

DOI: 10.11992/tis.201611008

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20170228.1323.004.html>

RGBD 人体行为识别中的自适应特征选择方法

龚冬颖^{1,2}, 黄敏^{1,2}, 张洪博³, 李绍滋^{1,2}

(1. 厦门大学 智能科学与技术系, 福建 厦门 361005; 2. 厦门大学 福建省仿脑智能系统重点实验室, 福建 厦门 361005; 3. 华侨大学 计算机科学与技术学院, 福建 厦门 361005)

摘要: 目前在 RGBD 视频的行为识别中, 为了提高识别准确率, 许多方法采用多特征融合的方式。通过实验分析发现, 行为在特定特征上的分类效果好, 但是多特征融合并不能体现个别特征的分类优势, 同时融合后的特征维度很高, 时空开销大。为了解决这个问题, 提出了 RGBD 人体行为识别中的自适应特征选择方法, 通过随机森林和信息熵分析人体关节判别力, 以高判别力的人体关节的数量作为特征选择的标准。通过该数量阈值的筛选, 选择关节特征或者关节相对位置作为行为识别特征。实验结果表明, 该方法相比于特征融合的算法, 行为识别的准确率有了较大提高, 超过了大部分算法的识别结果。

关键词: 人体行为识别; 自适应特征选择; 信息熵; 随机森林

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2017)01-0001-07

中文引用格式: 龚冬颖, 黄敏, 张洪博, 等. RGBD 人体行为识别中的自适应特征选择方法[J]. 智能系统学报, 2017, 12(1): 1-7.

英文引用格式: GONG Dongying, HUANG Min, ZHANG Hongbo, et al. Adaptive feature selection method for action recognition of human body in RGBD data[J]. CAAI transactions on intelligent systems, 2017, 12(1): 1-7.

Adaptive feature selection method for action recognition of human body in RGBD data

GONG Dongying^{1,2}, HUANG Min^{1,2}, ZHANG Hongbo³, LI Shaozi^{1,2}

(1. Intelligent Science & Technology Department, Xiamen University, Xiamen 361005, China; 2. Fujian Key Laboratory of Brain-like Intelligent Systems, Xiamen University, Xiamen 361005, China; 3. Computer Science & Technology School, Huaqiao University, Xiamen 361005, China)

Abstract: Many methods adopt the technique of multi-feature fusion to improve the recognition accuracy of RGBD video. Experimental analyses revealed that the classification effect of certain behavior in some features is good; however, multi-feature fusion cannot reflect the classification superiority of certain features. Moreover, multi-feature fusion is highly dimensional and considerably expensive in terms of time and space. This research proposes an adaptive feature selection method for RGBD human-action recognition to solve this problem. First, random forest and information entropy were used to analyze the judgment ability of the human joints, whereas the number of human joints with high judgment ability were chosen as the feature selection criterion. By screening the threshold number, either the joint feature or the relative positions of the joints was used as the recognition feature of action. Experimental results show that compared with multi-feature fusion, the method significantly improved the accuracy of action recognition and outperformed most other algorithms.

Keywords: action recognition of human body; adaptive feature selection; information entropy; random forest

人体行为识别在计算机视觉领域上有很重要的研究意义, 广泛应用于异常人体行为识别、用户身份识别、基于内容的体育视频检索、智能家居环

境等。目前, 人体行为识别主要困难在于用于人体行为识别的视频持续在一系列的时间和空间, 视频基于一个很高维的时间和空间, 行为表示困难, 同时同一个动作受执行个体、速度、衣着、光照、遮挡和摄像头拍摄角度的影响, 行为的类内距离大, 类间差异小。新推出的 3-D 体感摄像机(例如

收稿日期: 2016-11-07. 网络出版日期: 2017-02-28.

基金项目: 国家自然科学基金项目(61572409, 61571188, 61202143); 福建省自然科学基金项目(2013J05100); 中医健康管理福建省 2011 协同创新中心项目.

通信作者: 李绍滋. E-mail: szlig@xmu.edu.cn.

Kinect) 除了提供RGB视频外,还有深度信息、关节点信息,能快速分割前景背景,为人体行为识别带来了新的思路。因此,目前许多行为识别的工作都是基于RGB视频序列的。

现有RGBD视频的行为识别中,特征主要分为两种:深度表观特征和关节点运动特征。深度表观特征主要基于某一时空领域内的深度点云信息,侧重描述时空域内的深度表观细节,可以是某一时空域内的占有模式^[1]、4-D法向量统计^[2]和运动历史图梯度统计^[3]等。关节点运动特征主要基于关节点3-D位置,通过关节点位置的时间变化^[1]和空间偏移值^[4]来描述行为,此方法特征紧凑,避免了大量的冗余计算。但在一些人物交互的行为中,仅采用关节点运动特征无法刻画交互物体,需要引入深度表观特征。

因此,很多学者尝试采用特征融合的方法,同时使用多模态特征来描述行为。在一些方法中^[5-7]结合了多种从2-D图像提取的特征,使用单个分类器进行分类。还有一些方法中设计了多种分类器。文献[8-9]通过对各种分类器的分类结果进行整合,从而提高了行为识别的准确率。

但是这些特征融合方法普遍存在3个缺陷:1) 视频存在类内类间的区别,用多个特征串联可能存在结果受其中某个特征影响,结果达不到预期效果;2) 无法对特征与类别之间的关系进行分析;3) 多特征融合的维度很高,对于训练以及测试需要大量的时间,而目前普遍采用的降维方式如PCA会降低准确率。因此,本文设计了一种特征选择方式,根据关节特征分析关节点的判别类型,选择不同的行为特征用于识别行为。

本文提出的自适应特征选择方法,分别采用HON4D关节点表观特征^[10]和关节点相对距离特征作为行为表示特征。要实现上述目的,主要存在的问题是在测试过程中,需要确定特征选择的依据。对于MSR-Daily Activity 3-D数据集来说,尽管已经知道HON4D关节点特征在喝水、吃东西等类别上效果较好,相对距离特征对于看书、打电话、玩游戏等行为的识别较好,但是在测试过程中测试样本却是不知道该使用什么特征来表现。通过实验分析发现,采用HON4D关节点特征分类较好时,对应于高判别力的关节点数量较多。因此在本文中采用随机森林对HON4D关节点进行建模,根据关节点的类别投票结果熵,将其结果作为关节点的判别力,提出了自适应熵阈值,选择高判别力的关节点;同时基于高判别力关节点数量的特征选择方法,用

于表示行为,进而提高行为的识别率。

1 相关技术介绍

目前,在RGBD的人体行为识别中主要采用基于关节点序列的运动轨迹特征以及基于局部或者整体的表观特征。

关节点运动轨迹特征是指骨架在时间序列上的变化所提取的特征。Yang^[11]提出基于关节点3-D位置的位置特征、运动特征和位移特征,使用贝叶斯最邻近法则(NBNN)进行分类。Wang等^[1]在Actionlet Ensemble中使用到3-D关节点相对位置特征,并且添加了傅里叶时间金字塔提高精度。Xia等^[4]在3-D关节点球形位置计算直方图,通过隐马尔可夫模型建立3-D关节点的时间演变模型进行分类。

另一种主要特征是基于局部或整体的表观表示。Wang等^[1]提出LOP作为关节点的深度表观信息,用此特征来描述人体与物体的交互。HOPC是Rahmani^[12]提出的基于3-D点云的表示方法,围绕该点可以得到自适应时空支撑体。SNV是Yang等^[2]在时空体内描述的深度表观信息和变化信息,该特征是由深度序列划分成时空网格,并聚类底层法向量获得的。

不同的特征在不同行为上具有不同的识别强度。不少文章也选择了多特征的方法,使得这些特征存在互补性。除了一些单纯使用不同特征简单连接进行融合以外^[1,13-14],还有一些特征进行了融合处理。Gao等^[15]提取深度序列图的不同特征,然后进行多特征映射并且使用字典学习模型。Liu等^[16]提出基于3D²的CNN框架,这个框架可以自动从原始深度视频序列提取时空特征。Li等^[17]使用组合稀疏正则化,提出用多特征稀疏融合(MFSF)排序以获得特征的共享和特定结构的重要性。

现有方法中,结合关节点运动信息和局部表观信息能很好地描述行为。关节点运动信息虽然可以表现人体运动姿态,但是在捕获变化的时候无法避免运动速度的影响。在关节点局部表观信息上,虽然可以捕捉一些人体的姿态信息,但是描述运动特点有限。HON4D的特征较短,经过傅里叶变化后,保留的低频信息不仅特征短,而且判别力强。因此,在本文提出的模型中,采用关节点相对距离作为运动信息模型,关节点位置的HON4D特征作为表观信息模型,根据关节点的判别力,选择不同的特征,进而增强分类性能。在特征处理上采用

Actionlet Ensemble 提出的傅里叶时间金字塔,此方法能够很好地去掉时间上对分类带来的影响。

2 自适应特征选择方法

在本节中,我们将详细介绍自适应特征选择方法。本文考虑了两种特征,即体现关节点表观信息的 HON4D 特征^[10]和运动变化的关节点相对距离特征。在特征选择上,基于关节点熵进行自适应特征选择,人体行为识别方法如图 1 所示。

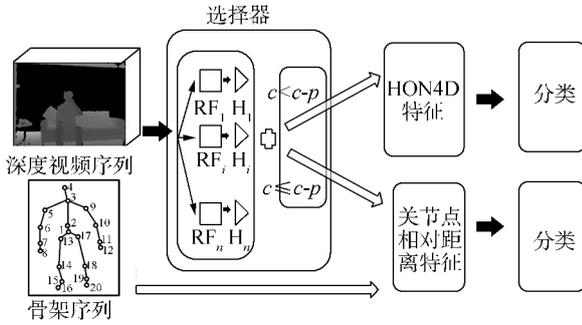


图 1 自适应特征选择方法框架

Fig.1 The frame of adaptive feature selection method

2.1 特征介绍

首先,关节点相对距离特征以臀部为参考点,计算每一帧内各关节点相对参考点的位移。记第 i 个关节点在第 f 帧的坐标为 $J_i(f) = (x_i^f, y_i^f, z_i^f)$, n_{Joint} 为关节点总数量,则每一帧的相对关节点位移为

$$V(f) = \{J_i(f) - J_j(f) \mid i = 2, 3, 4, \dots, n_{\text{Joint}}; j = 1\} \quad (1)$$

再将各帧的相对关节点位移串联,进行傅里叶变换,即

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (2)$$

取变换后的低频信息,最终得到关节点相对距离特征。其次,在表观特征上,我们对 HON4D 进行了改进。根据关节点判别力筛选出前 N 个判别力强的关节点特征。关节点判别力是指某个关节点的特征在数据集上准确率的评判值。为了得到不同关节点判别力评价,我们将每个关节点 HON4D 特征随机森林(RF)中进行模型训练,得到各个模型中的袋外估计错误率,以使用来评价关节点的判别力。随机森林是一个具有 T 棵决策树的集成模型。本文采用自举采样法,在训练阶段,每棵决策树都会在训练集中生成自举采样,大约丢弃 37% 的样本,根据这些丢弃样本去计算袋外估计错误率。

对于给定行为 $c(c = 1, 2, \dots, c_{\text{Cls}})$, c_{Cls} 为数据集的类别,根据各关节点 $p(p = 1, 2, \dots, n_{\text{Joint}})$ 在相应的 RF 模型上计算的该行为上的袋外估计错误率,每个行为将得到一个关节点错误率升序的排序表。则

最终数据为 $(1+c_{\text{Cls}})$ 个排序表,其中第 1 列为所有数据集对于其中一个行为的平均错误率。最终,从基于整个数据集上的关节点升序表中取出前 N 个关节点,将其 HON4D 特征串联,得到表观特征。

2.2 基于熵的自适应特征选择方法

关节点相对位移特征,适用于运动变化较大的行为,如静止、站起、坐下等,而表观特征则可以刻画交互物体和局部细节变化的行为,如看书、打电话等。因此,应根据行为的特点选择合适的特征。为此,本文提出了基于熵的自适应特征选择方法。

首先,为了评估表观特征对于各行为的判别力,我们将各关节点的表观特征输入 RF 模型,从训练模型中得到决策树的投票结果,由投票结果的不确定性去判断该类特征在行为分类上是否具备代表性。确定性强,表示该特征能够充分代表样本,反之,则使用关节点相对位移。由于任何信息都存在冗余,冗余的程度与不确定性有关,排除冗余后的平均信息量称为“信息熵”。不确定性函数为

$$f(p) = \log \frac{1}{p} = -\log p \quad (3)$$

对于整个信息源而言,熵的定义为

$$H(U) = E(-\log p_i) = -\sum_{i=1}^n p_i \log p_i \quad (4)$$

因此,在本文的算法中,我们将在关节点特征的 RF 模型中引入熵的概念。行为 c 在训练数据集的第 $s(s = 1, 2, \dots, S)$ 个样本中的第 $n_j(n_j = 1, 2, \dots, n_{\text{Joint}})$ 个关节点,通过使用随机森林模型得到的 T 棵决策树,利用决策树进行投票分类,则每个关节点都会对训练样本产生投票结果。该样本的投票结果为 $(n_{\text{Joint}} \times c_{\text{Cls}})$ 个,每个投票结果表示为 $v_i^{n_j, s, f}(i = 1, 2, \dots, c_{\text{Cls}})$,投票的概率为

$$p_i^{n_j, s, f} = v_i^{n_j, s, f} / \sum_{k=1}^{c_{\text{Cls}}} v_k^{n_j, s, f} \quad (5)$$

由此得到每个样本的信息熵,并求出 S 个样本的平均信息熵,因此在 c_{Cls} 类行为中可得到在该模型的信息熵 $E_{n_j}(n_j = 1, 2, \dots, n_{\text{Joint}})$ 。

在训练模型中可以得到平均信息熵,根据此信息熵得到针对于此关节点模型的熵的阈值。

作为特征自适应选择的选择器时,测试样本在经过 20 个关节点模型时,可以类似于训练模型得到 20 个熵。本文中熵的意义为,HON4D 特征的熵越大,信息不确定性越大,表示测试样本使用 HON4D 不能准确描述行为,因此不使用关节点特征,而选择关节点相对距离特征。

在实验中,本文设定当测试样本的熵超过训练

模型中得到的平均信息熵的个数为 C ,设置阈值为 C_p , $C < C_p$,选择 HON4D 关节点表观特征 ,否则选择关节点相对距离特征。

算法详见算法 1。

算法 1 基于熵的自适应特征选择算法

输入 训练样本平均熵 ,测试样本 HON4D 特征与关节点相对距离特征 ,20 个关节点模型 ,关节点相对距离模型 ,HON4D 关节点组合模型 ,阈值 C_p ,平均信息熵 $E_{n_j}(n_j=1, 2, \dots, n_{joint})$

输出 预测标签。

- 1) 初始化 C 为 0。
- 2) for 第 n_j 个关节点 in 所有关节点数目 do。
- 3) 第 n_j 个关节点的 HON4D 特征作为测试样本特征 ,由相对应的关节点模型测试得到 c_{cls} 个投票结果。
- 4) 计算 c_{cls} 个投票结果的投票概率。
- 5) 计算该模型下的投票熵。
- 6) 比较 E_{n_j} 与该模型下的投票熵 ,若 E_{n_j} 小 ,则 $C+1$ 。
- 7) end for。
- 8) 比较 C 与 C_p ,若 C 小 ,则使用关节点相对距离特征 ;若 C 大 ,则使用 HON4D 特征。
- 9) 输出预测标签。

3 实验结果

为了验证本方法的可用性 ,我们将在 MSR-Daily Activity 3D 数据集上验证。该数据使用 Kinect 设备

拍摄 ,是日常活动的数据集 ,如图 2。在这个数据集中 ,动作覆盖手、腿、躯干以及其他一些互动行为。其中数据集中每个行为由 10 个人完成 ,每个人分别以不同的姿态完成。该数据集样本数为 320。在本文的实验中 ,我们将在每个行为中选取 5 个人所完成的 10 个样本 ,也就是其中 160 个作为训练样本 ,其余 160 个作为测试样本。

在设置参数方面 ,HON4D 中网格大小为 $60 \times 60 \times 10$,步长为 $3 \times 3 \times 1$,傅里叶变换取低频信息 10。在本文的方法中 ,需要设置阈值 :每个关节点的信息熵阈值和高判别力关节点的数量阈值。其中 ,每个关节点的信息熵阈值采用各自的平均信息熵作为阈值 ,是一种自适应的方法 ,不需要手动调节。同时 ,通过实验对比 ,关节点数量的阈值 $C_p = 16$ 时 ,行为识别的准确率最高。

3.1 特征介绍

在 HON4D 的关节点判别力表示中 ,我们采用错误率的方式进行衡量 ,可以从 20 个关节点模型中分别得到 16 个行为的错误率。从图 2 可以看出 ,某个行为中错误率较低的关节点在该行为中占据比较重要的位置 ,比如图 2(a) 的行为为喝水的动作 ,重要的关节点主要集中在手部、头部 ,这些关节点的错误率也较低。

根据这 16 个行为的关节点错误率可以得到针对整个数据集的平均错误率 ,由平均错误率可以得到相应关节点的排序为 $\{2, 10, 11, 1, 12, 5, 9, 3, 17, 4, 7, 13, 6, 8, 14, 18, 15, 19, 16, 20\}$ 。

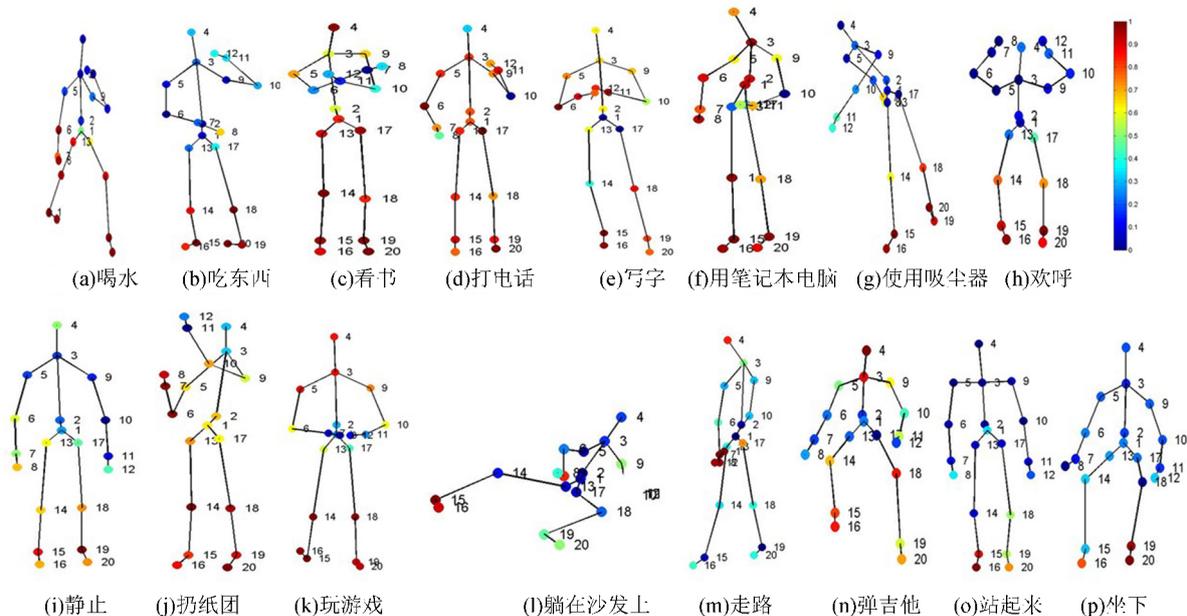


图 2 关节点错误率骨架图

Fig.2 The skeleton diagram of joint point error

实验中,我们采用交叉验证法,发现取前 12 个关节子序列时,分类模型结果最佳。图 3 给出了测试集在不同关节个数下的准确率变化,准确率随关节个数增加而提升,但到达某个范围值后趋于稳定。在后续实验中,我们选取前 12 个关节子序列的 HON4D 特征作为表观特征。

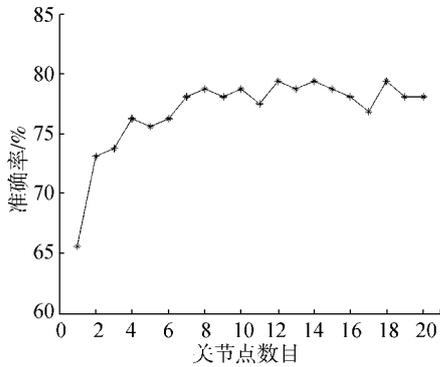


图 3 HON4D 模型关节准确率
Fig.3 The joint accuracy of HON4D model

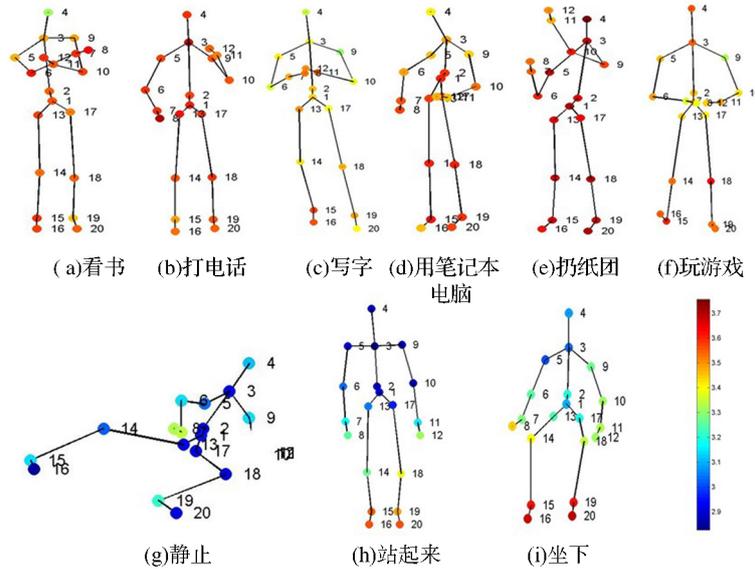


图 4 关节熵骨架图
Fig.4 The skeleton diagram of joint point entropy

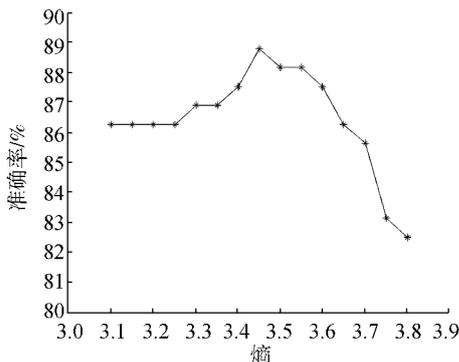


图 5 相同阈值准确率
Fig.5 The accuracy of same threshold

为了判断关节局部特征是否能够代表该行为,本文引入关节的信息熵。从图 2 可以清楚看到 (c) (d) (e) (f) (j) (k) 图中的行为中大部分的关节错误率都较高,而 (i) (o) (p) 图中行为的关节错误率都很低。如图 4 中所示,这几个行为的熵在错误率高的行为中,熵同样高,这样说明了关节表观特征在这些行为中,不确定性大,无法将其分到正确的类别。相反的,在其错误率较低的行为中,其熵普遍也比较低,可以使用关节表观特征代表该行为。因此,可以看出使用熵作为两个特征模型的判断标准是可行的。

同时,每个关节信息熵的阈值,对实验结果有着较大影响。采用统一阈值的方式虽然简单,但是并没有区分不同关节的差异,而每个关节采用不同的阈值手动调节的方法,参数较多。因此,本文采用平均信息熵作为阈值的自适应方法,不需要设置参数。

由图 5 可以看出,在设置相同的阈值中准确率最高为 88.75%,而采用本文提出的方法,在该数据集上准确率可以达到 91.25%,提高了 2.5%。

3.2 行为识别分析结果

本文使用了两种特征,一种是基于关节判别力的 HON4D 局部表观特征,另一种是使用关节相对距离变化来表达该数据集。根据高判别力关节的数量选择不同的特征用于行为识别。在实验中,数量阈值 C_p 从 1~20,结果如图 6 所示。

通过图 6 的对比可以看出,在 $C_p = 16$ 时,准确率是 91.25%,达到最高。分类的混淆矩阵如图 7~图 9 所示。

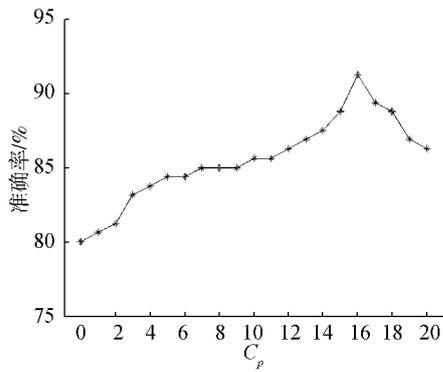


图 6 Cp 取值准确率

Fig.6 The accuracy of Cp's value

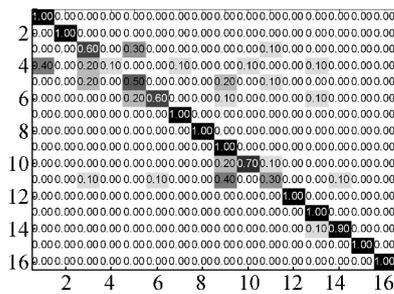


图 7 HON4D 特征的混淆矩阵

Fig.7 The confusion matrix of HON4D feature

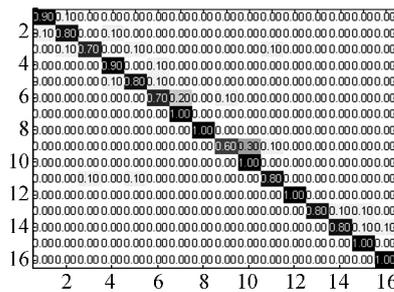


图 8 相对距离特征的混淆矩阵

Fig.8 The confusion matrix of the relative positions feature

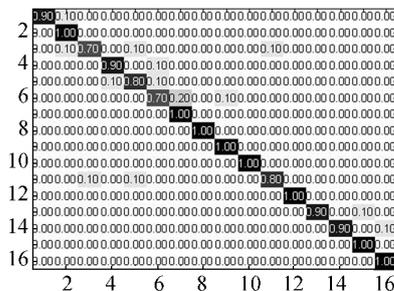


图 9 自适应特征选择方法的混淆矩阵

Fig.9 The confusion matrix of adaptive feature selection method

从混淆矩阵可以看出,在图 2 中(c) (d) (e) (f) (j) (k) 这几种行为上, HON4D 特征的分类效果较差,由此可以验证本文引入熵来作为判断该特征是

可以成立的。关节点相对距离特征在图 2(c) (d) (e) (f) (j) (k) 这些行为上较有优势。而本文提出的特征选择方法,将这两种特征的优势进行了充分的合并,结果有了很大的提高。

目前,基于该数据集的方法非常多。由表 1 可知,目前大部分使用直接串联的特征效果并不理想,使用本文中的两个特征时,整体的准确率只达到了 86.25%,由于特征关节点相对距离线性可分,直接导致结果受到关节点相对距离的影响,其中 HON4D 特征就无法发挥作用。

表 1 对比实验准确率

Table1 The accuracy of contrast experiment

分类	具体方法	准确率 / %
单特征	HOPC	81.7
	SNV	86.25
特征融合	Actionlet	86.75
	RGGP ^[17]	90.4
本文算法	HON4D	78.5
	HON4D(12 个关节点)	79.375
本文算法	关节点相对距离	86.25
	直接串联	86.25
	自适应特征选择法	91.25

4 总结

文中提出了特征自适应选择方法,此方法包括 4 个组成部分,即 HON4D 关节点模型、HON4D 关节点组合模型、关节点相对距离模型和自适应特征选择器。首先我们根据 HON4D 的关节点特征训练得到关节点随机森林的模型,根据平均熵结果为测试投票结果的阈值,以此来权衡 HON4D 关节点特征在此测试样本中是否存在较高的识别率。本文提出的方法在 MSR-Daily Activity 3D 这个数据集上评估。实验结果表明,本文的方法确实有效。

参考文献:

[1] WANG Jiang, LIU Zicheng, WU Ying, et al. Mining actionlet ensemble for action recognition with depth cameras [C]// Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, USA, 2012: 1290-1297.

[2] YANG Xiaodong, TIAN Yingli. Super normal vector for activity recognition using depth sequences [C]// Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition.

- Columbus, USA, 2014: 804–811.
- [3] CHEN Chen, JAFARI R, KEHTARNAVAZ N. Action recognition from depth sequences using depth motion maps-based local binary patterns [C]//Proceedings of 2015 IEEE Winter Conference on Applications of Computer Vision. Waikoloa, USA, 2015: 1092–1099.
- [4] XIA LU, CHEN C C, AGGARWAL J K. View invariant human action recognition using histograms of 3D joints [C]//Proceedings of 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, USA, 2012: 20–27.
- [5] LIU Jingen, ALI S, SHAH M. Recognizing human actions using multiple features [C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA, 2008: 1–8.
- [6] WANG Liang, ZHOU Hang, LOW S C, et al. Action recognition via multi-feature fusion and Gaussian process classification [C]//Proceedings of 2009 Workshop on Applications of Computer Vision. Snowbird, USA, 2009: 1–6.
- [7] LIU Jia, YANG Jie, ZHANG Yi, et al. Action recognition by multiple features and hyper-sphere multi-class SVM [C]//Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, Turkey, 2010: 3744–3747.
- [8] BENMOKHTAR R. Robust human action recognition scheme based on high-level feature fusion [J]. Multimedia tools and applications, 2014, 69(2): 253–275.
- [9] TRAN K, KAKADIARIS I A, SHAH S K. Fusion of human posture features for continuous action recognition [C]//Proceedings of the 11th European Conference on Trends and Topics in Computer Vision. Heraklion, Greece, 2010: 244–257.
- [10] OREIFEJ O, LIU Zicheng. HON4D: histogram of oriented 4D normals for activity recognition from depth sequences [C]//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 716–723.
- [11] YANG Xiaodong, TIAN Yingli. Effective 3D action recognition using EigenJoints [J]. Journal of visual communication and image representation, 2014, 25(1): 2–11.
- [12] RAHMANI H, MAHMOOD A, HUYNH D Q, et al. Real time action recognition using histograms of depth gradients and random decision forests [C]//Proceedings of 2014 IEEE Winter Conference on Applications of Computer Vision. Steamboat Springs, USA, 2014: 626–633.
- [13] YU Gang, LIU Zicheng, YUAN Junsong. Discriminative orderlet mining for real-time recognition of human-object interaction [M]//CREMERS D, REID I, SAITO H, et al. Computer Vision—ACCV 2014. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 50–65.
- [14] CHAARAOU A A, PADILLA-LOPEZ J R, FLOREZ-REVUELTA F. Fusion of skeletal and silhouette-based features for human action recognition with RGB-D devices [C]//Proceedings of 2013 IEEE International Conference on Computer Vision Workshops. Sydney, Australia, 2013: 91–97.
- [15] GAO Zan, ZHANG Hua, LIU A A, et al. Human action recognition on depth dataset [J]. Neural computing and applications, 2016, 27(7): 2047–2054.
- [16] LIU Zhi, ZHANG Chenyang, TIAN Yingli. 3D-based deep convolutional neural network for action recognition with depth sequences [J]. Image and vision computing, 2016, 55(2): 93–100.
- [17] LI Meng, LEUNG H, SHUM H P H. Human action recognition via skeletal and depth based feature fusion [C]//Proceedings of the 9th International Conference on Motion in Games. Burlingame, USA, 2016: 123–132.

作者简介:



龚冬颖,女,1992年生,硕士研究生,主要研究方向为行为识别、机器学习。



黄敬,女,1982年生,博士研究生,主要研究方向为行为识别、机器学习、目标检测和图像检索。



张洪博,男,1986年生,讲师,博士,主要研究方向为人体行为识别,主持国家自然科学基金青年项目和福建省自然科学基金面上项目各1项,发表学术论文多篇,其中被SCI、EI检索20余篇。