

文章编号: 2095-2163(2020)04-0035-05

中图分类号: TP391

文献标志码: A

基于 Kinect2.0 的健身动作训练系统

刘长喜, 付少锋

(西安电子科技大学 计算机科学与技术学院, 西安 710126)

摘要: 为减少在健身训练指导中存在的主观因素影响, 促进健身运动的标准化、规范化, 提出通过 Kinect 获取人体主要关节的运动, 进一步转化为表征运动的线条图。利用 Harris 角点检测得到特征点, 基于 MOPS 描述符进行特征匹配, 采用等特征点比例分段法进行分析, 计算各项指标, 为学员提供有效地建议。实验结果表明, 对于一系列连续动作, 对比结果与人为对比相比在细节方面更加优秀。

关键词: Kinect; 连续动作对比; 训练辅助; 特征匹配; 关节姿态

Fitness exercise training system based on Kinect v2.0

LIU Changxi, FU Shaofeng

(School of Computer Science and Technology, XiDian University, Xian 710126, China)

[Abstract] In order to reduce the influence of subjective factors in fitness training guidance and promote the standardization and standardization of fitness exercises, it is proposed to use Kinect to obtain the movements of the main joints of the human body, and further convert them into line diagrams that represent sports. Feature points were obtained using Harris corner detection, feature matching was performed based on MOPS descriptors, and equal feature point proportion segmentation methods were used to analyze and calculate various indicators to provide effective suggestions for students. The experimental results show that for a series of continuous actions, the comparison results are better in detail than the artificial comparison.

[Key words] Kinect; Continuous motion comparison; training assistance; feature matching; joint pose

0 引言

随着国家经济的发展, 人民生活水平的提高, 政策的鼓励^[1], 健身在人们生活中的地位越来越重要。然而, 传统的健身房或者家用健身产品都存在一些缺点。例如, 教练的指导取决于主观因素, 动作校正不够具体直观, 健身器材无法贴合学员需要并且很难独立学习。由此出现了许多依托于互联网+的智能健身设施^[2]。如: 英派斯, 舒华等, 但随之产生了许多问题。

侵入式设备对人体感觉不友好, 而且设备价格昂贵^[3]; 非侵入设备, 如光学运动式捕捉^[4]对人体运动无影响, 但精度低。针对这些问题, 本文使用 Microsoft 的 Kinect2.0 体感器, 结合特征匹配算法实现了一套健身运动训练系统。Kinect 提供了一系列可供开发者使用的 API, 可以方便的获取人体的深度数据和关节姿态信息, 在匹配完成后, 得到一组评价数据, 用来指导学员的动作。本文将 Kinect 体感器及模式识别技术与传统的健身学习模式相结合,

为健身行业注入了新的活力。

1 Kinect2.0 简介

1.1 kinect 深度图像

Kinect2.0, 即微软第二代 Kinect for Windows, 发布于 2014 年。Kinect2.0 有 3 个镜头, 分别为彩色摄像头、深度(红外)摄像头和红外线发射器^[5], 如图 1 所示。

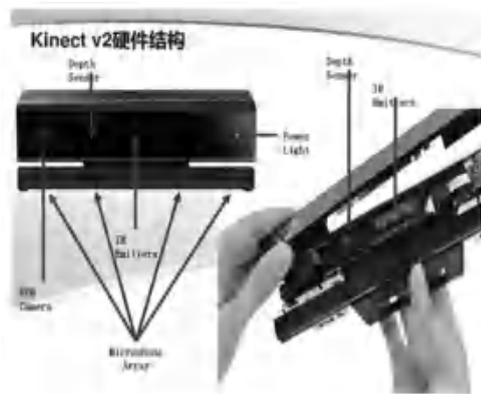


图 1 Kinect 外表及内部结构

Fig. 1 Kinect appearance and internal structure

基金项目: 国家大学生创新创业训练计划(201910701098)。

作者简介: 刘长喜(1999-), 男, 本科生, 主要研究方向: 计算机视觉、模式识别; 付少锋(1975-), 男, 硕士, 副教授, 主要研究方向: 信息安全、嵌入式系统。

通讯作者: 刘长喜 Email: liucxi@foxmail.com

收稿日期: 2020-01-06

在获取深度信息方面, Kinect 使用光编码技术。编码具有高度的随机性和唯一性。深度摄像头可以接收这些信息, 进而读取物体的深度信息^[6]。

1.2 基于深度图像的动作捕捉

Kinect2.0 定义了 25 个人体关节, 如 SpineBase、HandLeft、HandRight 等。人体部位明确定义了左右, 使分类算法具有区分人体左右侧的能力。人体部位的定义可以根据实际场景做出调整。

Kinect 定义了描述身体部位的特征, 其在不同部位有不同的响应。单个特征可以提供像素所属于身体部位的微弱信号, 在随机森林中组合后, 就可以准确区分所有人体部位^[7], 即确定每个像素的部位信息。将所有可能的像素汇聚起来, 就可以形成人体骨架关节推测结果。

Kinect 定义的关节名称及序号如图 2 所示。使用 Kinect 可获得骨骼关节的旋转信息。

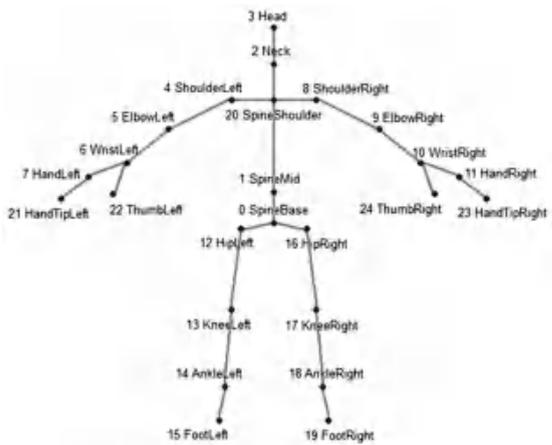


图 2 kinect 人体关节示意图

Fig. 2 Kinect human joint diagram

2 动作特征匹配

2.1 数据预处理

首先, 利用 Kinect 捕捉人体运动, 获得一系列关节旋转数据, 这些数据在 Kinect 中以四元组的形式存储。有关提取特征的方法较多, 如: 考虑关节自由度的角度特征^[8]、关注距离和角度的组合特征^[9]等。为方便后续处理, 本文将其转为以欧拉角形式存储。

在此, 本文舍弃了一些对动作匹配没有影响的关节, 保留的关节如表 1 所示。

2.1.1 四元数与欧拉角

在图形学中, 四元数和欧拉角是最常用的旋转表示方式, 与矩阵相比具有节省存储空间和方便插值的优点^[10]。

表 1 保留关节编号及名称

Tab.1 Joint number and name

原编号	新编号	关节名称
0	0	SpineBase
4	1	ShoulderLeft
5	2	ElbowLeft
6	3	WristLeft
7	4	HandLeft
8	5	ShoulderRight
9	6	ElbowRight
10	7	WristRight
11	8	HandRight
12	9	HipLeft
13	10	KneeLeft
14	11	AnkleLeft
16	12	HipRight
17	13	KneeRight
18	14	AnkleRight

四元数: 在三维建模中, 四元数占有非常重要的地位。因它不仅有着明确的几何意义, 且表征方式简单, 只用到了四个数。Kinect 中关节的位置即采用此方式。其定义如下:

$$\begin{cases} q = [w \ x \ y \ z]^T, \\ |q|^2 = w^2 + x^2 + y^2 + z^2. \end{cases} \quad (1)$$

欧拉角: 使用绕三维空间中相关垂直的坐标轴的三个旋转组成的序列, 来表示一个角位移, 使用 Roll、Yaw、Pitch 分别表示物体绕 x 、 y 、 z 的旋转角度, 记为 a 、 b 、 c 。

2.1.2 四元数到欧拉角的转换

四元数转换为欧拉角的公式如下:

$$\begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} \hat{e} \\ \hat{e} \\ \hat{e} \end{pmatrix} \begin{pmatrix} \arctan2(2(wx + yz), 1 - 2(x^2 + y^2)) \\ \arcsin(2(wy + zx)) \\ \arctan2(2(wz + xy), 1 - 2(y^2 + z^2)) \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{u} \\ \hat{u} \end{pmatrix} \quad (2)$$

对采集到的数据进行处理, 去突变、去边缘得三个通道角度变化。图 3(a)~3(c) 为 HandLeft 三通道路线图:

2.2 特征检测与特征描述

由此, 动作序列的对比转化为线条图特征的匹配。相比直接对比图像, 可以大大减少计算量。本文采用的 Harris 特征检测和 MOPS (Multi-scale oriented patches) 特征匹配方法, 与其它特征点检测和匹配方法, 如 DoG^[11]、OGB^[12]、SIFT^[13] 等相比, 对于线条图来说, 在保证准确率的基础上有高效的运算。

2.2.1 Harris 角点检测

采用 Harris 寻找不变的局部特征。Harris 角点检测寻找到的特征点在局部范围内具有较大的灰度变化, 同时具有平移、旋转不变形^[14]。Harris 算子定义为:

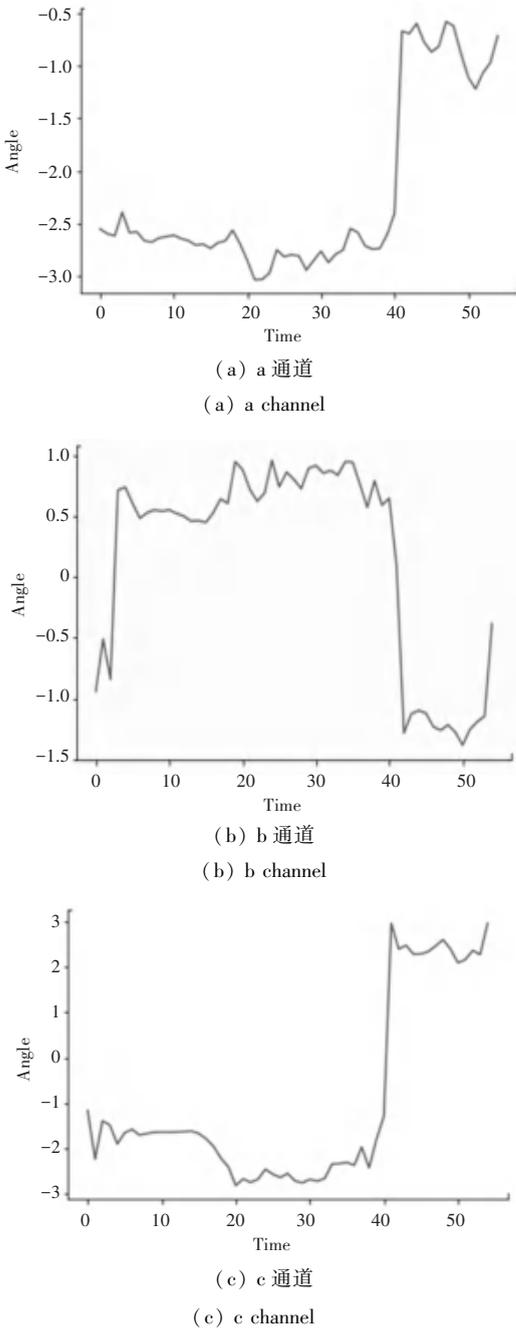


图 3 HandLeft 三通道线条图

Fig. 3 HandLeft three-channel line drawing

$$H = \begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix} \quad (3)$$

$$C = \det(H) - k (\text{trace}(H))^2$$

其中, I_x 、 I_y 为图片沿 x 、 y 方向的梯度, k 为常数 (通常取值为 0.04 ~ 0.06), 当 C 的值大于阈值时, 认为该点为特征点^[15]。

角点检测核心步骤包括^[16]:

- (1) 计算图像 x 方向与 y 方向导数;
- (2) 计算导数的平方;

- (3) 高斯滤波器平滑;
- (4) 寻找角点;
- (5) 非最大抑制^[16]。

2.2.2 MOPS 描述符

MOPS 是简单且高效的特征描述符, 它采集特征点周围的像素特征, 存储为 64 维向量。计算 MOPS 描述符的核心步骤包括^[17]:

- (1) 选取所检出特征周围 40×40 的方形窗口;
- (2) 由 2.2.1 可得特征点的梯度方向, 将该窗口旋转到水平;
- (3) 在以特征点为中心的 8×8 方形窗口内采样;
- (4) 规格化^[17]。

2.3 特征匹配

计算两特征向量间的欧式距离, 计算最优与次优匹配的比值, 通过指定合适的阈值过滤掉错误匹配。

3 动作评判指标

考虑到学员动作与标准动作频率不一, 在此定义等特征点比例分段方法^[18]。

等特征点比例分段, 即两组动作对应分段特征点, 占总特征点数量的比例相同。具体步骤为: 将标准动作序列等间隔分为 N 段。动作越复杂, 分段数量越高。学员动作分段需先计算标准动作段特征点占总特征点的比例, 再在学员动作中选取相同比例的特征点作为对应段。

完成分段后, 就可以对动作进行全面的对比。对比指标如下:

- (1) 动作总体匹配比例

标准动作与学员动作匹配的特征点数量 Num_0 / 标准动作特征点的总数量 Sum_0 。

- (2) 局部动作匹配比例

标准动作中第 i 段与学员动作第 i 段匹配的特征点数量 Num_i / 标准动作中第 i 段特征点的总数量 Sum_i 。

由此, 可得动作评分:

$$Score = 0.6 \sum_{i=1}^N \frac{Num_i}{Sum_i} + 0.4 \frac{Num_0}{Sum_0} \quad (4)$$

(3) 动作复杂程度: 检测到的特征点数量在一定程度上可以体现动作的复杂程度。

(4) 动作提前与延后: 通过特征点匹配图像, 可以观察到学员动作相对于标准动作的提前与延后, 从而对学员动作进行指导。

4 实验结果分析

本实验测试总流程如图 4 所示。

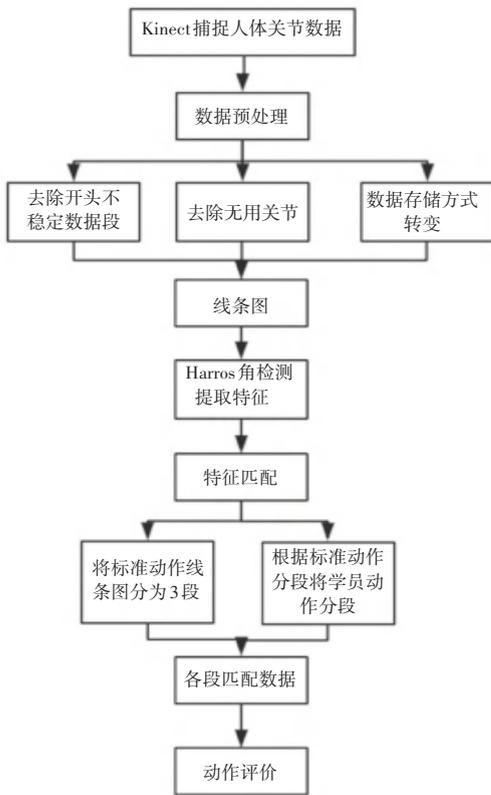


图4 系统总流程

Fig. 4 System overall process

本文采集了两组动作,第一组作为标准动作,第二组作为学员动作,在捕捉过程中截取连续两帧图像,如图5(a)、5(b)所示。特征检测结果如图6(a)、6(b)所示。综合3个通道的匹配关系,按照特征点比例分段方法,得到分段结果如表2所示。



(a) 第一帧

(a) First frame

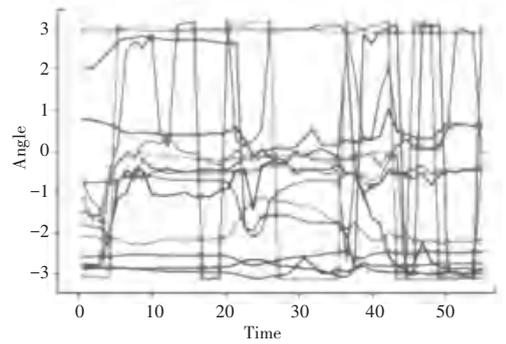


(b) 第二帧

(b) Second frame

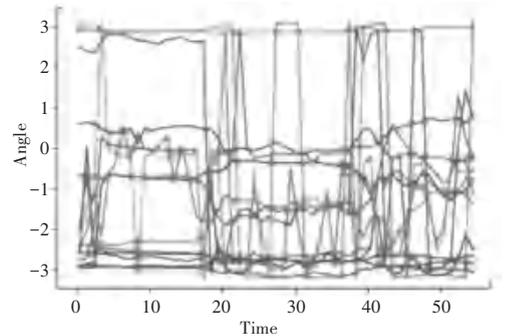
图5 动作捕捉示意图

Fig. 5 Schematic of motion capture



(a) 标准动作

(a) Standard motion



(b) 学员动作

(b) Student motion

图6 特征检测结果

Fig. 6 Feature detection results

表2 分段结果表

Tab. 2 Segmentation table

组别	第1段	第2段	第3段	第4段	第5段
标准动作	46	34	62	102	113
学员动作	32	44	58	86	82

在本次实验中,选取 $N = 5$ 。各段匹配结果如表3所示。

其中,每一行为标准动作的段,每一列为学员动作的段。如:学员动作第一段与标准动作第二段匹配数据位于第一列第二行,有5个特征点匹配。由此可得:

表3 各段匹配结果

Tab. 3 Matching results

N	第1段	第2段	第3段	第4段	第5段
第1段	6	1	0	2	1
第2段	5	4	2	1	0
第3段	4	2	17	8	7
第4段	6	0	4	18	7
第5段	1	0	2	12	10

(1) 动作总体匹配比例。 $Num_0/Sum_0 = 118/357 = 33.05\%$,即学员动作与标准动作的整体匹配程度为33.05%。

(2) 局部动作匹配比例。由指标可知,各段局

部匹配比例依次为:13.04%、11.76%、27.41%、17.64%、8.84%。总得分为:0.258 852。

(3)动作复杂程度。由表2可知,标准动作有357个特征点,而学员动作有302个特征点,即标准动作比学员动作要复杂、多变。

(4)动作提前与延后。由式(2)可知,学员动作与标准动作匹配最好的为第三段,其次为第四段,最差的为第五段。可见该学员对复杂动作的掌握不到位,而较简单动作,如第一段、第二段有没有做好,所以得分较低。观察表3(5,4)和(4,5)可知,学员动作第四段有12个特征落在了第五段,即第四段动作落后于标准动作,同理,第五段动作稍提前于标准动作。

5 结束语

本系统将 Kinect 和特征匹配相结合,实现了对连续动作的对比和评价。首先利用 Kinect 获取骨骼的旋转信息,再将其转换为欧拉角的形式存储,在特征匹配方面,选用 Harris 与 MOPS 相结合,准确而又高效。在后面的实验中,对两段动作进行了比较,看到了许多凭肉眼看不到的细节,为动作的比较与评价提供了科学客观的数据。

参考文献

- [1] 李杨,姜藤野.我国全民健身运动发展现状和对策研究[J].中国传媒科技,2012(02):229.
- [2] 陈兴兴.基于“互联网+”的智能家用健身器材设计研究[D].济南:齐鲁工业大学,2019.
- [3] 陈晓,宋晓梅,张意华.可穿戴运动捕捉系统[J].国外电子测量技术,2017,36(10):60-63.

- [4] 谭菁华,晏怡果.光学式三维运动捕捉在运动系统相关研究中的应用[J].海南医学,2018,29(14):2029-2032.
- [5] 陈昌伟.基于 Kinect 的人体动作比对分析及生物力学分析[D].天津:天津大学,2014.
- [6] 薛智宏,张利英,程振华,等.基于 Kinect 的原地太极拳辅助训练系统[J].河北科技大学学报,2017,38(02):183-189.
- [7] SHOTTON, J. Real-time human pose recognition in parts from single depth images [C]//Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011.
- [8] 李红波,李双生,孙船源.基于 Kinect 骨骼数据的人体动作姿势识别方法[J].计算机工程与设计,2016,37(04):969-975.
- [9] 郑轩宇,史畅,崔文成.基于 Kinect 的运动训练辅助系统[J].计算机与现代化,2019(08):12-16,56.
- [10] 刘正存.面向大众体育运动示教的三维人体动作捕捉与分析[D].天津:天津大学,2013.
- [11] 邸男,李桂菊,陈春宁,等.结合归一化差分高斯特征的图像匹配技术研究[J].电子测量与仪器学报,2014,28(06):585-590.天津大学,2014.
- [12] 姚晋晋,张鹏超,王永鑫,等.基于局部一致性的特征匹配算法[J].数字技术与应用,2019,37(07):128-130.
- [13] 陈凯扬,罗志灶,王建兴.基于 Kinect 三维重构的特征点提取改进[J].计算机与现代化,2019(11):34-37.
- [14] 陶冶江,黄华.基于改进的 Harris 和 SIFT 算法的眼底图像拼接[J].计算机工程与设计,2012,33(09):3507-3511.
- [15] 侯建辉,林意.自适应的 Harris 棋盘格角点检测算法[J].计算机工程与设计,2009,30(20):4741-4743,4801.
- [16] HARRIS C, STEPHENS M. A Combined Corner and Edge Detector[C]//Proceedings of the 4th Alvey Vision Conference, 1988.
- [17] Brown, M. Multi-image matching using multi-scale oriented patches[C]//Computer Vision and Pattern Recognition(CVPR) 2005. IEEE Computer Society Conference on, 2005.
- [18] 任洪林.基于 Kinect 的个性化人体三维动作重现与动作细节比对研究[D].天津大学,2014.

(上接第34页)

着搜索层数的加深,搜索的局面也会成指数级别的增加。然而对于一些局面没有必要再去对它们进行搜索。本文选择 Alpha Beta 剪枝算法,从而增加搜索的层数,提高博弈模型的强度。

虽然本文的设计达到一定的使用效果,但还有待进一步完善。如评估函数的设计还较为朴素,静态评估考虑的棋盘因素有限。在设计博弈模型时还可采用深度学习和强化学习相结合的方法;可采用 Alpha Zero 的方法来对博弈模型进行自博弈训练,通过大量的自对弈让模型通过自我学习的方式提升棋力。

参考文献

- [1] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587):484.

- [2] 杨周凤.国际跳棋完备信息博弈关键技术研究及系统设计[D].沈阳:沈阳航空航天大学,2018.
- [2] 西洋跳棋规则.《互联网文档资源(<http://wenku.baidu.com>)》,2017
- [3] 邓丁朋,孟坤.在线棋牌用户智能化等级评定算法研究[J].数码世界,2016(1):75-76,77.
- [4] 张利群.机器博弈中的数据结构与基本方法[M].辽宁科学技术出版社,2014:1-6.
- [5] 吴岳.一种高性能西洋跳棋引擎设计[J].电脑编程技巧与维护,2014,32(2):63-68.
- [6] 郑昌松,贾丽娟,权贺.等.基于西洋跳棋的博弈程序研究[J].哈尔滨理工大学学报,2016,21(3):24-28.
- [7] 程宇.人机博弈中评估函数的研究[J].福建电脑,2011,27(9):1-2.
- [8] 徐心和,王骄.中国象棋计算机博弈关键技术分析[J].小型微型计算机系统,2006,27(06):961-969.
- [10] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676):354.