

基于 RGB 彩色和深度信息的人体关节点定位

肖秦琨 赵艳 高嵩

(西安工业大学电子信息工程学院 西安 710021)

摘要: 在人体实时定位问题中,关节点定位成为了构建人体平面骨架模型、识别人体动作中一个重要研究课题。所以,提出了一种利用 RGB 彩色和深度信息来定位人体关节点的方法,首先利用 Kinect 来获取运动人体的 RGB 彩色和深度图像,接着根据人体所处背景的复杂程度来采用不同的人体分割方法将人体分割出来,再对得到的人体轮廓图像进行平滑处理,从而得到平滑且连通的人体轮廓二值图像,再采用细化算法提取出人体轮廓骨架,然后对关节点周围 8 个邻域的像素值进行分析,并结合人体各刚体部件的比例关系,对人体的 16 个关节点进行定位,进而构建出人体平面关节点模型。实验结果表明,与现有方法相比,该方法简单有效,能够从多种运动状态人体中提取出准确的关节点坐标,并具有良好的准确性和鲁棒性。

关键词: Kinect; 人体轮廓二值图像; 细化算法; 人体平面关节点模型

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.99

Human body joint points calibration based on the RGB and depth information

Xiao Qinkun Zhao Yan Gao Song

(College of Electronic Information and Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: The calibration of joint points is an important research topic in the solution of construction of human two-dimensional skeleton of models and recognition of human actions. In this paper, a novel approach about human body joint points calibration based on the RGB color and depth information was proposed. First of all, with the help of Kinect to get RGB color images and depth images, depending on the complexity of human background to divide the human body segmentation with different methods, and smoothing the obtained body contour images, body contour binary images are smooth and connected. Then body outline frames are extracted by thinning algorithm. Analyzing the joint 8-neighborhood pixel values, and combining with proportions of rigid body parts, sixteen joint points are calibrated. Finally, human two-dimensional joint of models are built. Experimental results show that this approach is simple and available compared with existing methods. Joint points are able to be located from human body with different states of motions, and with good accuracy and robustness.

Keywords: Kinect; body contour binary images; thinning algorithm; human two-dimensional joint of models

1 引言

近年来,人体动作分析^[1]广泛的应用于人机交互、虚拟现实、视频监控等领域。随着图像处理技术和计算机视觉的发展,从无标记的图像序列中定位人体关节点来对人体动作进行分析也受到了广泛关注。由于关节点在构建人体骨架中起着关键的作用,而骨架可以有效的反应出人体所处的运动状态,因此人体关节点模型包含了可用于人体动作识别^[2]的重要特征,它可以用来描述人体的运动信

息,作为识别人体动作的有效方法。

由于人体运动的复杂性,因此现有的研究方法对人体和背景^[3]均加以限制,如背景单一、人体的服装颜色不要与背景颜色相同或相近等。故直接从人体前景图像中提取出人体平面关节点较为困难,必须先得到人体二值图像,再进行相应地处理才能较好的提取出人体平面关节点。以往定位人体关节点的方法是通过标记点进行跟踪,定位标记点的方法主要分为两类:一类是利用传感器获取标记点位置;另一类是利用数字图像处理技术^[4]定位

标记点。利用上述方法在很多应用上虽获得了成功,但通过仪器对标记点进行定位通常会使得人体运动变得不自然,而在对标记点进行识别与定位过程中,这些标记点可能由于变形导致定位不准确,从而对测试效果造成影响。

提出的基于 RGB 彩色和深度信息来定位人体关节点的方法,使用 Kinect 摄像机采集运动人体的 RGB 彩色和深度图像,再进行背景差分、平滑等预处理后,采用细化算法提取出人体轮廓骨架,且细化后的人体骨架为单位像素宽。实验证明该算法鲁棒性好,算法效率高,容易实现。

2 基于 RGB 彩色和深度信息的人体骨架提取算法

2.1 信息初始采集与分割

本文对 RGB 彩色与深度信息采集使用的设备是微软公司推出的一款 Xbox360 的体感外设——Kinect^[5],使用该设备可以同时获取物体的 RGB 彩色和深度信息^[6]。Kinect 有 3 个摄像头:中间的 RGB 彩色摄像头、两边的红外发射器和 CMOS 摄像机。其中 RGB 镜头可以采集到 RGB 彩色图像,分辨率为 640×480 ;红外发射器和红外 CMOS 摄像头可以采集到 3D 深度图像,分辨率为 320×240 。

其中,RGB 彩色空间^[7]是由 R(Red)、G(Green)、B(Blue)三基色来表示,其他任何一种颜色则是由这三分量按照一定的比例关系组成。此外,RGB 彩色空间中的色度学标准主要有以下 3 条:

1)混合色的亮度是其 3 个分量 R、G、B 颜色亮度之和。

2) R、G、B 3 个分量按照一定比例混合可以形成任何一种颜色,并且此混合得到的颜色唯一。

3)只要 2 种颜色的 R、G、B 分量对应相等,那么这 2 种颜色必然是同一种颜色,并且每种分量以相同的倍数增加或以相同的缩减比例缩小,形成的颜色也都是相同的。

在 RGB 彩色空间中,RGB 彩色图像中的每一个像素都是由 R、G、B 三基色按照不同的比例混合而成的。而深度图像也称距离图像,指从观察视角看,图像包含信息与场景物体表面距离相关的一种图像或一个图像通道。

接下来将所采集到的 RGB 彩色图像中的人体分割出来,具体流程图如图 1 所示。本文针对人体所处背景的复杂程度^[8],分别采用了不同的人体分割方法。当 RGB 彩色图像背景简单时,利用背景减除法^[9]直接将 RGB 彩色图像的当前帧 $I_k(x,y)$ 与已知背景 $I_{bg}(x,y)$ 做差,则差分后图像 $d_k(x,y) = I_k(x,y) - I_{bg}(x,y)$;当 RGB 彩色图像背景复杂^[10]时,利用深度数据预先设定一个合理的阈值 T_k ,把当前像素 $d_k(x,y)$ 同预先给定阈值进行比较,当

$|d_k(x,y)| > T_k$ 时, (x,y) 为背景,当 $|d_k(x,y)| \leq T_k$ 时, (x,y) 为运动目标,然后再将 RGB 彩色图像中的背景去掉,从而将人体分割出来,得到分割后的人体轮廓,如图 2 所示。

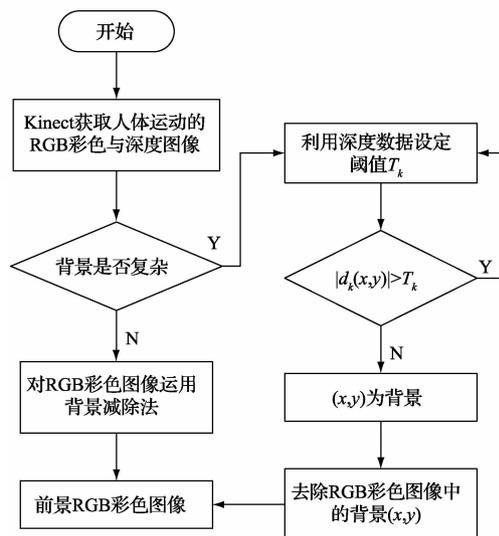


图1 信息初始采集与分割流程



图2 分割后的人体轮廓

2.2 人体骨架提取

2.2.1 人体轮廓图像的平滑处理

由于得到的人体轮廓图像并不光滑,这样会使得提取的人体骨架中有较多的分支点,因此,需要先对获得的人体轮廓图像进行滤波。采取对轮廓图像进行 2 次滤波的方法,先进行中值滤波^[11],接着对轮廓图像进行降采样,将其高度和宽度均缩小 4 倍,然后再对其进行一次滤波。

2.2.2 人体轮廓骨架的提取

本文采用细化算法^[12]提取轮廓骨架^[13],且该细化算法满足 3 个条件:

- 1)连通性:整个骨架连通在一起,不能中间断开;
- 2)单线性:整个骨架的宽度只有单位像素宽;
- 3)简化性:骨架越简化越好,分支点越少越好。

运行细化算法后,得到的二值图像如图 3 所示,而在二维空间中,一幅二值图像可被认为仅仅包括背景和目标 2 种像素,其中背景像素和目标像素的灰度值分别为 0 和 1。此外,一幅 $M \times N$ 的二值图像可表示成一个二维数组 $P[M,N] = [p_{xy}]$,每个像素 p_{xy} 的坐标为 (x,y) ,其中, $p_{xy} = 0$ 的像素为背景点, $p_{xy} = 1$ 的像素为目标点。

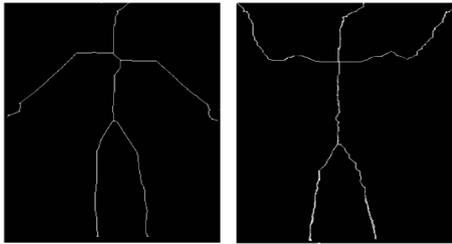


图3 人体骨架二值

得到人体骨架二值图后,将其以二维数组的形式存储起来。如果像素 (x, y) 是白色的,那么 $p_{xy}=1$,否则 $p_{xy}=0$ 。在此,定义二值图像左上角为坐标原点,水平向右为X轴正方向,竖直向下为Y轴正方向。

2.3 人体关节点定位算法

为了方便使用骨架,引用标准骨架模型,模型如图4所示。此模型能很好地表示人体结构,能适应于任意形变,具有很好的通用性。因此,选择此模型作为标准人体骨架。为了建立标准人体骨架,最为关键的是对关节点进行定位,找准每个关节点,再用线把关节点连接起来,就能获得标准人体骨架。

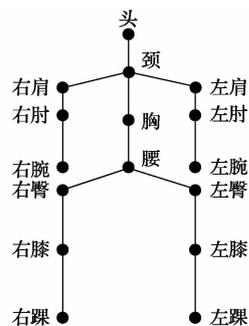


图4 标准人体骨架模型

设被检测的像素为 P ,其灰度值为 $f(p)$,且其八邻域为 $N_i(i=1,2,3,\dots,8)$, N_i 的位置如表1所示。

表1 P 的八邻域

N_1	N_2	N_3
N_8	P	N_4
N_7	N_6	N_5

遍历整幅人体骨架线图像,找到灰度值 $f(p)=1$ 的目标像素 P ,并通过式(1)计算出 P 的八邻域^[13]中非零像素个数 $N(P)$ 。

$$N(P) = f(N_1) + f(N_2) + f(N_3) + f(N_4) + f(N_5) + f(N_6) + f(N_7) + f(N_8) \quad (1)$$

1) 定位头部和手脚关节点

若像素点 P 的八邻域中有且只有一个像素点的灰度值大于0,即 $N(P)=1$,那么该像素点 P 就是人体骨架

线上的末端关节点,即头部或手脚关节点。接着找出这5个像素点的坐标位置,根据人体头部以及手脚的位置关系来确定它们的准确位置,其中纵坐标 y 值最小的像素点为头部关节点, y 值最大的像素点为脚部关节点, y 值介于二者之间的像素点为手部关节点。

2) 定位颈部和腰部关节点

若像素点 P 的八邻域中只有4个邻接点的灰度值大于0,即 $N(P)=4$,那么像素点 P 就是人体骨架线上的颈部关节点;若像素点 P 的八邻域中只有3个邻接点的灰度值大于0,即 $N(P)=3$,那么像素点 P 就是人体骨架线上的腰部关节点。然后提取出颈部和腰部关节点的坐标。

3) 定位其他关节点

在找出其他关节点之前,本文先提出一种“线上最短距离”^[14]算法。“线上最短距离”是指2个关节点之间的以线段相连的最短线段长度。如果2点间有多条路径,则取最短的距离^[15]。它保证了只在线段上计算距离,而且是最短距离。

首先把图像复制到矩阵中,标记白色像素点为1,黑色像素点为0。以颈部关节点开始,在此关节点处,把距离加1,判断此点,如果没有到达左手关节点,则进行递归运算,如果到达左手关节点,则返回颈部到左手的距离。如果递归过程中某点上、下、左、右都没有白色点,则返回-1,继续递归,直到所有路径已计算过。同样的方法,可以找到颈部到右手、腰部到左脚和腰部到右脚的距離。

利用以上得到的一些距离和人体各刚体部件的比例关系,再结合来“线上最短距离”算法,就可以提取出人体其他关节点。人体各刚体部件比例关系如下:

$$L_{Neck-Shoulder} : L_{Shoulder-Elbow} : L_{Elbow-Hand} = 1 : 1 : 2$$

$$L_{Waist-Hip} : L_{Hip-Knee} : L_{Knee-Foot} = 1 : 2 : 2$$

式中: $L_{Neck-Shoulder}$ 为从脖子到肩的最短距离, $L_{Shoulder-Elbow}$ 为从肩到肘的最短距离, $L_{Elbow-Hand}$ 为从肘到手掌的最短距离, $L_{Waist-Hip}$ 为腰到臀的最短距离, $L_{Hip-Knee}$ 为臀到膝的最短距离, $L_{Knee-Foot}$ 为膝到脚的最短距离。

而胸部关节点为颈部和腰部中点。所以,就提取出了人体其他关节点,并将所有关节点联接起来,就得到了人体平面骨架。

3 实验结果

实验使用Kinect设备采集运动人体的RGB彩色和深度图像,其场景为正常环境下室内任意角落,图像中的背景可以简单也可以复杂,图像分辨率为 640×480 ,每秒最多可获取300帧图像。实验结果主要分为3部分,即运动人体轮廓提取、运动人体骨架线提取和运动人体关节点定位,如图5所示。

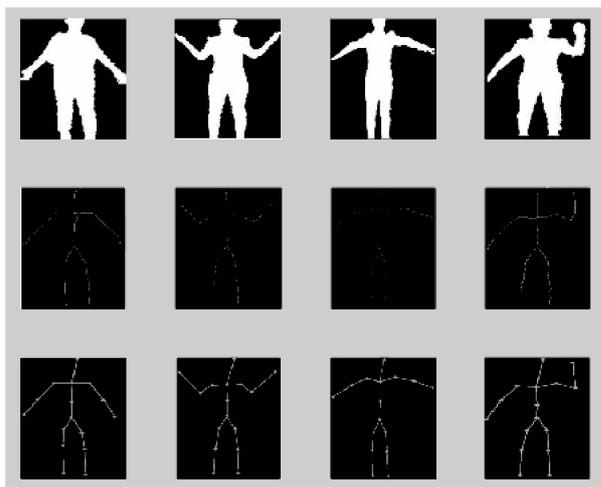


图5 实验结果

通过实验证明,本文定位人体关节的方法具有以下特点:

1) 本文利用细化法得到人体骨架,而细化后的人体骨架为单位像素宽,所以,该方法可以非常鲁棒地实现人体关节的定位;

2) 本文所使用的方法无须任何手工干预,能够自动地将运动人体关节定位出来;

3) 该方法具有普遍适用性,能够从不同运动状态的人体 RGB 彩色和深度图像中定位出人体关节。但在某些方面,此方法仍须进一步完善,当运动人体出现自遮挡时,细化后的人体骨架线就会出现肢体不全等现象,那么利用本文方法将无法得到全部的人体平面关节。

4 结论

提出了一种利用 RGB 彩色和深度信息来定位人体关节的方法。首先利用细化法对人体轮廓二值图像进行细化处理,得到的人体骨架为单位像素宽,再对关节周围八个邻域的像素值进行分析,并结合人体各刚体部件的比例关系,从而提取出人体的 16 个关节。该方法保证了人体关节定位的准确性,并为人体动作识别奠定了坚实的基础。但如果运动人体出现自遮挡时,那么利用该方法将会丢失部分关节。为了解决该问题,需要利用深度图像找出被遮挡的肢体部分,进而确定丢失的关节,这也是在今后研究中需要做的工作。

参考文献

[1] 刘开余,夏斌. 基于 Kinect 的实时人体姿势识别[J]. 电子设计工程,2014,22(19):31-34.

- [2] 李红岩,毛征,袁建建,等. 一种基于算法融合的运动目标跟踪算法[J]. 国外电子测量技术,2013,32(12): 36-40.
- [3] 孙斌,黄神治. 移动背景下运动目标检测与跟踪技术研究[J]. 电子测量与仪器学报,2011,25(3): 206-210.
- [4] XIAO ZH D, NAIT-CHARIF H, ZHANG J J. Automatic estimation of skeletal motion from optical motion capture data [J]. Computer Science, 2008: 144-153.
- [5] 石曼银. Kinect 技术与工作原理的研究[J]. 哈尔滨师范大学自然科学学报,2013(3): 83-86.
- [6] SHOTTON J, SHARP T, KIPMAN A, et al. Real-time human pose recognition in parts from single depth images[J]. Communication of the ACM, 2013, 56(1):116-124.
- [7] JAYARAM S, SCHMUGGE S, SHIN M C, et al. Effect of color space transformation, the illuminance component, and color modeling on skin detection[J]. Communication of the ACM, 2013,56(1):116-124.
- [8] 李响,谭南林,王天雷,等. 复杂场景下基于局部运动补偿的目标检测[J]. 仪器仪表学报,2014,35(7): 1555-1563.
- [9] 李宁,黄山,张先震,等. 基于背景差分的人体运动检测[J]. 微计算机信息,2009,25(21):256-257.
- [10] 雷涛,罗薇薇,樊养余,等. 复杂环境下的运动人体骨架提取算法[J]. 计算机应用研究,2010,27(8): 3194-3197.
- [11] 陈健,郑绍华,余轮,等. 基于方向的多阈值自适应中值滤波改进算法[J]. 电子测量与仪器学报,2013, 27(2): 156-161.
- [12] 余樑,王广伟. 二值图像的快速细化算法[J]. 科技世界,2012,5(14): 111-112.
- [13] 蔡建军,孔令富,李海涛. 基于欧式距离变换的人体 2D 关节点标定[J]. 计算机仿真,2012,29(7): 243-246.
- [14] 程洪,庄浩洋. 3D 人机交互中的骨架提取和动态手势识别[D]. 成都:电子科技大学,2013:21-22.
- [15] 魏玮,王丹丹,刘静,等. 无需人工干预的关节点提取和三维重建[J]. 计算机科学,2013,40(4): 292-294.

作者简介

肖秦琨,1973 年出生,工学博士后,教授,主要研究方向为模式识别、计算机动画与多媒体分析与处理等。

赵艳,1986 年出生,工学硕士,学生,主要研究方向为模式识别、计算机动画与多媒体分析与处理。

E-mail: xiaoqinkun10000@163.com