

文章编号: 1004-4353(2016)03-0235-06

基于 Kinect 深度和骨架信息的指尖检测方法

毛雁明, 阮群生, 章立亮

(宁德师范学院 计算机系, 福建 宁德 352100)

摘要: 为克服指尖检测方法易受不同光照、复杂背景和手腕信息的影响, 提出了一种基于深度和骨架信息的指尖检测方法. 首先对 Kinect 获取的深度图像进行中值滤波和形态学闭操作处理, 以消除噪声和填充空洞; 接着通过骨架跟踪得到的右手关节点锁定用户并进行手势分割; 然后在计算手心和最高指尖点位置的基础上, 利用 Freeman 链码提取手势左右轮廓; 最后根据指尖点之间的轮廓曲线特征提取其他指尖点. 实验结果表明, 该方法具有良好的指尖检测效果, 且对光照、背景和手腕信息鲁棒.

关键词: 骨架信息; 深度信息; 手势分割; 指尖检测

中图分类号: TP391.41

文献标识码: A

A fingertip detection method based on Kinect depth and skeleton data

MAO Yanming, RUAN Qunsheng, ZHANG Liliang

(Department of Computer, Ningde Normal University, Ningde 352100, China)

Abstract: To overcome the fingertip detection method which is susceptible to the different illumination, complex background and wrist data, a fingertip detection method based on skeleton and depth data was proposed. First, in order to eliminate noises and fill holes, it used median filter and morphological closing operation to process the depth image captured by Kinect sensor; subsequently, it used right hand joint point obtained by the skeleton tracking for lock user and gesture segmentation; then, on the basis of calculating the positions of palm and the highest fingertip, it used Freeman Chain Code (FCC) to extract the left and right contours of the gesture; finally, the other fingertips were extracted according to the contour curve features between fingertips. Experimental results show that the proposed method has good finger detection effect and also robust to the illumination, background and wrist information.

Keywords: skeleton data; depth data; gesture segmentation; fingertip detection

0 引言

由于手势相对于键盘和鼠标具有自然、直观的特点, 因此, 近年来手势识别成为国内外众多科研工作者在人机交互领域的重要研究课题之一. 手势识别的研究与开发可以应用到虚拟现实、视频监控、远程会议、手语识别以及机器人控制等诸

多领域^[1]. 手指特征点的运动是手势识别和理解的基础, 具有动作表示方式简单、含义丰富等特点, 因此, 对指尖检测方法的研究引起了研究者广泛关注, 并已取得了一定的研究成果^[2-6]. 传统的手势分割一般采用肤色模型来实现^[7-9], 该方法易受复杂背景、光照条件以及类肤色的干扰, 而且人种肤色的差异也限制了其应用范围. 目前的指尖

收稿日期: 2016-08-29

作者简介: 毛雁明(1982—), 男, 讲师, 研究方向为计算机视觉、模式识别.

基金项目: 福建省自然科学基金资助项目(2015J01660); 福建省教育厅 A 类科技项目(JA15543); 宁德师范学院青年科技项目(2015Q04)

检测算法主要采用手部轮廓特征,例如:文献[10]在肤色区域的轮廓上根据曲率进行指尖检测,并通过很多限制条件来判定指尖点,该方法实时性较差;文献[11]利用肤色和运动信息提取手区域,采用自适应 K 曲率检测指尖,但该方法由于增加了过多条件,如指宽、指长、指向和手掌宽度信息等,影响了指尖检测的效果。

为解决上述手势分割和指尖检测存在的问题,本文首先利用 Kinect 传感器获取深度图像和人体骨架信息,通过骨架跟踪获取的右手关节位置信息锁定用户并分割手势;然后在计算手心和最高指尖点位置的基础上,通过 Freeman 链码提取手势的左右轮廓,最后利用指尖点之间的轮廓特征提取了其他指尖点。

1 深度图像和骨架信息获取

Kinect 传感器通过红外激光源发出有编码的红外线激光标记场景,而后接收器接收返回来的红外线激光信息交给 PS1080 芯片进行处理,算出深度图. 人体骨架信息的获取是在 Kinect 传感器获取目标深度图后,对每个像素进行概率推测,并分类它们的特征变量,然后在一个随机决策库中进行搜索,匹配出它是人体的哪一部位^[12]. 本文通过 OpenNI 提供的深度生成器获取深度图像,通过骨架功能 SkeletonCapability 提供的成员函数 GetSkeletonJointPosition() 获取骨架节点绝对坐标位置,再通过深度生成器中的 Convert-RealWorldToProjective() 函数将其转换为屏幕坐标. 图 1 为深度图像和人体骨架叠加图.



图 1 深度图像和人体骨架叠加图

2 手势分割

2.1 深度图像预处理

由于 Kinect 传感器获取的深度图像不稳定,

存在噪声和空洞,直接处理会对后续的手心位置计算、轮廓提取以及指尖检测效果造成影响,因此需要先对深度图像进行预处理. 通过分析均值滤波、低通滤波和中值滤波等滤波器的特性,并针对深度图像零值噪声点较多的情况,本文采用中值滤波对深度图像进行预处理以去除噪声.

中值滤波是一种非线性平滑技术,通过滑动窗口扫描整幅图像,将窗口内的像素值从大到小排列,将其中位数作为窗口中心点的像素值;但是,中值滤波无法解决深度图像空洞的问题. 形态学中的膨胀操作是将与物体接触的所有背景点合并到该物体中,使边界向外部扩张的过程,可以用来填充物体中的空洞;腐蚀操作是一种消除边界点,使边界向内部收缩的过程,可以用来消除小且无意义的物体. 因此,为了填充空洞而又不使边界向外扩张,采用先膨胀后腐蚀的形态学闭操作对深度图像进行处理.

2.2 手部提取

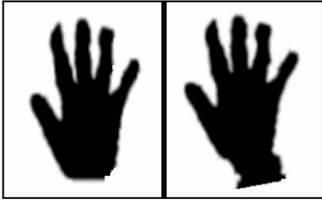
针对基于肤色的手部分割易受不同光照、复杂背景和类肤色的影响,以及基于 OpenNI 提供的手部生成器获取手心位置的手部分割^[1,13]易受其他用户进入视野的影响,造成手部的分割错位,本文采用了基于骨架信息的手部分割方法. 由于深度图像中的每个像素点的低 3 位代表了用户的 ID 信息,且 OpenNI 提供的骨架功能能够实时跟踪人体骨架关节位置,因此,可通过搜索特定的 ID 信息锁定用户,并实时跟踪该用户,以消除其他用户进入该区域造成的影响;然后根据骨架跟踪获取的右手位置(设其在深度图像中的像素值为 rH)设定阈值实现对手部的分割. 图 2 为通过式(1)进行手部分割的效果图.

$$Hands(x, y) = \begin{cases} 0, & \text{if } rH - 5 \leq \text{Dep}(x, y) \leq rH + 5; \\ 255, & \text{otherwise.} \end{cases} \quad (1)$$

式(1)中: $Hands(x, y)$ 表示分割后深度图像在像素点 (x, y) 的灰度值, rH 表示右手关节的深度值, $\text{Dep}(x, y)$ 表示深度图像中像素点 (x, y) 的灰度值.

从图 2 可以看出,指尖的方向与 Kinect 摄像头方向有一定倾斜时,分割的手部会包含手腕信息,但这并不影响后续的指尖检测,这是因为在后

续的指尖检测之前会通过提取手势左右轮廓消除手腕信息.



(a) 不包含手腕 (b) 包含手腕

图 2 手部分割图

3 指尖检测

3.1 手心位置计算

本文利用几何重心来获取手心位置,几何重心位置的计算需要利用图像的几何特征矩.图像的几何特征矩是通过对手部所有点进行积分运算而得到的一个粗略特征,该方法能够有效地描述对象的形态特征.其中,图像的零阶矩代表了图像的质量,图像的一阶矩用来确定图像重心,图像的 (p, q) 矩计算公式为: $m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$. 对于通过 2.2 得到的手部二值图像 $Hands(x, y)$,可利用式(2)—(5) 计算手部重心:

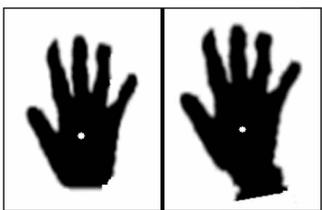
$$m_{00} = \sum_x \sum_y Hands(x, y), \tag{2}$$

$$m_{10} = \sum_x \sum_y x Hands(x, y), \tag{3}$$

$$m_{01} = \sum_x \sum_y y Hands(x, y), \tag{4}$$

$$\bar{x} = m_{10} / m_{00}, \bar{y} = m_{01} / m_{00}. \tag{5}$$

其中, m_{00} 表示手部图像零阶矩, m_{10} 和 m_{01} 表示手部图像一阶矩, (\bar{x}, \bar{y}) 表示手部重心 Palm. 由上式计算得到的手心(白色小圆圈)如图 3 所示.



(a) 不包含手腕 (b) 包含手腕

图 3 手心位置

3.2 最高指尖点检测

按 2.2 方法进行手势分割可以得到手部的基

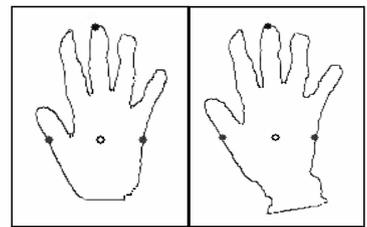
本信息. 由于人们在做手势时,基本都是指尖朝上,且手势的最高点为指尖点,因此可通过从左到右、从上到下的顺序对分割得到的手势图像进行搜索.一旦找到某个像素点的灰度值为零时,则可确定该像素点为最高指尖点,记为 Top. 图 4 为通过上述算法分别对手势 1—5 搜索得到的最高指尖点示意图.



图 4 手势最高指尖点(灰色圆圈代表最高指尖点)

3.3 轮廓提取

从图 2 中的手势分割图可以看出,由于人们做手势时,手部不会总能与 Kinect 摄像头接近垂直,因此这样得到的手势就可能包含手腕.为了消除手腕对轮廓提取效率和后续指尖检测算法的影响,本文根据前面获取的手心位置 Palm,在保持手心纵坐标 \bar{y} 不变的条件下,分别从左到右和从右到左查找第一次与手势相交的像素点,分别设为 lEndContourPt 和 rEndContourPt,其像素点分别记为 (x_l, y_l) 和 (x_r, y_r) ,如图 5 所示.轮廓提取时,首先从最高指尖点 Top 开始,然后分别向左和向右提取轮廓点,直到提取到手势轮廓点 lEndContourPt 和 rEndContourPt 结束.按此方法提取的轮廓不再包含手腕信息.



(a) 不包含手腕 (b) 包含手腕

图 5 轮廓终点提取

为了提取轮廓,首先需要利用边缘检测算子检测出手势的轮廓线.本文采用 Canny 算子进行

边缘检测,这主要是因为 Canny 算子得到的边缘是单像素的.轮廓提取算法采用 Freeman 链码,以最高指尖点 (x_t, y_t) 作为参考点,与其相邻的像素分别在 8 个不同的位置上,并赋予它们方向值 0—7(称为 0—7 位链码方向值),如图 6 所示.对于往左方向的轮廓提取,开始时只向 4、5、6 这 3 个方向提取轮廓点,其他轮廓点都用 8 链码进行提取,直到提取到像素点 lEndContourPt 结束;对于往右方向的轮廓提取,开始时只向 0、6、7 这 3 个方向提取轮廓点,之后的轮廓点都用 8 链码进行提取,直到提取到像素点 rEndContourPt 结束.图 7 为手势轮廓提取效果图.

3	2	1
4	p	0
5	6	7

图 6 p 点的 8 链码示意图

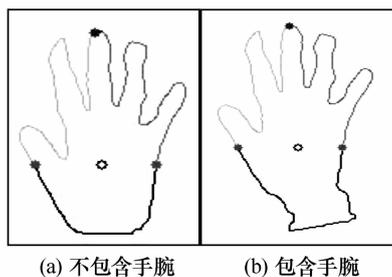


图 7 轮廓提取图(最高点到左终点的轮廓为左轮廓,最高点到右终点的轮廓为右轮廓)

3.4 指尖提取

假设提取的手势轮廓的左轮廓为 LeftContour,包含 m 个轮廓点,则其存储了从最高点 $\text{Top}(x_t, y_t)$ 到像素点 (x_l, y_l) 的 m 个轮廓点,并按轮廓提取时的顺序存放,记为 (lx_i, ly_i) , $i = 0, \dots, m-1$; 同样,假设所提取的手势轮廓的右轮廓为 RightContour,包含 n 个轮廓点,则其存储了从最高点 $\text{Top}(x_t, y_t)$ 到像素点 (x_r, y_r) 的 n 个轮廓点,并按轮廓提取时的顺序存放,记为 (rx_i, ry_i) , $i = 0, \dots, n-1$.

从提取的手势轮廓特征可以发现,2 个指尖点之间的所有轮廓点的纵坐标大部分都满足先递

减后递增的趋势,对于少部分不满足要求的轮廓点,本文按以下方法解决.以左轮廓 LeftContour 指尖点 Top 左边存在一个指尖点为例:采用选取步长和概率的方法,首先计算从指尖点 Top 开始的前 8 个轮廓点纵坐标 ly_i , $i = 0, \dots, 7$,若满足条件 $ly_{i+1} - ly_i \leq 0$ 的 i 有 5 个,则认为该段轮廓像素点呈递减状态;否则,呈递增状态.按照该方法继续查找,若连续出现 3 次递增的情况,即可判定该 3 次递增中的第 1 次的 8 个轮廓点存在 1 个指谷点;继续上述步骤,直到连续出现 3 次递减的情况,即可判定该 3 次递减中的第 1 次的 8 个轮廓点存在一个指尖点,此时该 8 个轮廓点中纵坐标最大的点即为指尖点.按上述方法继续查找下一个指尖点,若满足上述条件则为指尖点,否则继续查找,直到查找到像素点 lEndContourPt 结束.

3.5 指尖检测算法

基于 Kinect 深度和骨架信息的指尖检测算法具体描述如下:

步骤 1 利用 Kinect 传感器采集深度图像和骨架关节信息.

步骤 2 采用中值滤波和形态学闭运算对深度图像进行预处理.

步骤 3 根据骨架追踪得到的右手关节信息锁定用户并分割手势.

步骤 4 计算手心位置、轮廓左右终点位置以及搜索最高指尖点位置.

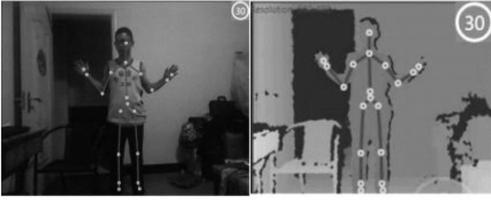
步骤 5 采用 Freeman 链码提取手势的左右轮廓.

步骤 6 根据指尖点之间的轮廓曲线提取其他指尖点.

4 实验结果及分析

为验证本文算法的有效性和指尖检测的准确性,以 VS2010 为编译环境,利用 C++、OpenNI v1.7 和 OpenCV v2.3.1 进行算法设计.实验时手指向上,人距离 Kinect 摄像头约 1.5 m.

图 8 中(a)、(b)、(c)分别为人体在强光、黑暗和复杂背景干扰下的骨骼跟踪效果图.从图 8 可以看出,该骨骼跟踪对不同光照条件和复杂背景具有较好的鲁棒性.



(a) 强光下骨骼跟踪效果图



(b) 黑暗下骨骼跟踪效果图



(c) 复杂背景下骨骼跟踪效果图

图 8 骨骼跟踪效果图

图 9 为手势轮廓提取效果图. 从图 9 可以看出, 本文的轮廓提取方法能够较好地将手势轮廓提取出来, 剔除了手腕对后续指尖检测效果的影响.

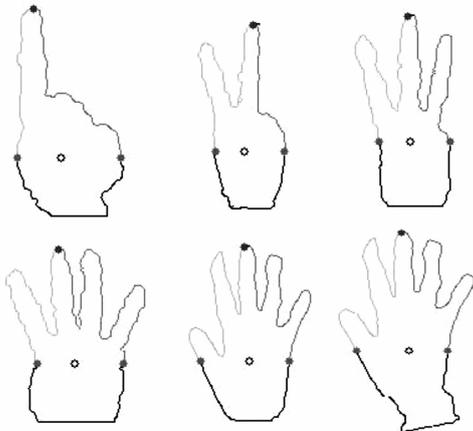


图 9 手势轮廓提取效果图(最高点ToLeft终点的轮廓为左轮廓,最高点toRight终点的轮廓为右轮廓)

图 10 为指尖检测实验结果图. 从图 10 可以看出, 本文算法能够实现对手势 1—5 指尖的正确识别, 即使分割的手势包含手腕信息, 也能够正确识别出指尖位置.

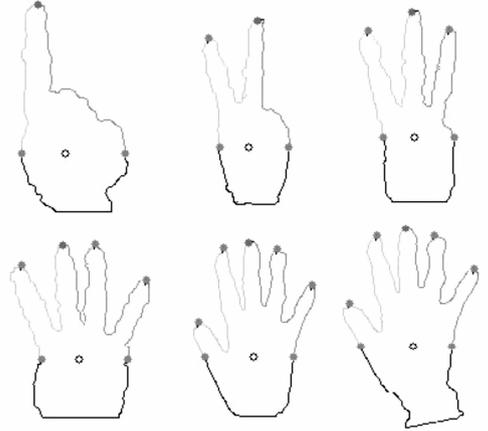


图 10 手势 1—5 指尖检测结果图

表 1—表 3 为 20 个人分别采用文献[11]、文献[14]方法和本文算法按手势 1—5 各做 10 次的指尖检测结果. 从表 1—表 3 可以看出, 本文算法的指尖平均检测正确率(99.2%)均高于文献[11](正确率为 91.9%)和文献[14](正确率为 97.5%)的方法. 本文方法中手势 5 出现检测正确率相对低的原因是: 分割得到的手掌相比正常手掌较短时, 计算出来的手心偏上, 使得左轮廓终点临近大拇指指尖点, 从而影响了大拇指指尖的提取.

表 1 文献[11]方法的指尖检测正确率

伸出手指数	1	2	3	4	5
正确的帧数	190	185	188	180	176
检测正确率/%	95	92.5	94	90	88

表 2 文献[14]方法的指尖检测正确率

伸出手指数	1	2	3	4	5
正确的帧数	196	193	193	195	198
检测正确率/%	98	96.5	96.5	97.5	99

表 3 本文算法的指尖检测正确率

伸出手指数	1	2	3	4	5
正确的帧数	200	200	200	200	192
检测正确率/%	100	100	100	100	96

5 结束语

本文研究了基于 Kinect 深度和骨架信息的指尖检测方法. 在手势分割阶段, 通过 OpenNI 中骨架功能 SkeletonCapability 提供的成员函数 GetSkeletonJointPosition() 获取骨架右手关节坐标位置, 通过 ID 信息锁定用户并分割手势; 指

尖检测阶段,在计算手心和最高指尖点位置的基础上,通过 Freeman 链码提取手势的左右轮廓,然后利用指尖点之间的轮廓特征提取其他指尖点.该方法使用设备简单,计算量少,检测效果较好,且对复杂背景和光照具有很好的鲁棒性,因此可应用在二维空间、单指尖或多指尖的交互系统中.

本文方法存在的问题有:1)要求做手势时手指必须向上;2)两指靠扰时,只能检测到一个指尖;3)只适用于手势 1—5 的指尖检测.因此,如何解决上述存在问题是今后的研究重点,以更好地完善该方法.

参考文献:

- [1] 毛雁明,章立亮.基于 Kinect 深度信息的手势分割与识别[J].系统仿真学报,2015,27(4):830-835.
- [2] Raheja J L, Chaudhary A, Singal K. Tracking of fingertips and centers of palm using Kinect[C]//Proc of 2011 the 3rd International Conference on Computational Intelligence, Modeling & Simulation, 2011:248-252.
- [3] Li Yi. Hand gesture recognition using Kinect[C]//Proc of 2012 IEEE the 3rd International Conference on Software Engineering and Service Science, 2012: 196-199.
- [4] Oikonomidis I, Kyriazis N, Argyros A A. Efficient model based 3D tracking of hand articulations using Kinect[C]//Proc of the 33nd British Machine Vision Conference, 2011:1-11.
- [5] 梅萍华,李斌,朱中的,等.基于径向对称变换的实时指尖检测算法[J].中国科学技术大学学报,2011,41(2):101-107.
- [6] Fahh C S, Sun H. Development of a fingertip glove equipped with magnetic tracking sensors[J]. Sensors, 2010,10(2):1119-1140.
- [7] 谭同德,郭志敏.基于双目视觉的人手定位与手势识别系统研究[J].计算机工程与设计,2012,33(1): 259-264.
- [8] 张秋余,王道东,张墨逸,等.基于特征包支持向量机的手势识别[J].计算机应用,2012,32(12):3392-3396.
- [9] 薛俊韬,纵蕴瑞,杨正瓴.基于改进的 YCbCr 空间及多特征融合的手势识别[J].计算机应用与软件,2016,33(1):151-155.
- [10] 李博男,林凡.基于曲率的指尖检测方法[J].南京航空航天大学学报,2012,44(4):587-591.
- [11] 徐义春,沈沅瑜,陈炼,等.鲁棒性指尖检测框架及其算法改进与分析[J].计算机工程与应用,2013,49(15):195-200.
- [12] 毛雁明,章立亮.基于 Kinect 骨架追踪技术的 PPT 全自动控制方法研究[J].海南大学学报(自然科学版),2015,33(3):215-220.
- [13] 李哲,彭四伟.基于深度信息的指尖跟踪与轨迹识别[J].计算机应用与软件,2016,33(4):155-172.
- [14] 徐鹏飞,张红英.基于 Kinect 深度图像信息的手势分割和指尖检测算法[J].西南科技大学学报,2014,29(1):49-54.