

文章编号 1004-924X(2015)增-0703-10

基于 Kinect 数据主成分分析的人体动作识别

刘志强, 尹建芹*, 张 玲, 魏 军, 刘小丽

(济南大学 信息科学与工程学院 山东省网络环境智能计算技术重点实验室, 山东 济南 250022)

摘要: 为了提高家庭环境下人体动作识别的效率和精度, 提出并实现了基于 Kinect 数据主成分分析的动作识别方法。首先, 通过 Kinect 采集人体动作特征描述的时间序列数据, 并构造人体姿态描述向量; 然后, 运用主成分分析方法分析不同时间点的特征值的差异, 获得重构的特征值, 使得不同类型的动作之间有明显区别。同时减少了冗余和噪声, 有利于动作的判断和识别。最后, 依据重构的特征和最近邻原则, 进行动作的识别分类。实验结果表明, 该方法对简单的人体动作识别精度可达 80% 以上, 单个样本识别时间分别是 1.67 ms 和 3.93 ms, 基本满足对人体动作识别的精度、抗干扰能力和实时性等要求。

关键词: 机器人; Kinect 数据; 动作识别; 主成分分析(PCA)

中图分类号: TP391 **文献标识码:** A **doi:** 10.3788/OPE.20152313.0703

Human action recognition based on Kinect data principal component analysis

LIU Zhi-qiang, YIN Jian-qin*, ZHANG Ling, WEI Jun, LIU Xiao-li

(Shandong Provincial Key Laboratory of Network Based Intelligent Computing,
School of Information Science and Engineering, University of Jinan, Jinan 250022, China)

* Corresponding author, E-mail: ise_yinj@ujn.edu.cn

Abstract: To improve the efficiency and accuracy of human action recognition in the home environment, a method of action recognition based on Principal Component Analysis(PCA) of Kinect data was proposed and realized. Firstly, the time series data of human action feature description was collected by Kinect, and human pose description vector was constructed. Then, the PCA method was used to analyze the difference between the feature values at different time points and to obtain the reconstructed eigenvalues, by which the distinction between different types of actions can be more obvious. Moreover, the action description feature of the redundancy was filtered to reduce the redundancy and noise, which was conducive to the judgment and recognition for the human action. Finally, the recognition and classification of the action were performed based on the features of the reconstruction and the nearest neighbor principle. The experimental results show that the accuracy of this method is more than 80% for the simple human motion recognition. The identification time of a single sample is

收稿日期: 2015-06-05; 修订日期: 2015-06-21.

基金项目: 国家自然科学基金资助项目(No. 61203341, No. 61375084)

1.67 ms and 3.93 ms, respectively. It can satisfy the human action recognition requirements for higher precision, strong anti-interference ability and real-time recognition.

Key words: robot; Kinect data; action recognition; Principal Component Analysis(PCA)

1 引言

在服务机器人领域,主动化和智能化是当前家庭服务发展的主要方向。为了能让机器人的服务更加主动和智能,满足人类的不同需求,实现人与机器顺畅自如的交流与互动,首先需要让机器识别人体动作,通过判别人在干什么,需要什么,以达到读懂人的行为信息的目的。另外,人体动作识别在高级人机交互、智能监控、体感游戏控制等方面,也有广泛的应用。

目前,在人体行为识别领域里还面临着许多困难和挑战:首先,传统的基于普通彩色图像或视频的研究方法,可获得的作为判别依据的有效数据信息非常有限,且容易受到光照、遮挡等噪声的干扰,很难取得较理想的动作识别效果;其次,当待识别动作相对于摄像机发生位置改变时,难以提取平移、旋转及缩放不变特征,不同的动作就容易混淆,影响识别的精度;最后,对于一些较相似的动作,直接依据采集的动作描述特征数据很难区分判别。

针对上述问题,本文使用 Kinect 采集与人体关节相关的动作描述特征数据。一方面,由于 Kinect 不受光照以及复杂背景的影响,具有较好的抗干扰能力,使本文算法对于光照变化等复杂背景具有良好的不变性;另一方面,采用人体结构向量构造的关节角度及模比值作为描述动作的特征,具有平移旋转缩放不变性^[1];最后,针对相似动作难区分的问题,采用 PCA (Principal Component Analysis) 对动作描述特征进行过滤、重组,在新的特征维度上,提高不同类别动作的差异,一定程度上提高了相似动作区分判别的精度。

目前使用 Kinect 进行人体行为识别的研究

越来越流行,这主要得益于它所能提供的深度信息。行为识别领域里的两个关键问题是动作特征描述和分类器的设计。Lv 等人^[2]采用类似动画中虚拟的关键姿势建立动作模型,但是关键姿势的数量不容易确定,而且计算复杂。Raptis 等人^[3]采用 16 个主要关节的运动作为特征,识别人的跳舞的姿态。Zhao 等人^[4]采用质心半径模型作为每一帧姿势的形状描述。Yun 等人^[5]利用所有关节对之间的距离作为特征,并且通过多实例学习进行特征选择。Ellis 等人^[6]提出了基于姿势序列中的单个典型姿势识别动作的方法。Lin 等人^[7]通过计算分析人体运动过程中不同部位之间的速度关系进行人体行为识别。Sun 等人^[8]通过比较相邻两帧之间成对 SIFT 特征提取轨迹特征,进行人体行为识别。Dollar 等人^[9]提出了立方体方法,将每个立方体包含局部邻域的编码信息作为动作描述特征。Xia 等人^[10]采用基于三维关节位置的直方图描述人体姿态。Li 等人^[11]提出基于三维关节词包模型进行动作识别。还有一些基于稀疏表示动作识别方法^[12-13]。但这些研究很少考虑不同人之间体态差异所导致的同类动作特征描述聚集性变差。Cohen 等人^[14]构建人体三维轮廓图形并采用 SVM 技术进行分类。行为识别算法则主要有模板匹配法和状态空间法两类^[1]。基于状态空间生成模型的算法中比较有代表性的是隐马尔可夫模型 (Hidden Markov Models, HMM) 和条件随机场 Xiong 和 Liu 等人^[15]采用基于人体轮廓的马尔可夫模型进行简单的动作识别。Han 等人^[16]采用基于人体三维关节的条件随机场进行人体行为识别。状态空间法的优点是比较稳定,但是计算过程复杂,如果使用的训练数据过少,容易出现过拟合问题。动态时间规整 (Dynamic Time Warping, DTW) 方法是比较常用的模板匹配方法,其

是处理时间序列信号较为经典的一种算法,文献[17-19]都采用 DTW 算法进行人体动作识别。它是将测试样本每个时刻的特征值与已知类别的标准模板任意时刻的特征值进行匹配,要求匹配距离最小,但是由于顺序搜索要求逐点扫描,计算量较大,而且还需要判断奇异点,这些因素都会影响运算效率。

针对 DTW 算法计算量大的问题,本文用 PCA 对时间帧序列数据进行降维,从而减少了计算量,提高了运算效率。

2 基于 Kinect 的人体动作特征描述

文献[1]从对人体姿态的认知上入手,利用 Kinect 采集关节点数据,借鉴人的认知特点进行人体日常行为识别研究,重点研究了小范围内的人体行为识别;针对人体结构特点,利用关节点数据构造人体结构向量作为 3 维空间下的人体模型,通过计算人体结构向量间的向量角和向量模比值提取行为表示特征量,进行人体动作特征的描述,选择结构向量间的夹角有效消除不同人或不同位置的人体结构向量间的差异,对于角度值不能描述的某些行为的细节部分,采用了 4 组结构向量间的模比值作为辅助信息刻画角度信息十分相似的动作。实验说明了上述动作描述特征向量且具有旋转和缩放不变性,为行为分类器的设计提供了良好的基础。

3 基于主成分分析的动作特征提取

人体行为动作的原始特征向量空间包含大量的姿态属性,具有高维性和时序性的特点,但并不是所有时间点的姿态属性对分类决策都有贡献,因此,提取有利于动作分类的时间点的姿态属性,舍弃冗余的属性,有利于提高分类结果的准确性。

3.1 动作主成分的含义

原始动作数据降维之前,在不同的时间维度上的特征值具有时序相关性,即随着动作的执行,

同一种特征值随时间的变化,形成一条连续的曲线。对不同类型的动作而言,在有些时间段特征的幅值及其变化趋势是非常接近的,而有些时间段,却存在较大差异,这对于不同类型的动作分析来说,存在较大冗余成分,影响决策的效率。通过 PCA 去除冗余,保留差异较大的维度上的特征数据作为决策判别的依据。

一种动作在新的维度上的特征值,是对原来不同时间维度上的特征值分别加权求和得到,表示此类动作在新的维度上的分量,可以证明不同类型的动作样本,在顺序排列的不同的几个主成分维度上,类间方差较大。

3.2 动作主成分的提取

3.2.1 基于 K-均值的代表性训练样本选择

由于对于同一种动作而言,不同的执行对象之间存在差异,即不同的人做同一种动作,会因为人的体态差异,或执行动作的速度快慢有差别,而导致同类个体之间的差异性。通过 K 均值聚类的方法,首先在同一类的动作中分别聚出 K 个子类,其类心作为同类动作中比较有代表性的模板,再用这些模板数据构建用于 PCA 的原始矩阵,即把这 K 个子类当做不同的类来处理。

在得到代表性训练数据集的基础上,基于该数据集提取动作主成分,而主成分提取的关键在于协方差矩阵的构造,本文采用了两种协方差矩阵构造方法进行主成分分析。

3.2.2 串特征构造协方差矩阵

首先将连续帧中的所有特征首尾相接,由此得到的协方差矩阵为基础进行动作主成分提取。假设动作的类别个数为 M ,原始的动作描述特征维数为 N ,每一维度的特征的帧数为 T 。

如图 1 所示,一行表示一类动作样本,它是将所有特征值按时间序列串起来排开,构成的一个高维度的向量,定义其为串特征,记为 $a(i)$ (i 是动作模板标号),同一种颜色代表相同类别的动作,不同的颜色代表不同的动作,构造这样的原始数据的矩阵 A ,其大小为 $K * M * T * N$ 。

对原始数据矩阵 A 进行预处理(使其每一列

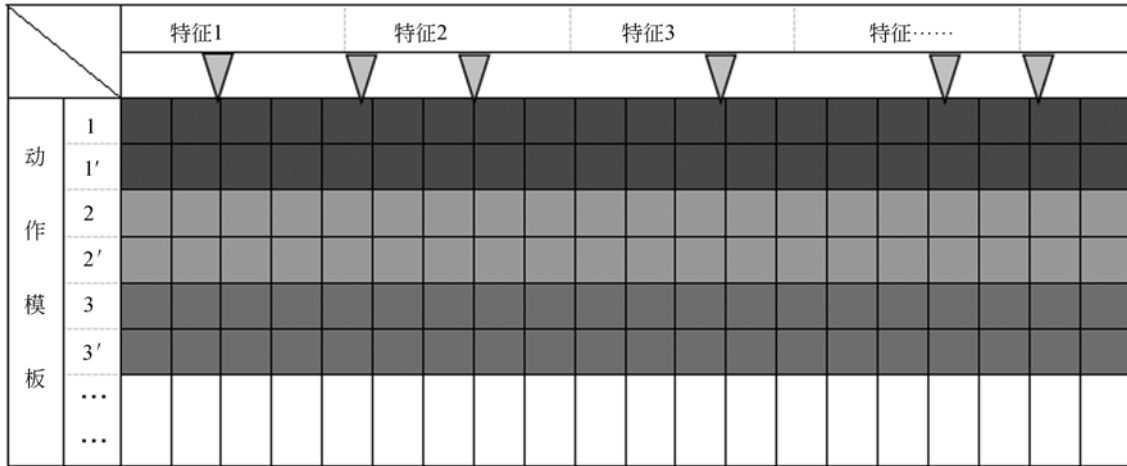


图 1 串特征原始模板构造

Fig. 1 Original template constructing of series feature

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}(1 \times 1) \\ \vdots \\ \mathbf{a}(K \times 1) \\ \vdots \\ \mathbf{a}(K \times M) \end{bmatrix} = \begin{bmatrix} a_{1 \times 1 \times 1} \cdots a_{1 \times 1 \times T} & a_{1 \times 2 \times 1} \cdots a_{1 \times 2 \times T} & \cdots & a_{1 \times N \times 1} \cdots a_{1 \times N \times T} \\ \vdots & \vdots & \vdots & \vdots \\ a_{K \times 1 \times 1} \cdots a_{K \times 1 \times T} & a_{K \times 2 \times 1} \cdots a_{K \times 2 \times T} & \cdots & a_{K \times N \times 1} \cdots a_{K \times N \times T} \\ \vdots & \vdots & \vdots & \vdots \\ a_{K \times M \times 1} \cdots a_{K \times M \times T} & a_{K \times M \times 2} \cdots a_{K \times M \times T} & \cdots & a_{K \times M \times N \times 1} \cdots a_{K \times M \times N \times T} \end{bmatrix}$$

的均值为零),转换成矩阵 \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1 \times N \times T} \\ b_{21} & b_{22} & \cdots & b_{2 \times N \times T} \\ \vdots & \vdots & \vdots & \vdots \\ b_{K \times M \times 1} & b_{K \times M \times 2} & \cdots & b_{K \times M \times N \times T} \end{bmatrix},$$

其中:

$$b_{ij} = a_{ij} - \frac{1}{K \times M} \sum_{i=1}^{K \times M} a_{ij}, \quad (1)$$

构建协方差阵 \mathbf{COV} :

$$\mathbf{COV} = \frac{1}{K \times M} \mathbf{B}^T \mathbf{B}, \quad (2)$$

特征值 λ_i 对应的特征向量为:

$$\mathbf{A}_2 = \mathbf{B} \mathbf{A}_1 = \begin{bmatrix} \sum_{i=1}^{N \times T} b_{1i} \times v_{i1} & \sum_{i=1}^{N \times T} b_{1i} \times v_{i2} & \cdots & \sum_{i=1}^{N \times T} b_{1i} \times v_{ip} \\ \sum_{i=1}^{N \times T} b_{2i} \times v_{i1} & \sum_{i=1}^{N \times T} b_{2i} \times v_{i2} & \cdots & \sum_{i=1}^{N \times T} b_{2i} \times v_{ip} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{N \times T} b_{K \times Mi} \times v_{i1} & \sum_{i=1}^{N \times T} b_{K \times Mi} \times v_{i2} & \cdots & \sum_{i=1}^{N \times T} b_{K \times Mi} \times v_{ip} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1(1) \\ \mathbf{a}_1(2) \\ \vdots \\ \mathbf{a}_1(K \times M) \end{bmatrix}, \quad (5)$$

$$\mathbf{u}_i = (v_{1i}, v_{2i}, \dots, v_{N \times Ti})^T. \quad (3)$$

3.2.3 主成分的提取

按照特征值由大到小,即 $\lambda_1 > \lambda_2 > \dots > \lambda_{N \times T}$,顺序排列对应的特征向量,并根据累计贡献率 $\sum_{i=1}^p \lambda_i / \sum_{i=1}^{N \times T} \lambda_i > 95\%$,取前 p 个特征向量构成矩阵 \mathbf{A}_1 :

$$\mathbf{A}_1 = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p). \quad (4)$$

矩阵 \mathbf{B} 乘以矩阵 \mathbf{A}_1 ,得到矩阵 \mathbf{A}_2 ,即将不同类别的动作样本投影到了新的空间,可以证明在此空间内,在新的维度上不同样本之间方差较大。

$$\frac{1}{K * M} \mathbf{A}_2^T \mathbf{A}_2 = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_p \end{bmatrix}. \quad (6)$$

这样每个动作就可以用一个低维度的向量 $\mathbf{a}_1(i)$ (i 是动作模板标号) 表示, 即为动作主成分。

3.2.4 单特征构造协方差矩阵

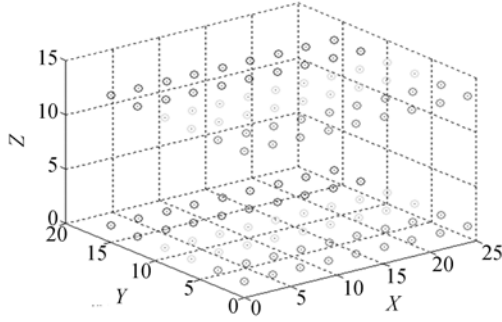


图 2 单特征原始模板图

Fig. 2 Original template of single feature

如图 2 所示, 三维坐标系里, Z 轴每一个刻度表示一个特征维度, 共计 N 个, X 轴表示时间帧序列, 标号为 $1 \sim T$, Y 轴表示动作模板, 共计 $K * M$ 个。在垂直于 Z 轴且对应 Z 轴每一个刻度的平面上, 同一种颜色代表同一类型的动作, 它包含 K 个子类, 这样, 在每一个特征维度上, 每一个动作模板用向量 $\mathbf{a}^j(i)$ 表示, 定义其为单特征:

$$\mathbf{a}^j(i) = (a_{11}^j, a_{12}^j, \dots, a_{1T}^j), \quad (7)$$

$$1 \leq i \leq K * M, 1 \leq j \leq N.$$

把向量 $\mathbf{a}^j(i)$ 按行顺序排列, 即每一个模板向量作为一行, 构成 $K * M * T$ 的矩阵, 记为 \mathbf{A}^j ,

$$\mathbf{A}_2^j = \mathbf{B}^j \mathbf{A}_1^j = \begin{bmatrix} \sum_{i=1}^T b_{1i}^j \times v_{11}^j & \sum_{i=1}^T b_{1i}^j \times v_{12}^j & \dots & \sum_{i=1}^T b_{1i}^j \times v_{1p}^j \\ \sum_{i=1}^T b_{2i}^j \times v_{11}^j & \sum_{i=1}^T b_{2i}^j \times v_{12}^j & \dots & \sum_{i=1}^T b_{2i}^j \times v_{1p}^j \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^T b_{k * M * i}^j \times v_{11}^j & \sum_{i=1}^T b_{k * M * i}^j \times v_{12}^j & \dots & \sum_{i=1}^T b_{k * M * i}^j \times v_{1p}^j \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^j(1) \\ \mathbf{a}_1^j(2) \\ \vdots \\ \mathbf{a}_1^j(K * M) \end{bmatrix}, \quad (12)$$

$$\frac{1}{K * M} \mathbf{A}_2^{jT} \mathbf{A}_2^j = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_p \end{bmatrix}. \quad (13)$$

作为原始数据矩阵:

$$\mathbf{A}^j = \begin{bmatrix} a_{11}^j & a_{12}^j & \dots & a_{1T}^j \\ a_{21}^j & a_{22}^j & \dots & a_{2T}^j \\ \vdots & \vdots & \vdots & \vdots \\ a_{k * M * 1}^j & a_{k * M * 2}^j & \dots & a_{k * M * T}^j \end{bmatrix}.$$

对矩阵 \mathbf{A}^j 进行预处理, 使其每一列的均值为零, 转换成矩阵 \mathbf{B}^j :

$$\mathbf{B}^j = \begin{bmatrix} b_{11}^j & b_{12}^j & \dots & b_{1T}^j \\ b_{21}^j & b_{22}^j & \dots & b_{2T}^j \\ \vdots & \vdots & \vdots & \vdots \\ b_{k * M * 1}^j & b_{k * M * 2}^j & \dots & b_{k * M * T}^j \end{bmatrix},$$

其中:

$$b_{1i}^j = a_{1i}^j - \frac{1}{K * M} \sum_{i=1}^{K * M} a_{1i}^j, \quad (8)$$

构建协方差阵 \mathbf{COV} :

$$\mathbf{COV} = \frac{1}{K * M} \mathbf{B}^{jT} \mathbf{B}^j, \quad (9)$$

特征值 λ_i^j 对应的特征向量为:

$$\mathbf{u}_i^j = (v_{1i}^j, v_{2i}^j, \dots, v_{Ti}^j)^T. \quad (10)$$

3.2.5 提取主成分

按照特征值由大到小 ($\lambda_1^j > \lambda_2^j > \dots > \lambda_T^j$) 顺序排列对应的特征向量, 并根据累计贡献率 $\sum_{i=1}^p \lambda_i^j / \sum_{i=1}^T \lambda_i^j > 95\%$ 取前 p 个特征向量构成投影矩阵 \mathbf{A}_1^j :

$$\mathbf{A}_1^j = (\mathbf{u}_1^j, \mathbf{u}_2^j, \dots, \mathbf{u}_p^j). \quad (11)$$

矩阵 \mathbf{B}^j 乘以矩阵 \mathbf{A}_1^j , 得到降维之后的模板矩阵 \mathbf{A}_2^j , 即将不同类别的动作样本投影到了新的空间, 可以证明在此空间内, 在新的维度上不同的样本之间方差较大, 便于分类。

这样, 每个动作在每一个特征维度上就可以用一个低维度的向量 $\mathbf{a}_1^j(i)$ 表示:

$$\mathbf{a}_1^j(i) = (a_{11}^j, a_{12}^j, \dots, a_{1p}^j), \quad (14)$$

$$1 \leq i \leq K * M, 1 \leq j \leq N.$$

将测试样本在每一个特征维度上, 由高维的

向量 $a^i(i)$ 经过上述转换变成一个低维度的向量,作为提取的动作主成分,记为 $a_i(i)$ 。

4 实验与结果

4.1 实验数据说明

对人的 6 种简单的行为动作进行了识别,包括:打电话、喝水、使用遥控器、倒水、起立和坐下。参照文献[1]提取人姿态描述特征向量,如表 1 所

表 1 人体姿态描述向量组成

Tab.1 Composition of gesture description vector

| 位置 | 属性名称 | 位置 | 属性名称 |
|----|--------------------------|----|-------------------------|
| 1 | Aneck_rshoulder_relbow | 13 | Arhip_rknee_rfoot |
| 2 | Arshoulder_relbow_rhand | 14 | Alshoulder_torso_lhip |
| 3 | Aneck_lshoulder_l elbow | 15 | Atorso_lhip_lknee |
| 4 | Alshoulder_l elbow_lhand | 16 | Alhip_lknee_lfoot |
| 5 | Aneck_rshoulder_torso | 17 | Aneck_torso_rhip_rknee |
| 6 | Arshoulder_relbow_torso | 18 | Aneck_torso_rknee_rfoot |
| 7 | Arelbow_rhand_torso | 19 | Aneck_torso_lhip_lknee |
| 8 | Aneck_lshoulder_torso | 20 | Aneck_torso_lknee_lfoot |
| 9 | Alshoulder_l elbow_torso | 21 | Rrat_head_to_rhand |
| 10 | Aelbow_lhand_torso | 22 | Rrat_torso_to_rhand |
| 11 | Arshoulder_torso_rhip | 23 | Rrat_head_to_lhand |
| 12 | Atorso_rhip_rknee | 24 | Rrat_torso_to_lhand |

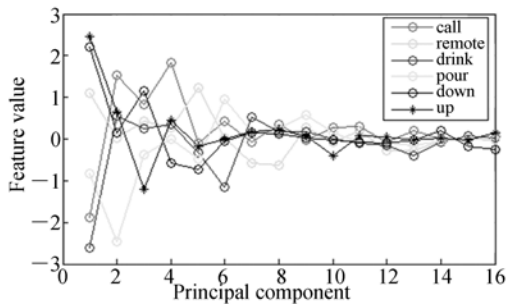


图 3 串特征模板图

Fig. 3 Series feature template

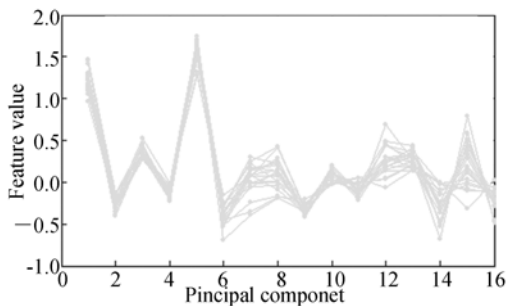


图 5 用遥控器图

Fig. 5 Using remote control

示。利用 Kinect 采集了这 6 种动作的 24 组三维特征向量数据,每种动作采集 50 组样本,其中 30 组为训练样本,20 组为测试样本。

4.2 串特征数据处理效果及识别结果

将串特征数据处理效果,绘制于图 3~9 中,图 3 不同颜色的曲线代表不同的动作模板,图 4~9 表示不同动作的测试样本变换后的轨迹,可以看出经过 PCA 之后,同类型的动作在主成分上较好的集中在一起,不同类型的动作又具有不同的变化趋势,而且差异较明显,利于分析判别。对于部分行为动作的特征描述向量,非常相似,例如喝水和打电话,所以,识别率也较低,见表 2。该方法适合于分析识别多种类型在小范围空间内具有较多细节变化的人体行为。

表 2 串特征动作识别混淆矩阵

Tab.2 Confusion matrix of series feature motion identification

| | 打电话 | 用遥控器 | 喝水 | 倒水 | 起立 | 坐下 |
|------|------|------|------|----|----|----|
| 打电话 | 0.95 | | 0.05 | | | |
| 用遥控器 | | 1 | | | | |
| 喝水 | 0.2 | | 0.8 | | | |
| 倒水 | | | | 1 | | |
| 起立 | | | | | 1 | |
| 坐下 | | | | | | 1 |

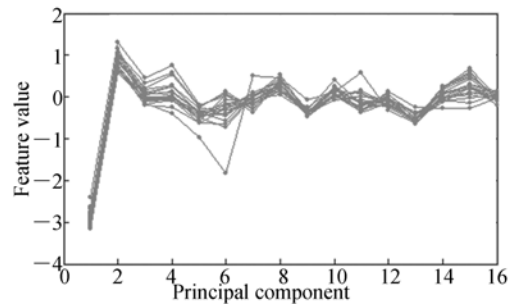


图 4 打电话图

Fig. 4 Call

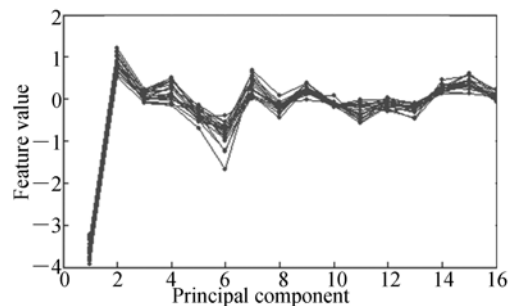


图 6 喝水图

Fig. 6 Drink

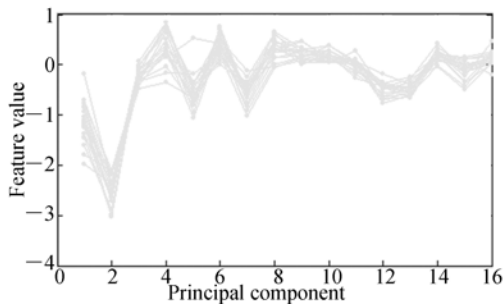


图 7 倒水图
Fig. 7 Pour

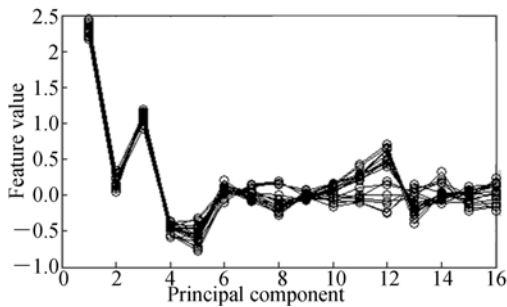


图 8 坐下图
Fig. 8 Sit down

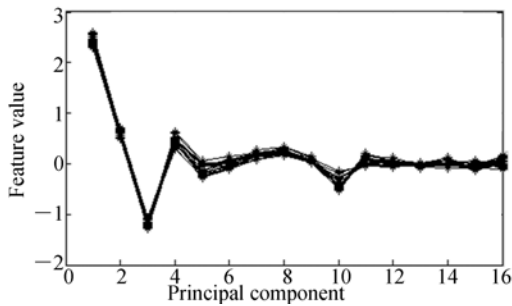


图 9 起立图
Fig. 9 Stand

4.3 单特征数据处理效果及识别结果

以倒水动作为例,图 10~19 分别绘制了 6 类动作模板,对应的倒水动作的测试样本,以及经过 PCA 处理后,在不同单特征主成分上的变化轨迹,其中特征标号见表 1。在模板图上,不同的颜色代表不同类型的动作,由图可以看出,通过对各类动作每一个特征的时间序列值做主成分分析之后,不同类型的动作在同一种特征的几个主要的时间点上有较大的差异,尤其是有些特征可以作为某类动作的标志性特征,例如,倒水动作,与左臂相关的特征,其变化趋势与其他类型的动作有较大差异。通过这些特征,可以很容易地将一类

动作与其他动作区分开。由图可以看出同类型的动作具有较好的聚集性,而且与模板的分布大致相同。这样,测试样本能够很好地被与之同类的模板捕获,有利于识别。最后,通过 K 近邻算法,综合所有特征的分析判别给出最终的识别结果,见表 3。该方法识别精度也较高,但是,计算量较大,影响识别效率。

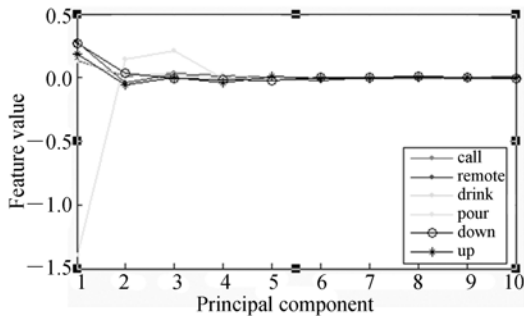


图 10 模板 4 号特征图
Fig. 10 Feature of the No. 4 template

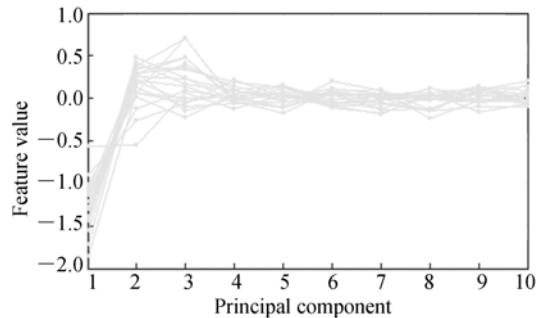


图 11 倒水动作 4 号特征图
Fig. 11 No. 4 feature of pour movement

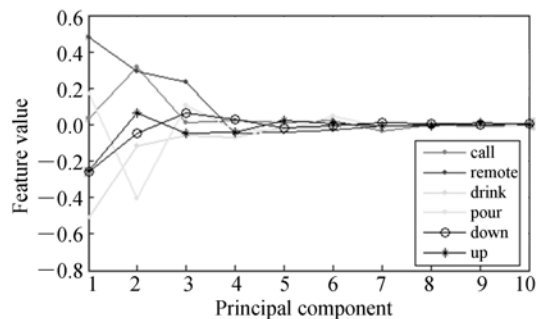


图 12 模板 7 号特征图
Fig. 12 No. 7 feature of template

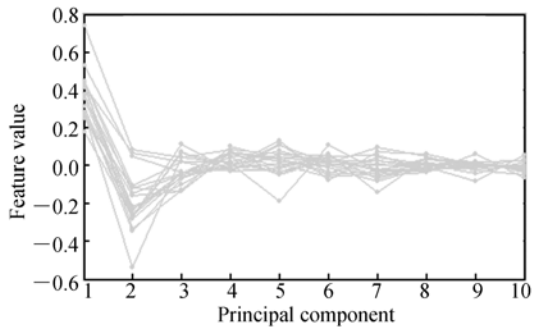


图 13 倒水动作 7 号特征图

Fig. 13 No. 7 feature of pour movement

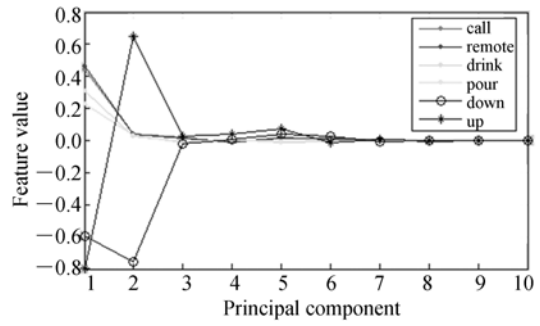


图 16 模板 13 特征图

Fig. 16 No. 13 feature of template

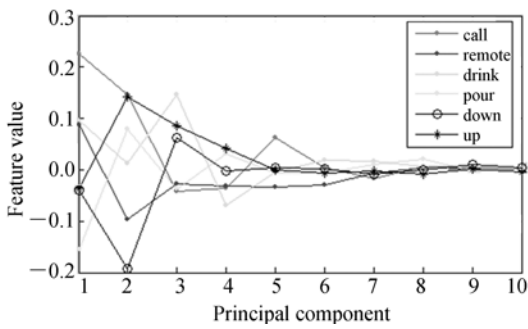


图 14 模板 8 号特征图

Fig. 14 No. 8 feature of template

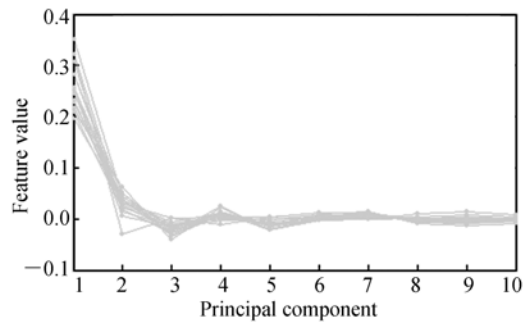


图 17 倒水动作 13 号特征图

Fig. 17 No. 13 feature of pour movement

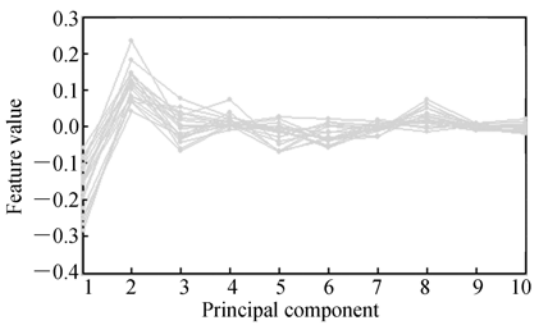


图 15 倒水动作 8 号特征图

Fig. 15 No. 8 feature of pour movement

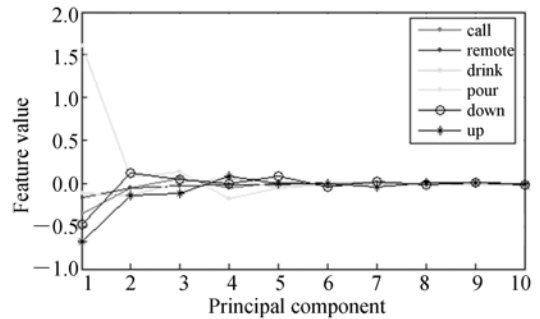


图 18 模板 21 特征图

Fig. 18 No. 21 feature of template

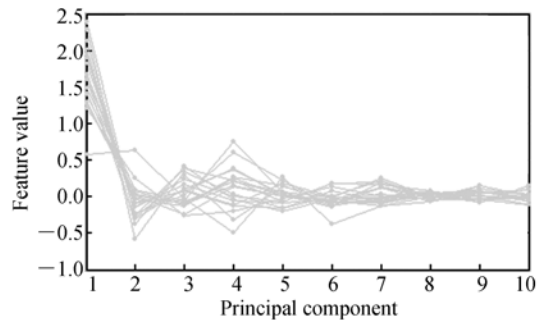


图 19 倒水动作 21 号特征图

Fig. 19 No. 21 feature of pour movement

表 3 单特征动作识别混淆矩阵

Tab. 3 Confusion matrix of single feature motion identification

| | 打电话 | 用遥控器 | 喝水 | 倒水 | 起立 | 坐下 |
|------|------|------|------|----|----|----|
| 打电话 | 0.9 | | 0.1 | | | |
| 用遥控器 | | 1 | | | | |
| 喝水 | 0.15 | | 0.85 | | | |
| 倒水 | | | | 1 | | |
| 起立 | | | | | 1 | |
| 坐下 | | | | | | 1 |

5 结 论

本文使用 Kinect 采集人体动作特征描述数据,并通过两种方法做主成分分析,结合 K 近邻算法进行了动作识别。实验结果表明,识别精度都达到 80% 以上,单个样本识别时间分别是 1.67 ms 和 3.93 ms,两种方法提取的主成分对同类动作聚集性较好,对不同类型的动作离散程度也较好,而且算法运行速度快,利于实时识别。

参考文献:

- [1] 田国会,尹建芹,韩旭,等.一种基于关节点信息的人体行为识别新方法[J].*机器人*,2014,36(3):285-292.
TIAN G H, YIN J Q, HAN X, *et al.*. Novel human activity recognition method using joint points information [J]. *Robot*, 2014, 36(3):285-292. (in Chinese)
- [2] LV F J, NEVATIA R. Single view human action recognition using key pose matching and viterbi path searching [C]. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR' 07, IEEE*, 2007:1-8.
- [3] RAPTIS M, KIROVSKI D, HOPPE H. Real-time classification of dance gestures from skeleton animation [C]. *Proceedings of the 2011 ACM SIGGRAPH. Eurographics Symposium on Computer Animation*, 2011:147-156.
- [4] ZHAO H Y, LIU ZH J. Human action recognition based on non-linear SVM decision tree [J]. *Journal of Computational Information Systems*, 2011, 7(7):2461-2468.
- [5] YUN K, HONORIO J, CHATTOPADHYAY D, *et al.*. Two-person interaction detection using body-pose features and multiple instance learning [C]. *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, *IEEE*, 2012:28-35.
- [6] ELLIS C, MASOOD S Z, TAPPEN M F, *et al.*. Exploring the trade-off between accuracy and observational latency in action recognition[J]. *International Journal of Computer Vision*, 2013,101(3):420-436.
- [7] LIN S Y, SHIE C K, CHEN S C, *et al.*. Human action recognition using action trait code [C]. *2012 21st International Conference on Pattern Recognition (ICPR) IEEE*, 2012:3456-3459.
- [8] SUN J, WU X, YAN SH C, *et al.*. Hierarchical spatio-temporal context modeling for action recognition[C]. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, IEEE*, 2009:2004-2011.
- [9] DOLLÁR P, RABAUD V, COTTRELL G, *et al.*. Behavior recognition via sparse spatio-temporal features [C]. *2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, IEEE*, 2005:65-72.
- [10] XIA L, CHEN C C, AGGARWAL J K. View invariant human action recognition using histograms of 3d joints [C]. *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, *IEEE*, 2012:20-27.
- [11] LI W Q, ZHANG ZH Y, LIU Z C. Action recognition based on a bag of 3d points[C]. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, *IEEE*, 2010:9-14.
- [12] GUHA T, WARD R K. Learning sparse representations for human action recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012,34(8):1576-1588.
- [13] CASTRODAD A, SAPIRO G. Sparse modeling of human actions from motion imagery [J]. *International Journal of Computer Vision*, 2012,100(1):

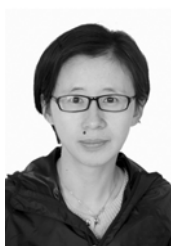
- 1-15.
- [14] COHEN I, LI H. Inference of human postures by classification of 3D human body shape [C]. *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, AMFG 2003. IEEE, 2003:74-81.
- [15] XIONG J, LIU ZH J. *Human Motion Recognition Based on Hidden Markov Models* [M]. Berlin: Springer Berlin Heidelberg, 2007: 464-471.
- [16] HAN L, WU X X, LIANG W, *et al.*. Discriminative human action recognition in the learned hierarchical manifold space [J]. *Image and Vision Computing*, 2010, 28(5): 836-849.
- [17] MÜLLER M, RÖDER T. Motion templates for automatic classification and retrieval of motion capture data [C]. *Proceedings of the 2006 ACM SIGGRAPH. Eurographics Symposium on Computer Animation*, Eurographics Association, 2006:137-146.
- [18] MÜLLER M, BAAK A, SEIDEL H P. Efficient and robust annotation of motion capture data [C]. *Proceedings of the 2009 ACM SIGGRAPH, Eurographics Symposium on Computer Animation*, ACM, 2009:17-26.
- [19] REYES M, DOMINGUEZ G, ESCALERA S. Feature weighting in dynamic timewarping for gesture recognition in depth data [C]. *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2011: 1182-1188.

作者简介:



刘志强(1988—),男,山东枣庄人,硕士研究生,2012年于山东大学机电一体化工程专业获得学士学位,主要从事图像处理 and 机器视觉领域的研究。E-mail: liuzq@qq.com

导师简介:



尹建芹(1978—),女,山东潍坊人,副教授,2000年于山东工业大学获得学士学位,2002年、2013年于山东大学控制科学与工程学院分别获得硕士、博士学位,主要从事图像处理及模式识别方面的研究。E-mail: ise_yinj@ujn.edu.cn