

文章编号:1671-6833(2018)05-0033-06

一种基于 RGB-D 的人体关节点定位方法

张青建, 韩建平

(杭州电子科技大学 计算机学院, 浙江 杭州 310016)

摘 要:提出了一种基于 RGB-D 的人体关节点定位方法:对 kinect 获取的深度图像滤波膨胀细化处理得到人体骨架线,根据角点检测算法得到人体骨架线的端点及折点,然后利用运动目标的跟踪算法对彩色图像中的人脸部跟踪定位,最后结合人体各刚架比例关系,定位出人体的 15 个关节点.试验结果表明,与直接根据人体各刚件比例关系定位关节点的方法相比,此方法的关节点平均定位准确率提高了 7%,其中头节点定位准确率提高了 42%.

关键词:RGB-D; kinect; 图像细化; 角点检测; 脸部跟踪; 关节点定位

中图分类号:TP391 **文献标志码:**A **doi:**10.13705/j.issn.1671-6833.2018.05.007

0 引言

Kinect 不仅价格低廉,而且能够获取彩色和深度信息,其深度信息具有抗光照扰动的优势,因此可用于人体关节点定位.文献[1]利用目标识别标记出人体模型,通过分类器定位出人体关节点,此方法是目前比较先进的关节点定位方法,但该方法需要大量的训练标记样本.文献[2]利用深度图像获取人体细化线,并根据人体刚架比例关系定位出人体关节点,此方法容易实现,但定位准确率有待提高.

笔者在文献[2]的基础上,加入了角点检测及分类、脸部跟踪等内容,提高了关节点定位的准确性和鲁棒性.文章首先对深度图像预处理以获取人体骨架线,从而得到骨架线角点并对角点筛选分类,然后在彩色图像中对人的脸部跟踪以定位头节点,最后结合人体各刚件比例关系,定位出人体关节点,方法如图 1 所示.

1 获取人体骨架线角点及头节点

1.1 深度及彩色图像预处理

用 kinect 摄取彩色图像和深度图像,由于此二者坐标不同,所以笔者以深度图像的坐标为基准,实现彩色图像坐标向深度图像坐标对齐^[3],以保证头节点(从彩色图像中获取)和其他关节点(从深度图像中获取)坐标一致.设置深度图像

的有效景深为 0.8 ~ 3.8 m,超过此范围的景深,在彩色图像中予以擦除,初步排除不必要的像素干扰.

Kinect 直接摄取的深度图像中人体轮廓边缘不光滑,人体区域内可能有黑洞,若不对深度图像预处理,细化后的人体骨架线有毛刺,甚至导致骨架线变形,因此需对深度图像进行滤波、膨胀腐蚀等预处理.首先对深度图像用中值滤波祛除噪声^[4];然后对深度图形膨胀腐蚀处理以消除人体区域内的黑洞^[5];最后对深度图像二值化^[6]处理以减少数据量,取阈值 128,大于阈值的像素置为 1,否则置为 0.

1.2 获取人体骨架线及角点

对预处理后的深度图像,利用 Zhang-Suen 细化算法^[7]获取人体骨架线,并对骨架线过滤,使得骨架线上每两个像素之间有一个空白像素,然后利用 Shi-Tomasi 角点检测算法^[8]获取骨架线的角点,最后将得到的角点分为折点和端点,并分别存储在 Point1 集合和 Point2 集合中.

1.3 定位头节点

在 1.2 节中获取了人体骨架线的角点,若直接将骨架线的头部端点作为头节点会导致头节点的定位有半个头长的误差,笔者利用 camshift 运动目标跟踪算法^[9]获取脸部中心位置,以此作为头节点. Camshift 跟踪算法分为 3 个步骤来完成,如图 1 所示.

收稿日期:2017-12-28;修订日期:2018-03-29

作者简介:张青建(1989—),男,河南济源人,杭州电子科技大学硕士研究生,主要从事视频图像处理研究,E-mail: zhangqjhenan@163.com.

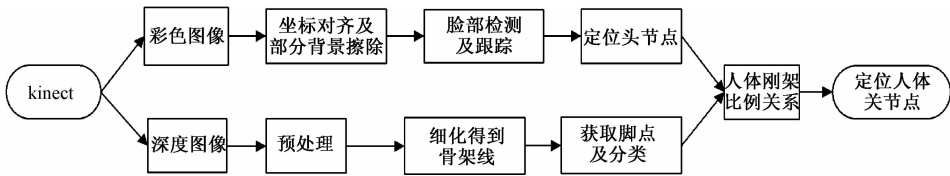


图 1 方法流程图

Fig. 1 Flow chat of this paper method

1.3.1 获取脸部区域(ROI)的颜色概率分布图

为减少光照影响,首先将彩色图像从 RGB 空间转换到 HSV 空间^[10],做出 H 分量直方图,如图 2 所示,直方图代表了 ROI 区域不同颜色在 H 分量上出现的概率,然后将 ROI 区域中每个像素的值用其颜色概率替换,得到颜色概率分布图^[11],如图 3 所示。

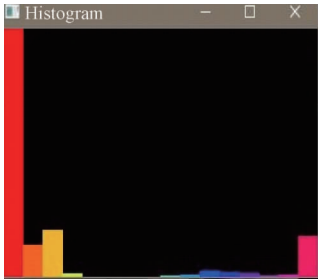


图 2 H 分量直方图

Fig.2 H compont histogram



图 3 颜色概率分布图

Fig.3 The distribution of color probability

1.3.2 meanshift 算法路径寻优

Meanshift 算法^[12]是一种密度函数梯度估计的非参数方法,通过迭代寻优找到概率分布的极值来定位目标.首先在图 3 中选取搜索窗口,计算窗口的质心 (x_c, y_c) ,

$$x_c = M_{10}/M_{00}; y_c = M_{01}/M_{00}, \tag{1}$$

式中: M_{00} 和 M_{10} 、 M_{01} 分别是零阶矩和一阶矩,

$$M_{00} = \sum_x \sum_y I(x, y), \tag{2}$$

$$M_{10} = \sum_x \sum_y xI(x, y), \tag{3}$$

$$M_{01} = \sum_x \sum_y yI(x, y). \tag{4}$$

把搜索窗口的中心设置在计算出的质心处,这样窗口的移动引起窗口内容的变化,继续重复

以上定位窗口中心的步骤,直到搜索窗的中心与质心间的距离小于预设的阈值则停止计算^[13].

1.3.3 camshift 跟踪

Camshift 算法是将视频中的每一帧做 mean-shift 运算,并将上一帧搜索窗的大小和中心,作为下一帧 meanshift 算法搜索窗的初始值.如此迭代下去,实现对运动目标的跟踪.在脸部区域跟踪过程中,将窗口的中心位置作为头节点,实现效果如图 4 所示。

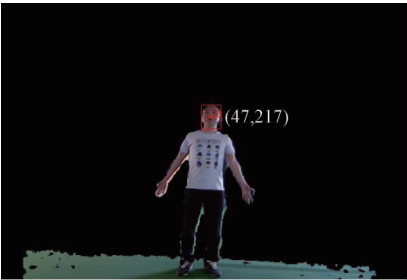


图 4 脸部跟踪效果图

Fig.4 The figure of tracing face

2 人体关节点定位

2.1 中国成年人各部件比例关系

在标准人体骨架模型(如图 5)中剩余的 14 个关节,需要结合骨架线角点和人体各刚件比例关系共同定位得出.《中国成年人人体尺寸》(GB/T 10000—1988)给出了我国成年人人体尺寸的比例关系,如图 6 所示,以此得到笔者所需的人体各刚件比例关系,以 l (头-颈)为基准.

$$\begin{aligned} l(\text{颈-肩}) &= 0.8l(\text{头-颈}), \\ l(\text{肩-肘}) &= 1.5l(\text{头-颈}), \\ l(\text{颈-尾椎}) &= 3.5l(\text{头-颈}), \\ l(\text{尾椎-臀}) &= 0.5l(\text{头-颈}), \\ l(\text{臀-膝}) &= 2l(\text{头-颈}), \\ l(\text{膝-足}) &= 2l(\text{头-颈}), \end{aligned} \tag{5}$$

其中, l (头-颈)为头节点到颈节点的距离; l (颈-肩)为颈节点到肩节点的距离; l (肩-肘)为肩节点到肘节点的距离; l (肘-手)为肘节点到手节点的距离; l (颈-尾椎)为颈节点到尾椎点的距离; l (尾椎-臀)为尾椎点到臀节点的距离;

$l(\text{臀-膝})$ 为臀节点到膝节点的距离; $l(\text{膝-足})$ 为膝节点到足节点的距离.

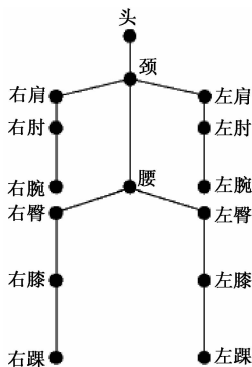


图 5 标准人体骨架模型

Fig.5 Standard human skeleton model

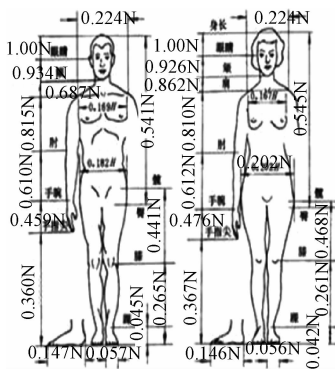


图 6 中国成年人人体尺寸比例关系

Fig.6 The proportion of human body size in China

2.2 关节点定位

上文得到的人体各刚件比例关系具有一般性和普遍性,但是个体之间存在差异,针对某一个体,如果仅仅根据此比例关系,定位出的人体关节点位置的准确率有待提高,所以本文中,笔者结合骨架线角点和人体各刚件比例关系共同定位,得出除头节点外的 14 个人体关节点,具体操作如下.

①头节点. 根据上文 camshift 的脸部跟踪算法获取脸部中心作为头节点的位置,并在端点集合 Point2 中将 y 坐标值大于此头节点且距离最近的点,也就是人体细化后得出的头部顶点,从端点集合 Point2 剔除.

②颈节点. 操作①定位了头节点,则在折点集合 Point1 中,到头节点距离最近的点一定是颈节点,以此可以定位出颈节点的坐标,并记录颈节点到头节点的距离为 $l(\text{头-颈})$. 最后将颈节点从 Point1 集合剔除.

③肩节点. 以颈节点为圆心,以 $l(\text{颈-肩}) \pm 20\%$ 的长度为半径画出两个圆,出现在这两圆之间区域的 Point1 集合中的折点则为肩节点. 由于肩节点的对称性,找到一侧肩节点,另一侧的肩节

点可以对称定位得出.

④肘节点、尾椎点、臀节点、膝节点. 此 3 个节点定位方法与肩节点方法类似,不再赘述.

⑤手节点. 经过上几步的节点剔除操作,此时的端点集合 Point2 中只剩下手节点和足节点,在站立状态下,手节点到肘节点的距离小于足节点到肘节点的距离,所以计算 Point2 集合中剩余端点到某一侧肘节点的距离,距离最小的端点即为该侧的手节点.

⑥足节点. 在端点集合 Point2 中,笔者已经剔除了有可能为头部端点的点,手部端点的点,那么剩下的端点只可能属于足节点,由此定位出左右足节点的位置. 至此,就完成了人体关节点的定位.

3 试验与分析

3.1 试验环境

笔者的试验环境是在室内,光线充足,无刺激性强光,背景环境复杂性一般,试验对象是单人,人距离 kinect 约 1.5 m 处正向面对 kinect 成站立姿势,kinect 距地约 1.2 m,使用 kinect 摄取 270 帧视频,帧率 30 fps,图像分辨率 640×480 ,视频中的试验人作出 4 个基本动作,分别标记为动作 1~4.

3.2 试验过程

由 kinect 摄取人体的彩色图像和深度图像,如图 7 和图 8 显示的是动作 1 的彩色及深度图像,接着实现彩色图像和深度图像的坐标对齐并对彩色图像擦除部分背景,如图 9 所示. 对深度图像进行滤波膨胀等预处理,如图 10 所示. 对预处理后的深度图像细化处理,得到人体骨架线,如图 11 所示. 若不对深度图像预处理,获得的人体骨架线有较多毛刺,如图 12 所示,不符合试验要求. 获取骨架线的角点,并将得到的角点分为折点和端点,并分别存储在 Point1 集合和 Point2 集合中,如图 13 所示,其中,红色为端点,蓝色为折点.



图 7 人体彩色图像

Fig.7 Color image of human body

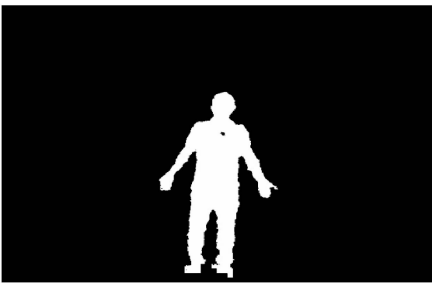


图 8 人体深度图像
Fig. 8 Depth image of human body



图 9 对彩色图像坐标对齐并擦除部分背景
Fig. 9 Align the color image coordinates and erase part of background

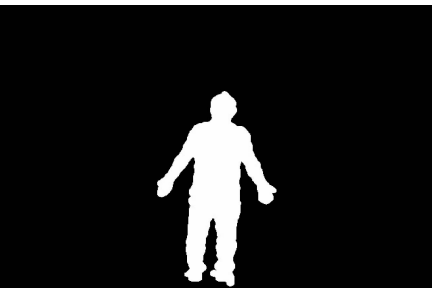


图 10 预处理后的深度图像
Fig. 10 The depth image after pre-processed

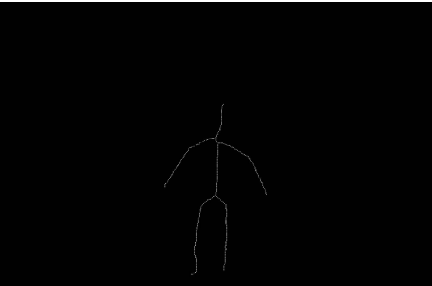


图 11 细化处理得到人体骨架线
Fig. 11 The human skeleton line after Thinning-processed

3.3 试验结果

本试验分别获取动作 1~4 对应的预处理后的深度图,人体骨架线角点分类图,关节点定位图,试验效果如图 14~17 所示。

3.4 结果分析

从试验效果图可以看出,笔者方法能够定位

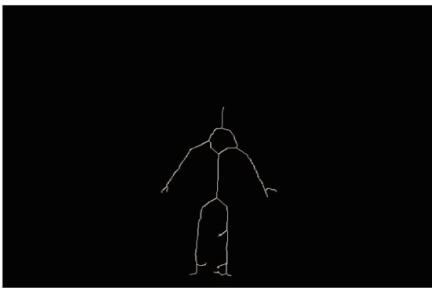


图 12 未经预处理得到的人体骨架线
Fig. 12 The human skeleton line without pre-processed

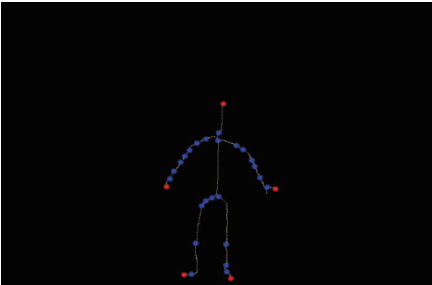


图 13 人体骨架线角点及分类
Fig. 13 The classified corners of human skeleton line

到人体的 15 个关节点. 为了验证本试验关节点定位的准确性,笔者做了两个对比试验:①以人工标记方式获取的关节点坐标为基准,对比笔者方法与文献[2]的方法(简称对比方法),得出两种方法获取的各关节点的平均像素误差;②以文献[1]的人体关节点定位方法(其功能实现已内置于 kinect 产品当中,是目前比较先进的人体关节点定位方法,简称 kinect 方法)为基准,得出笔者方法定位的各关节点的平均正确识别率。

对比一:首先对试验人的各个关节点做人工标记,然后在相同的试验条件下,对同一个动作,用笔者方法和对比方法分别获取对应关节点的坐标,再分别计算出两者的坐标与人工标记的对应关节点坐标间的欧氏距离^[14],若距离小于 15 个像素,则认定识别正确,否则认定识别失败,得到笔者方法和对比方法的关节点定位平均正确识别率,如图 18 所示。

对比二:在相同的试验条件下,同一个动作的对应关节点分别以 kinect 方法和笔者方法定位出关节点坐标,并以 kinect 方法获取的关节点坐标为基准,计算出笔者方法定位的关节点坐标与 kinect 方法获取的对应关节点坐标间的欧氏距离,当两者对应关节点坐标间的欧式距离小于 15 个像素时则认定识别正确,否则认定识别失败,得到笔者方法定位关节点的平均正确识别率,如图 19 所示。

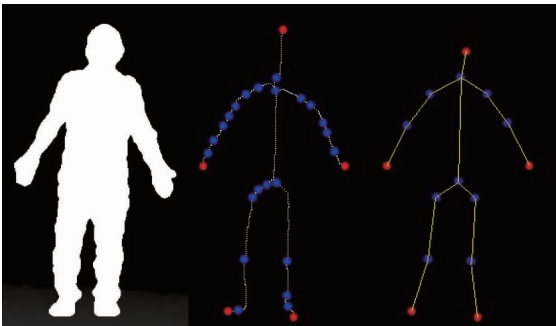


图 14 动作 1
Fig. 14 Action 1

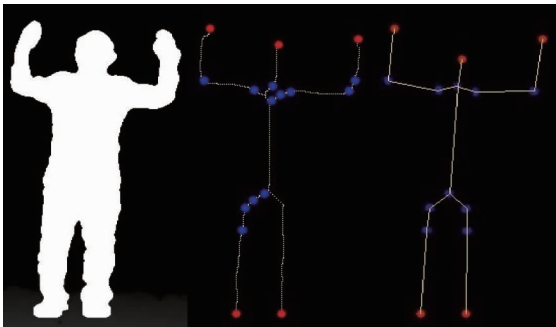


图 15 动作 2
Fig. 15 Action 2

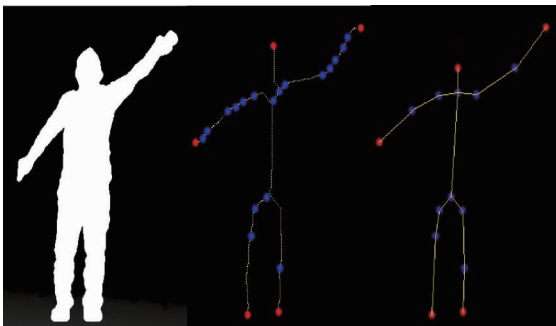


图 16 动作 3
Fig. 16 Action 3

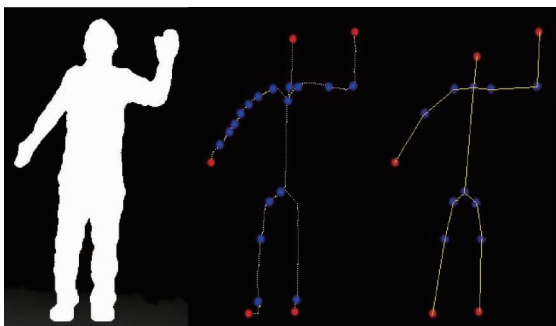


图 17 动作 4
Fig. 17 Action 4

从图 18 可以看出,相较于对比方法,笔者方法定位头节点的误差明显更小,笔者方法定位的其他关节点的误差也相对较小.对比方法的关节点的平均定位准确率是 69.7%,而笔者方

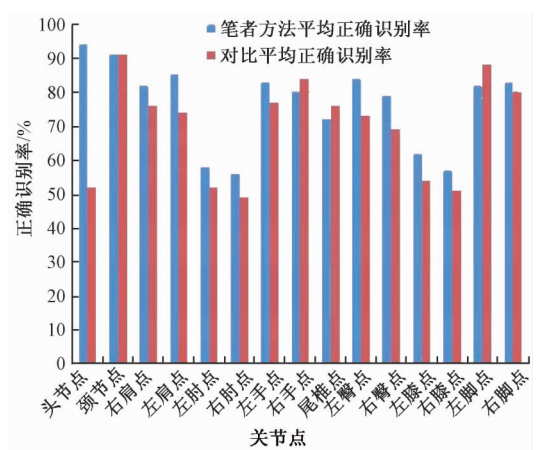


图 18 笔者方法和对比方法的关节点平均正确识别率
Fig. 18 The correct recognition rate of human joints of the proposed method and the comparison method

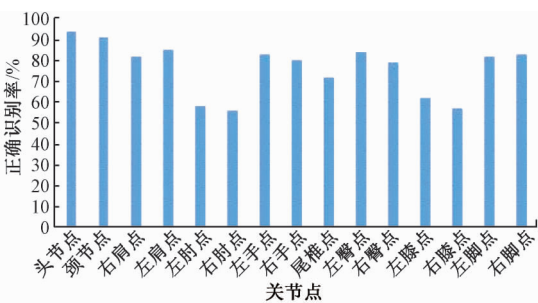


图 19 笔者方法定位的各关节点的平均正确识别率
Fig. 19 The correct recognition rate of human joints of the proposed method

法的关节点平均定位准确率为 76.5%,整体的定位准确率提高了 7% 左右,其中头节点定位准确率提高效果最为明显,达到了 42% .从图 19 可以看出,笔者方法除肘节点、膝节点、尾椎点外,其他节点的定位相较于 kinect 方法差别不大,基本吻合,尾椎点的定位效果一般,是由于人体细化后得到的尾椎点位置偏下所致.肘节点和膝节点的定位效果不够理想,平均正确识别率在 60% 左右,主要原因是当人体双臂伸展或双膝直立无弯曲站立的姿势下,折点定位的误差较大导致的.

4 结论

提出了一种利用深度图像和彩色图像定位人体关节点的方法,笔者的工作主要做了以下改进.

(1) 通过运动目标跟踪获取人体脸部区域中心点作为头节点,防止了直接以骨架线头部端点作为头节点产生的误差.

(2) 获取人体骨架线的角点,并将角点分类为端点和折点,并结合人体刚架比例关系定位出关节点,比直接通过一般性的人体刚架比例关系

定位关节点的方法具有更好的适用性.

笔者提出的人体关节点定位方法,还有较大的改进空间,比如角点检测算法的效率有待提高等.

参考文献:

[1] SHOTTON J, KIPMAN A, et al. Real-time human pose recognition in parts from single depth images[J]. Communications of the ACM, 2013, 56 (1): 116 - 124.

[2] 肖秦琨,赵艳,高嵩. 基于 RGB 彩色和深度信息的人体关节点定位[J]. 国外电子测量技术, 2015 (2): 27 - 30.

[3] 余涛. Kinect 应用开发实战用:最自然的方式与机器对话[M]. 北京: 机械工业出版社, 2012.

[4] KO S J, LEE Y H. Center weighted median filters and their applications to image enhancement [J]. IEEE transactions on circuits and systems, 1991, 38 (9): 984 - 993.

[5] JACKWAY P T, DERICHE M. Scale-Space properties of the multiscale morphological dilation-erosion [J]. IEEE transactions on pattern analysis & machine intelligence, 1996, 18(1): 38 - 51.

[6] 孙少林,马志强,汤伟. 灰度图像二值化算法研究 [J]. 价值工程, 2010(5): 142 - 143.

[7] ZHANG T Y, SUEN C Y. A fast parallel algorithm for

thinning digital patterns [J]. Communications of the ACM, 1984, 27(3): 236 - 239.

[8] SHI J, TOMASI C. Good features to track [C] // 1994 Computer Vision and Pattern Recognition. Seattle: IEEE, 2002:593 - 600.

[9] NOUAR O D, ALI G, RAPHAEL C. Improved object tracking with camshift algorithm [C] // 2006 IEEE International Conference on Acoustics, Speech and Signal Processing. Toulouse: IEEE, 2006: 1121 - 1127.

[10] SOLEIMANIZADEH S, MOHAMAD D, SABA T, et al. Recognition of partially occluded objects based on the three different color spaces (RGB, YCbCr, HSV) [J]. 3D Research, 2015, 6(3): 1 - 10.

[11] 姚红革,郝重阳,雷松则. 序列图像中彩色目标跟踪的加权颜色分布方法 [J]. 中国图象图形学报, 2009(1): 99 - 105.

[12] COMANICIU D, MEER P. Meanshift: a robust approach toward feature space analysis [J]. IEEE transactions on pattern analysis and machine intelligence, 2002, 24(5): 603 - 619.

[13] GARY B, ADRIAN K. 学习 OpenCV(中文版) [M]. 北京: 清华大学出版社, 2012.

[14] WANG L, ZHANG Y, FENG J. On the euclidean distance of images [J]. IEEE transactions on pattern analysis & machine intelligence, 2005, 27(8):1334 - 1339.

A Method of Positioning the Human Body Joints Based on RGB and Depth Information

ZHANG Qingjian, HAN Jianping

(School of Computer, HangzhouDianzi University, Hangzhou 310016, China)

Abstract: In this paper, a method of positioning the human joints based on RGB-D was proposed. The human body skeleton line was obtained by the deep image filtering, dilating and thinning processing of the kinect acquisition, the end and turning points of the human skeleton line were obtained according to the corner detection algorithm. Then tracked and positioned the human's face in the color image, at last the 15 joints of human body were positioned combining with the proportion of the rigid parts of the human body. The results showed that the accuracy of positioning joints of this method was improved by 7% compared with the method of positioning the joints directly according to the proportion of the rigid parts of the human body, and the accuracy of positioning the joint of head was improved by 42%.

Key words: RGB-D; kinect; image thinning; corner detection; facing tracking; positioning joints