

基于 Kinect 骨骼数据的人体动作姿势识别方法

李红波, 李双生, 孙舶源

(重庆邮电大学 计算机科学与技术学院, 重庆 400065)

摘要: 针对现有的复杂背景下人体动作姿势训练与识别中存在识别准确率不高和实时性不强等问题, 提出一种基于 Kinect 骨骼数据的人体动作姿势识别方法。从 Kinect 获取骨骼姿势特征点数据, 计算姿势特征矢量; 将当前的实时姿势特征矢量与预设的标准姿势特征矢量进行匹配比较, 实时反馈两者姿势的匹配度, 达到姿势识别的目的。实验结果表明, 该方法能有效地进行人体动作姿势的匹配识别, 识别准确率较高、实时性较强, 具有良好的可扩展性。

关键词: Kinect; 姿势识别; 模版匹配; 骨骼数据; 特征矢量

中图法分类号: TP391.41 **文献标识号:** A **文章编号:** 1000-7024 (2016) 04-0969-07

doi: 10.16208/j.issn1000-7024.2014.04.025

Posture recognition method based on Kinect skeleton data

LI Hong-bo, LI Shuang-sheng, SUN Bo-yuan

(College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: As the traditional methods of posture recognition are neither accurate nor efficient enough in complex environmental background, a human posture recognition method based on Kinect skeleton data was proposed. Posture characteristic points were abstracted from skeleton data obtained using Kinect, and the feature vectors of postures were calculated. The real-time posture feature vectors were matched with the presupposed feature vectors and the matching results were fed back in real time. The recognition was finished if there was a match satisfied. Experimental results show that the proposed method can accomplish posture recognition accurately, and it has high correct recognition rate and performs well on real-time capability and scalability.

Key words: Kinect; posture recognition; template matching; skeleton data; feature vector

0 引言

人体动作姿势识别针对的是静态对象, 比如人体体势或某个身体部位的姿态识别。本文的研究对象是人体整体的动作姿势。常用的人体动作姿势识别方法主要有 3 类^[1], 本文所使用的方法则是基于模板匹配的识别算法。

在现有的人体动作姿势识别研究中, 大多数工作是基于传统的单目视觉进行的^[2]。这些基于传统视觉的人体动作姿势识别过程易受光照、背景等因素的影响, 其识别的准确性和鲁棒性都不好。微软发布了一款基于 Windows 平台的 Kinect 摄像机^[3], 它不仅可以提供传统的彩色图像, 同时还可以提供物体的深度信息。在 Kinect for Windows SDK 支持下, 研究者可以获得人体骨架中 20 个骨骼关节点

信息^[4]。利用 Kinect 获取的人体骨骼关节点三维坐标数据进行人体动作姿势识别, 既可以避免光照变化的影响, 也可以免去在复杂背景下对目标对象进行分割提取等预处理过程。在基于 Kinect 的人体动作姿势识别中^[5-10], L. Miranda 等^[5]使用骨架关节间的角度来描述每种人体动作姿势, 并根据预先定制和训练过的姿势内核, 由多级支持向量机 (support vector machine, SVM) 进行姿势分类, 最后采用决策森林进行人体动作姿势识别。该方法的识别率依赖于姿势内核的制定及训练, 且对相似姿势的识别效果一般。T. Y. Lin 等^[6]提出的基于 Kinect 的物理康复辅助系统中, 通过计算标准动作和学习动作相应关节点间欧氏距离的平均值及骨骼关节间角度差值的平均值, 以此来评价康复者学习太极姿势的准确程度。该方法虽然能够较好

收稿日期: 2015-04-22; 修订日期: 2015-07-05

基金项目: 重庆市自然科学基金项目 (CSTC, 2012jjA40027); 重庆教委科学技术研究基金项目 (KJ130518); 国家社会科学基金项目 (13CGL146)

作者简介: 李红波 (1970-), 男, 甘肃兰州人, 硕士, 高级工程师, CCF 会员, 研究方向为数字媒体技术、机器视觉、增强现实; 李双生 (1989-), 男, 四川广元人, 硕士研究生, 研究方向为数字媒体技术、模式识别; 孙舶源 (1987-), 男, 河南许昌人, 硕士研究生, 研究方向为数字媒体技术、虚拟现实。E-mail: lihongbo@cqupt.edu.cn

地判定人体动作姿势的匹配程度,但实时性较差。

针对以上现有的人体动作姿势识别方法存在对环境背景要求较高、识别率不高和实时性不强等问题,本文提出了一种基于 Kinect 骨骼数据的人体动作姿势识别方法。该方法主要适用于对特定人体动作姿势进行模仿和学习的训练或游戏中。

1 相关技术

1.1 Kinect 骨骼数据简介

Kinect 摄像机能获取到人体动作姿势的彩色图像和深度图像,而文本研究中所使用的数据源为通过 Kinect SDK 提供的应用程序编程接口(application programming interface, API)获取到的 20 个人体骨骼关节点的三维坐标信息^[4],它们能完整且有效地表示人体构造以及各种动作姿势。

Kinect 摄像机捕获到的人体骨骼关节点分布如图 1 所示。其中,黑色圆圈代表的骨骼关节点依次为左手腕、左手肘、左肩膀、肩膀中心、右肩膀、右手肘、右手腕、脊柱、臀部中心、左臀、右臀、左膝、右膝、左脚踝、右脚踝、左脚、右脚。

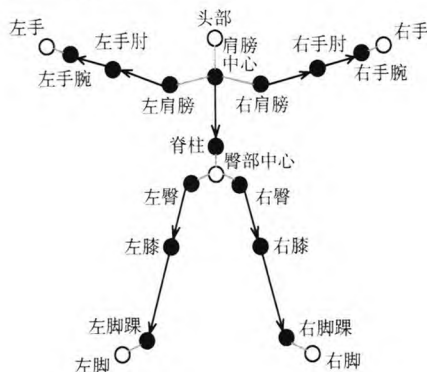


图 1 人体骨骼点分布

1.2 基于模型的姿势识别方法

在基于模型的姿势识别方法中, H. Jiang 等^[11]提出了一种基于一致性最大覆盖的单帧图像姿态估计方法。该方法结合边缘、颜色对称和肢体联动等线索,对经分割得到的人体动作姿势目标图像采用多个矩形块进行一致性最大覆盖处理,人体的每个身体部位(如头部,手臂,腿部,躯干部分等)对应一个矩形块,从而将人体动作姿势识别问题转化为矩形块覆盖与原分割图像的一致性匹配问题,使得根据各矩形块的分布识别人体动作姿势是一个线性可解的问题。一致性最大覆盖法可定义为如下公式的优化问题

$$\max_c \left\{ \sum_{(x,y) \in I} r_{x,y} - \alpha M(c) - \beta P(c) - \gamma S(c) \right\} \quad (1)$$

式中: $\left\{ \sum_{(x,y) \in I} r_{x,y} \right\}$ ——所有被人体部位所覆盖的像素点潜在值总和, $M(c)$ ——身体部位与局部图像特征匹配的损耗

值, $P(c)$ ——身体部位遵从人类身体构造的配置度, $S(c)$ ——对称身体部位间色差的惩罚因子,对称部位间色差越大, $S(c)$ 值越大。 α, β 和 γ 为非负常量值,用以控制各参数的权重。

式(1)中描述的一致性最大覆盖是一个 NP 难度的组合搜索问题, H. Jiang 等^[11]提出一种混合整型线性规划方法,优化后的一致性最大覆盖公式如下

$$\begin{aligned} & \max \left\{ \sum_{(x,y) \in I} r_{x,y} - \alpha \sum_{n \in V, i \in T(n)} c_{n,i} \cdot \eta_{n,i} - \right. \\ & \quad \beta \sum_{(n,m) \in E, i \in T(n), j \in T(m)} h_{n,i,m,j} \cdot \xi_{n,i,m,j} - \\ & \quad \left. \gamma \sum_{(n,m) \in H} \sum_{k=1}^3 (g_{n,m,k}^+ + g_{n,m,k}^-) \right\} \quad (2) \end{aligned}$$

式(2)中各变量定义请参见文献[11]。通过这种放宽法,将式(1)问题的求解转化成一一系列的线性规划问题的求解,最终求得表示身体部位的各矩形块在人体目标前景图像中的最佳一致性最大覆盖。

1.3 基于 Kinect 的姿势识别方法

T. Y. Lin 等^[6]提出了一种基于 Kinect 的物理康复辅助系统,主要用于帮助有运动障碍的病人进行运动康复训练。康复系统首先需要用 Kinect 录入标准的康复运动姿势形成数据库,然后康复者根据系统演示的动作姿势进行模仿学习,系统通过对数据库中标准姿势和病人模仿姿势进行计算比较,反馈给病人两者动作姿势的匹配程度。

标准姿势和模仿姿势是通过 Kinect 的骨骼数据信息进行匹配识别的。考虑到人的位置和身材不尽相同,所以首先需要在三维笛卡尔坐标系中使用齐次坐标系统对骨骼关节点进行几何转换,骨骼关节点的三维坐标点几何转换公式^[12]如下

$$\begin{aligned} \mathbf{v}' &= \mathbf{v}' + \mathbf{d} = \begin{bmatrix} 1 & 0 & 0 & dx \\ 0 & 1 & 0 & dy \\ 0 & 0 & 1 & dz \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \\ &= [x + dx, y + dy, z + dz, 1]^T \quad (3) \end{aligned}$$

式中: \mathbf{v} ——初始的骨骼关节点坐标, \mathbf{v}' ——转换后的骨骼关节点坐标。通过这种归一化处理,两个进行匹配比较的姿势骨架具有一致的身体比例。在匹配识别时,首先分别从两个姿势骨架中提取出头-肩-手部位的 10 个骨骼点。设标准姿势骨架为 A , 提取出的骨骼关节点集合为 $\mathbf{p}_A = \{a_1, a_2, a_3 \cdots a_{10}\}$, 模仿姿势为 B , 提取出的骨骼关节点集合为 $\mathbf{p}_B = \{b_1, b_2, b_3 \cdots b_{10}\}$, 则 \mathbf{p}_A 和 \mathbf{p}_B 对应骨骼关节点间欧式距离的平均值^[6]为

$$d_{joint} = \frac{1}{10} \sum_{i=1}^{10} \|a_i - b_i\| \quad (4)$$

依据所提取出的骨骼关节点集合,可以计算出相邻骨骼关节间形成的角度值。设标准姿势的角度值集合为 $\theta_A = \{\alpha_1, \alpha_2, \alpha_3 \cdots \alpha_9\}$, 模仿姿势的角度值集合为 $\theta_B = \{\beta_1, \beta_2,$

$\beta_3 \cdots \beta_9$ }, 则两者骨架间对应角度的差值平均值^[6]为

$$d_{angle} = \frac{1}{9} \sum_{k=1}^9 |\alpha_k - \beta_k| \quad (5)$$

这样, 通过比较归一化后的标准姿势与模仿姿势骨架间欧氏距离的平均值和相应骨骼关节间角度差值的平均值来识别和评价康复者的模仿手势或姿势。

2 基于 Kinect 骨骼数据的姿势识别方法

2.1 问题提出和总体方案

H. Jiang 等^[11]提出的基于一致性最大覆盖的单帧图像姿态估计方法, 虽然能提高对复杂姿势的识别率, 但其平均识别率比其它传统的识别方法并无明显的提升; T. Y. Lin 等^[6]提出的基于 Kinect 的物理康复辅助系统中, 在进行标准姿势和病人姿势的匹配识别之前, 需要将各自的骨骼关节点坐标几何转换到以躯干为中心原点的局部坐标系中, 再分别利用式 (4) 和式 (5) 进行两者姿势骨架的匹配比较。该方法虽然能够较好地识别和判定人体动作姿势的匹配准确程度, 但不能较好地满足人体动作姿势识别的实时性要求。

针对以上两类问题, 本文提出了一种较新颖的基于 Kinect 骨骼数据的人体动作姿势识别方法。该方法通过由肢体特征矢量集合组成的姿势特征描述符来唯一地表征人体动作姿势, 并通过标准姿势和实时姿势特征描述符中对应特征矢量间的角度值来表达两者姿势之间的差异程度, 并根据这种差异度的大小来判别人体动作姿势是否匹配识别。本文所提出的人体动作姿势识别方法的流程如图 2 所示。

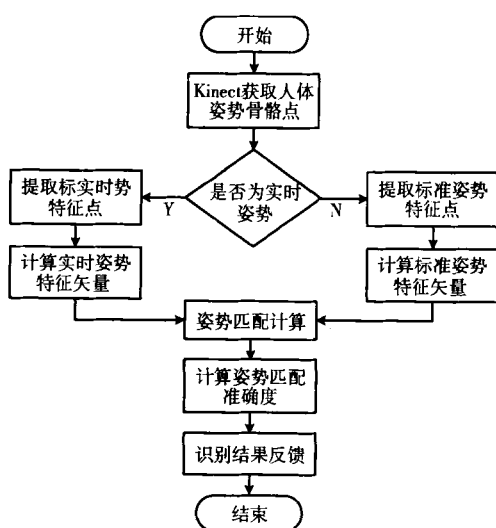


图 2 本文人体动作姿势识别方法流程

2.2 姿势特征描述符提取

通过 Kinect SDK 的“骨骼跟踪”功能可以获得 20 个人体骨骼关节点 (如图 1 所示), 但从运动力学的角度来

说, 各个关节点有着不同的自由度, 所以对人体动作姿势的描述贡献度上存在着差异。在不影响对动作姿势的完整性表达的基础上, 可以舍去一些在动作姿势表达中作用不明显的骨骼关节点。因此, 在提取姿势特征点时, 应该考虑以下姿势特征描述符应具有的性质^[13]: ①唯一性: 每个动作姿势对应一个唯一的姿势特征描述符; ②完整性: 每个姿势特征描述符必须包含一个动作姿势的基本特征; ③几何变换下的不变性: 姿势特征描述符对动作姿势位置的变化和身材大小的差异保持不变的特点; ④敏感性: 姿势特征描述符能容易地区分相似的动作姿势。

根据以上姿势特征描述符的性质, 考虑到人体运动和身体结构的特征, 本文提取出在人体动作姿势中表现活跃的 14 个骨骼关节点 (如图 1 黑色圆圈所示) 作为动作姿势特征点。通过去除人体骨架中冗余的骨骼关节点, 将欲处理的骨骼关节点个数减少至 14 个, 达到降低数据维数的目的。然而, 所提取出的动作姿势特征点的位置被限制在 Kinect 设备的坐标系中, 可以通过肢体特征矢量来解决这一问题。设相邻的两个动作姿势特征点分别为 (x_0, y_0, z_0) 和 (x_1, y_1, z_1) , 则对应的特征矢量值为

$$\mathbf{n} = (x_0, y_0, z_0) - (x_1, y_1, z_1) \quad (6)$$

通过式 (6) 进行特征矢量的计算, 得到一个由 9 个特征矢量组成的集合 (如图 1 黑色箭头所示) $\mathbf{P} = \{\mathbf{n}_0, \mathbf{n}_1, \mathbf{n}_2 \cdots \mathbf{n}_8\}$, 称这个矢量集合为动作姿势特征描述符。由矢量集合表示的姿势特征描述符可以唯一地表示一种人体动作姿势, 且在平移和缩放时保持不变, 身体运动也可以由每个特征矢量绕其起始关节点旋转模拟实现。

2.3 人体动作姿势匹配识别

得到动作姿势特征描述符后, 标准姿势和实时姿势的匹配识别问题就转化为比较两者姿势特征描述符之间的差异性。由于人体动作姿势是通过姿势特征描述符中的每一个特征矢量的旋转来表示的, 所以可以通过比较两者姿势特征描述符中对应特征矢量间的角度值来判断标准姿势和实时姿势的差异度。

如图 3 所示, 设标准姿势特征描述符中某一特征矢量为 $\mathbf{a}(x_s, y_s, z_s)$, 实时姿势特征描述符中对应的特征矢量为 $\mathbf{b}(x_r, y_r, z_r)$, 则两者之间的角度值的余弦值为

$$\cos \alpha_i = \frac{x_s \cdot x_r + y_s \cdot y_r + z_s \cdot z_r}{\sqrt{x_s^2 + y_s^2 + z_s^2} \cdot \sqrt{x_r^2 + y_r^2 + z_r^2}} \quad (7)$$

角度值 α_i 的取值范围为 $[0, 180^\circ]$, 该范围内的余弦函数值为单调递减的, 因此每个角度值唯一地对应一个余弦值。本文之所以使用角度值而非弧度值, 是因为角度值比弧度值有更大的数值取值范围, 能更为精确地反映出相应的特征矢量之间的差异性。通过式 (7) 对特征矢量间角度值的计算, 可得到一个角度值集合 $\theta = \{\alpha_0, \alpha_1, \alpha_2 \cdots \alpha_8\}$, 这个角度值集合就表示了标准姿势和实时姿势之间的差异程度。求得各角度值之和为

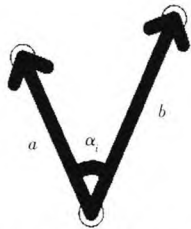


图3 特征矢量间角度值

$$\alpha_{si} = \sum_{i=0}^8 \alpha_i \quad (8)$$

由于特征矢量间角度值 α_i 不尽相同, 它们对衡量姿势匹配的贡献度也存在差异, 所以不能直接使用角度值和 α_{si} 来表示标准姿势和实时姿势之间的差异程度。可以通过给每一个特征矢量间角度值 α_i 赋予不同的权重值来抵偿这种差异性, 各角度值的权重计算公式如下

$$\omega_i = \frac{1 - e^{-\frac{\alpha_i}{\alpha_{si}}}}{\sum_{i=0}^8 1 - e^{-\frac{\alpha_i}{\alpha_{si}}}} \quad (9)$$

根据式 (9) 可以得到一个权重值集合 $W = \{\omega_0, \omega_1, \omega_2, \dots, \omega_8\}$, 可知 $\sum_{i=0}^8 \omega_i = 1$ 。并且, 特征矢量间的角度值 α_i 越大, 它被赋予的权重值 ω_i 也就越大。根据各特征矢量间的角度值及其对应的权重值, 可以得到一个归一化的角度值

$$D = \sum_{i=0}^8 \alpha_i \cdot \omega_i \quad (10)$$

由式 (10) 得到的归一化角度值参数 D , 它唯一地表示了标准姿势和实时姿势之间的差异度。在此, 引入 L. Liu 等^[13] 提出的基于 Kinect 的静态人体姿势评分方法中的评分计算公式, 并加以改进使其满足本文方法要求, 再将 D 值带入姿势匹配识别准确度计算公式中, 如下

$$S = \begin{cases} f(\alpha_{max}) \cdot [(D_s - D) \cdot \frac{100 - S_s}{D_s} + S_s] & 0 \leq D \leq D_s \\ 0 & D > D_s \end{cases} \quad (11)$$

姿势匹配识别准确度 S 的取值范围为 $[0, 100]$, 标准姿势和实时姿势的匹配度越高, S 的值越大。在式 (11) 中, D_s 为预设的标准角度差阈值, D_s 值越小, 姿势匹配准确度的判别越严厉。为避免出现负分的情况, 当 $D > D_s$ 时, 姿势匹配准确度将直接判为 0。 S_s 为预设的基准匹配度参数, 可以通过参数 S_s , 根据实际需要将姿势匹配识别准确度 S 控制在合适的范围内。 $f(\alpha_{max})$ 为肢体偏移极限函数^[13], 表示一个姿势匹配准确度的惩罚因子, 具体公式如下

$$f(\alpha_{max}) = \begin{cases} 1 - \frac{0.3}{M^2} \cdot \alpha_{max}^2 & 0 \leq \alpha_{max} \leq \sqrt{10/3}M \\ 0 & \alpha_{max} > \sqrt{10/3}M \end{cases} \quad (12)$$

式中: α_{max} ——角度值集合 θ 中的最大值, M ——预设的肢

体偏移的最大角度阈值。由式 (12) 可知: $f(0) = 1$, 当实时姿势和标准姿势完全相同时, 函数 $f(\alpha_{max})$ 的值为 1, 即姿势匹配准确度不会受到惩罚; $f'(\alpha_{max}) < 0$, $f''(\alpha_{max}) < 0$, 即最大角度值 α_{max} 的值越大, 函数 $f(\alpha_{max})$ 的值下降的越快; $f(M) = 0.7$, 当 $\alpha_{max} = M$ 时, 姿势匹配准确度会受到惩罚下降 30%。 α_{max} 值继续增大时, 函数 $f(\alpha_{max})$ 的值也会快速地下降, 直至为 0。

2.4 算法分析和描述

在本文方法中, 设预先设定的标准姿势骨骼数据为 A_s , 当前捕获的实时姿势骨骼数据为 B_r , 姿势匹配识别准确度为 S , 欲处理的骨骼关节点个数为 n 。本文算法描述如下所示, 可验证系统一次执行过程中, 基本操作的执行次数为 $3n$, 即该算法的时间复杂度为 $O(n)$, 且空间复杂度为 $O(1)$ 。本文所提出的人体动作姿势识别算法较简单、计算较快速。

算法名称: 基于 Kinect 骨骼数据的人体动作姿势匹配识别算法

输入: A_s, B_r

输出: S

方法: 算法按照以下步骤执行

步骤 1 提取出姿势特征点, 并初始化;

步骤 2 读取姿势特征点的三维坐标信息;

步骤 3 根据式 (6) 计算肢体特征矢量, 得到姿势特征描述符 P ;

步骤 4 根据式 (7) 计算标准姿势和实时姿势的特征描述符中对应的各特征矢量间的角度值, 得到角度值集合 θ ;

步骤 5 根据式 (8) ~ 式 (10), 将集合 θ 中的各角度值归一化为一个统一的角度值参数 D ;

步骤 6 依据步骤 5 中得到的 D 值, 根据式 (11) 和式 (12), 计算姿势匹配识别的准确度;

步骤 7 给出姿势匹配识别准确度 S (百分制), 算法结束。

3 实验验证及分析

3.1 实验设计及环境

为了体现本文算法的有效性、鲁棒性和改进性, 分别进行了两组验证实验和两组对比实验。对比对象为 H. Jiang 等^[11] 提出的基于一致性最大覆盖的单帧图像姿态估计方法, 和 T. Y. Lin 等^[6] 在基于 Kinect 的康复训练系统中, 提出的基于 Kinect 骨架的姿势识别方法。其中, 为了表明本文方法比传统基于视觉的人体动作姿势识别方法具有更好的识别准确率, 将本文方法和 H. Jiang 等的方法进行了一组有关识别率的对比实验; 为了表明本文方法比 T. Y. Lin 等的方法具有更好的实时性, 在相同条件下与本文方法进行了一组关于实时性的对比实验。

本文实验的硬件平台为: Kinect for Windows 摄像机一台, 所使用彩色图像和深度图像的分辨率均为 640×480 ; 实验所用计算机为 PC 机, 主要配置包括: Intel(R) Core (TM) i5-2320, 3.30 GHz 主频四核处理器, 内存为 4 GB, 显卡为 NVIDIA GeForce GTX 560; 系统开发环境为: Windows 7 操作系统, Visual Studio 2010 和 Kinect SDK v1.7。

3.2 姿势识别验证实验

本文使用了 6 组实验来验证本文方法的有效性。经过多次实验及分析发现, 当各阈值参数 $D_s=60$, $S_s=80$ 和 $M=30$ 时, 姿势匹配识别实验的效果最佳。并且, 当姿势匹配识别准确度 $S \geq 80$ 时, 则可以判定标准姿势和实时姿势之间实现最近匹配识别; 若 $60 \leq S < 80$ 时, 则可以判定标准姿势和实时姿势之间为次近匹配识别; 否则视为两者匹配识别失败(下文其它实验中各阈值参数设置和匹配识别规则均与此相同)。实验结果如图 4 所示, 每幅图的左边为预先设置的标准动作姿势, 右边图像为当前捕获的实时动作姿势, 右下方为标准姿势和实时姿势的匹配识别准确度(百分制)。

从图 4 (a) ~ (f) 中可知, 本文方法通过比较标准姿势特征描述符和实时姿势特征描述符之间的差异性, 能够有效地实现对人体动作姿势的匹配计算和识别。其中, 图 4 (e) ~ (f) 实验结果由于实时姿势和标准姿势在某一肢体上存在较为明显的差异(分别为右手小臂和左手小臂), 对应的特征矢量间角度值 α_i 受到其相应权重值 ω_i 和肢体偏移极限函数 $f(\alpha_{max})$ 的双重惩罚, 使得最终的姿势匹配识别准确度 S 产生较为明显的下降。

3.3 鲁棒性分析实验

本实验中, 使用与图 4 (a) ~ (d) 相同的标准姿势模板, 在暗光环境下采集实时姿势的骨骼关节点信息进行人体动作姿势的匹配识别。如图 5 所示, 姿势匹配识别结果基本没有受到影响, 本文方法仍能够准确地计算和判别标准姿势和实时姿势是否匹配识别。该实验结果表明, 本文方法的姿势匹配识别结果不受到光照变化的影响, 比传统的基于视觉的人体动作姿势识别方法具有更强的鲁棒性。

3.4 识别率分析实验

如图 6 所示, 将本文方法与 H. Jiang 等^[11]提出的基于一致性最大覆盖的单帧图像姿态估计方法在识别率方面进行了一组对比实验。其中, 图 6 (a) 为 H. Jiang 等的识别方法的实验结果, 图 6 (b) 为本文方法的识别实验结果, 且使用了更为复杂的背景。

从图 6 (a) 中选取 6 种动作姿势与本文方法进行识别率对比实验, 且每一种动作姿势进行多次实验, 并分别给出了两种方法的最低识别率、最高识别率及平均识别率。如表 1 所示, 实验结果表明本文方法具有更高的识别率, 识别准确性更强。



图 4 本文人体动作姿势识别方法验证实验

表 1 两种方法的识别率对比结果

方法类型	最低识别率/%	最高识别率/%	平均识别率/%
H. Jiang 等 ^[11] 的识别方法	87	93	89.3
本文方法	90.3	96.7	93.6

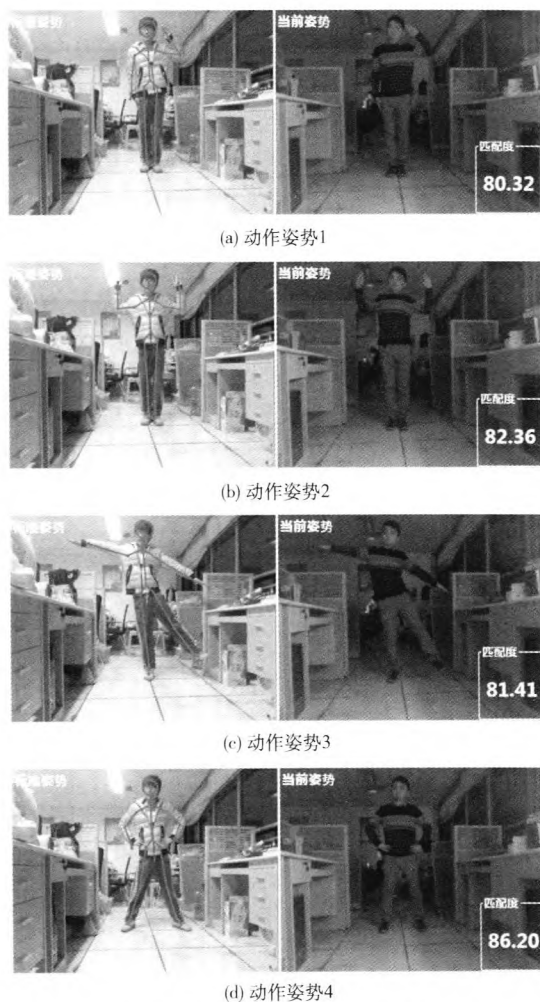


图5 光照变暗条件下的实验效果

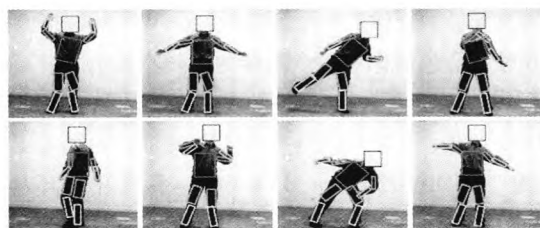
3.5 实时性分析实验

将本文方法与 T. Y. Lin 等^[6]提出的基于 Kinect 骨架的姿势识别方法在实时性方面进行一组对比实验,并使用如图 7 所示的 8 种动作姿势进行对比实验。实验过程中,在相同的实验条件下,分别用该方法和本文方法对图 7 中的每一种动作姿势进行多次实验,并计算出两种方法对每一种动作姿势进行识别的平均时间,实时性对比实验结果见表 2。

由表 2 实验结果数据可知,与 T. Y. Lin 等的识别方法相比,本文方法对每种动作姿势的平均识别耗时较短,前者对每种动作姿势进行识别的平均时间为 41.6 ms,而本文方法的平均识别时间为 32.4 ms。实验结果表明,两种方法的识别率较接近,而本文方法具有更好的实时性。

4 结束语

本文提出了一种基于 Kinect 骨骼数据的人体动作姿势识别方法。该方法通过从 Kinect 获取的骨骼数据中提取出的姿势特征点,并计算出姿势特征矢量,得到由特征矢量集合表示的姿势特征描述符,然后将当前的实时姿势与预

(a) H.Jiang 等^[11]的识别方法实验结果

(b) 本文方法识别实验结果

图6 两种方法的识别率对比实验

设的标准姿势特征描述符进行匹配比较,并实时地反馈两者姿势的匹配度,从而达到人体动作姿势识别的目的。实验验证和对比结果表明,本文方法能有效地进行人体动作姿势的匹配识别,识别准确率高、实时性较好和鲁棒性较强,且具有良好的可扩展性。但本文方法仍然在以下方面存在不足:当人体动作姿势中骨骼关节点发生自遮挡时,该方法的识别率会有较明显的下降;该方法只是针对静态的人体动作姿势进行匹配识别,下一步工作将在此方

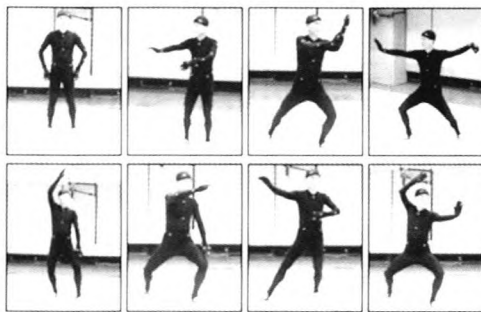


图 7 两种方法实时性对比实验的动作姿势

表 2 两种方法识别实时性对比结果

姿势 类型	T. Y. Lin 等 ^[6] 的识别方法		本文方法	
	平均识别 时间/ms	识别率/%	平均识别 时间/ms	识别率/%
姿势 1	38	95.3	29	95.8
姿势 2	42	92.4	31	91.3
姿势 3	45	90.2	37	89.3
姿势 4	39	93.1	30	94.5
姿势 5	40	92.7	31	92.2
姿势 6	45	91.4	38	90.1
姿势 7	43	92.3	32	93.4
姿势 8	41	92.7	31	93.4

法的基础上进行扩展和改进, 实现动态的人体动作行为过程的识别。

参考文献:

[1] HUANG Guofan, LI Ya. A survey of human action and pose recognition [J]. Computer Knowledge and Technology, 2013 (1): 133-135 (in Chinese). [黄国范, 李亚. 人体动作姿态识别综述 [J]. 电脑知识与技术, 2013 (1): 133-135.]

[2] HU Qiong, QIN Lei, HUANG Qingming. A summary of human action recognition based on visual [J]. Chinese Journal of Computers, 2013, 36 (7): 1-14 (in Chinese). [胡琼, 秦磊, 黄庆明. 基于视觉的人体动作识别综述 [J]. 计算机学报, 2013, 36 (7): 1-14.]

[3] Biswas K K, Basu S K. Gesture recognition using Microsoft

Kinect [C] //5th International Conference on Automation, Robotics and Applications. IEEE, 2011: 100-103.

[4] Shotton J, Sharp T, 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.

[5] Miranda L, Vieira T, Martínez D, et al. Online gesture recognition from pose kernel learning and decision forests [J]. Pattern Recognition Letters, 2014, 39: 65-73.

[6] Lin T Y, Hsieh C H, Lee J D. A Kinect-based system for physical rehabilitation: Utilizing tai chi exercises to improve movement disorders in patients with balance ability [C] //Proceedings of the 7th Asia Modelling Symposium. IEEE Computer Society, 2013: 149-153.

[7] Le T L, Nguyen M Q, Nguyen T T M. Human posture recognition using human skeleton provided by Kinect [C] //International Conference on Computing, Management and Telecommunications. IEEE, 2013: 340-345.

[8] TANG Zhibo, ZHANG Zhili, LIANG Feng, et al. Hand recognition in virtual maintenance training [J]. Computer Engineering and Design, 2014, 35 (12): 4278-4283 (in Chinese). [汤志波, 张志利, 梁丰, 等. 虚拟维修训练中的手势识别 [J]. 计算机工程与设计, 2014, 35 (12): 4278-4283.]

[9] LIN Peng, ZHANG Chao, LI Zhuliang, et al. Human body part identification based on depth images [J]. Computer Engineering, 2012, 38 (16): 185-188 (in Chinese). [林鹏, 张超, 李竹良, 等. 基于深度图像学习的人体部位识别 [J]. 计算机工程, 2012, 38 (16): 185-188.]

[10] Raptis M, Kirovski D, Hoppe H. Real-time classification of dance gestures from skeleton animation [C] //Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation. ACM, 2011: 147-156.

[11] Jiang H. Human pose estimation using consistent max covering [J]. Pattern Analysis and Machine Intelligence, 2011, 33 (9): 1911-1918.

[12] Liu T T, Hsieh C T, Chung R C, et al. Physical rehabilitation assistant system based on Kinect [J]. Applied Mechanics and Materials, 2013, 284: 1686-1690.

[13] Liu L, Wu X, Wu L, et al. Static human gesture grading based on Kinect [C] //5th International Congress on Image and Signal Processing. IEEE, 2012: 1390-1393.