算机应用研究, 2014, 31(1): 285-288.