

基于多学习器协同训练模型的人体行为识别方法*

唐超¹, 王文剑², 李伟³, 李国斌¹, 曹峰²

¹(合肥学院 计算机科学与技术系, 安徽 合肥 230601)

²(山西大学 计算机与信息技术学院, 山西 太原 030006)

³(厦门理工学院 计算机与信息工程学院, 福建 厦门 361005)

通讯作者: 唐超, E-mail: tangchao77@sina.com, http://www.hfuu.edu.cn

摘要: 人体行为识别是计算机视觉研究的热点问题, 现有的行为识别方法都是基于监督学习框架。为了取得较好的识别效果, 通常需要大量的有标记样本来建模。然而, 获取有标记样本是一个费时又费力的工作。为了解决这个问题, 对半监督学习中的协同训练算法进行改进, 提出了一种基于多学习器协同训练模型的人体行为识别方法。这是一种基于半监督学习框架的识别算法。该方法首先通过基于 Q 统计量的学习器差异性度量选择算法来挑取出协同训练中基学习器集, 在协同训练过程中, 这些基学习器集对未标记样本进行标记; 然后, 采用了基于分类器成员委员会的标记近邻置信度计算公式来评估未标记样本的置信度, 选取一定比例置信度较高的未标记样本加入到已标记的训练样本集并更新学习器来提升模型的泛化能力。为了评估算法的有效性, 采用混合特征来表征人体行为, 从而可以快速完成识别过程。实验结果表明, 所提出的基于半监督学习的行为识别系统可以有效地辨识视频中的人体动作。

关键词: 人体行为识别; 半监督学习; 协同训练; 学习器选择; 混合特征

中图法分类号: TP391

中文引用格式: 唐超, 王文剑, 李伟, 李国斌, 曹峰. 基于多学习器协同训练模型的人体行为识别方法. 软件学报, 2015, 26(11): 2939-2950. <http://www.jos.org.cn/1000-9825/4899.htm>

英文引用格式: Tang C, Wang WJ, Li W, Li G, Cao F. Multi-Learner co-training model for human action recognition. Ruan Jian Xue Bao/Journal of Software, 2015, 26(11): 2939-2950 (in Chinese). <http://www.jos.org.cn/1000-9825/4899.htm>

Multi-Learner Co-Training Model for Human Action Recognition

TANG Chao¹, WANG Wen-Jian², LI Wei³, LI Guo-Bin¹, CAO Feng²

¹(Department of Computer Science and Technology, Hefei University, Hefei 230601, China)

²(School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)

³(School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361005, China)

Abstract: Human action recognition is a hot topic in computer vision. Most of the existing work use the action models based on supervised learning algorithms. To achieve good performance on recognition, a large amount of labeled samples are required to train the sophisticated action models. However, collecting labeled samples is labor-intensive. This paper presents a novel semi-supervised learning algorithm named multi-learner co-training model (MCM) to recognize human actions. Two key issues are addressed in this paper. Firstly, the base classifiers in co-training are selected by Q statistic-based classifiers selection algorithm (QSCSA). Secondly, MCM is employed for the semi-supervised model construction. The new confidence score measure of unlabeled sample depends on estimating the classifier companion committee (CCC) accuracy on labeling the neighborhood of an unlabeled examples. To evaluate the proposed algorithm, mixed-descriptors are used to express actions so that the recognition algorithm can quickly complete the recognition process from a single

* 基金项目: 国家自然科学基金(61273291, 71031006, 41401521); 山西省回国留学人员科研资助项目(2012-008); 安徽高校自然科学基金研究项目(KJ2015A206); 合肥学院重点建设学科项目(2014xk08); 合肥学院学科带头人培养对象项目(2014dtr08); 合肥学院人才科研基金项目(15RC07); 厦门理工学院高层次人才项目(YKJ14014R)

收稿时间: 2015-05-29; 修改时间: 2015-07-14; 定稿时间: 2015-08-26

frame of visual image. Experimental results are presented to show that the proposed semi-supervised learning system can recognize simple human actions effectively.

Key words: human action recognition; semi-supervised learning; co-training; learner selection; mixed feature

人体行为识别是计算机视觉的研究热点问题,它在基于内容的视频检索、人机交互、运动视频分析和视频监控等方面具有广泛的应用价值.然而,由于人体动作的动态性造成动作表征困难以及由于视角、光照和遮挡等问题,都使得行为识别的研究成为难题.

传统的行为识别算法通常需要建立一个模型,而建模的过程中需要大量有标记的训练样本.要获得这些标记样本需要大量人工劳动,费时、费力,而无标记的训练样本的获取相对十分容易.目前,探索未标记数据来帮助提升整个学习系统的学习成效已成为一个热门的研究方向.在这种实际应用背景下,如何利用有标记样本,更重要的是充分利用大量未标记样本来提升整个识别系统的性能,成为一个关键问题.为了解决这个问题,我们提出了一种基于多学习器协同训练模型(multi-learner co-training model,简称 MCM)的人体行为识别方法,这是一种基于半监督学习(semi-supervised learning,简称 SSL)^[1]框架的识别算法.该方法首先通过基于 Q 统计量的学习器选择算法(Q statistic-based classifiers selection algorithm,简称 QSCSA)挑取出协同训练中基学习器集,在学习过程中,这些基学习器集对未标记样本进行标记;然后,采用基于分类器成员委员会(Classifier Companion Committee,简称 CCC)的标记近邻置信度计算公式来评估未标记样本的置信度,选取一定比例置信度较高的未标记样本加入到已标记的训练样本集并更新学习器来提升整个模型的泛化能力.经过迭代后,最后对学习器集融合进行决策输出.

相对于目前的识别算法,本文提出的识别算法具有以下优点:

- (1) 本文提出的基于半监督学习的行为识别方法可以有效利用少量有标记样本和大量未标记样本学习得到泛化能力较强的分类模型,满足了一些实际的应用要求.
- (2) 不同于标准的协同训练算法,MCM 算法不需要标准协同训练算法中两个独立冗余视图的要求,这更符合实际的行为识别问题.
- (3) 基于分类器成员委员会的标记近邻的置信度计算公式来评估未标记样本的置信度,增加了协同训练中基学习器之间的差异性,从而可以更好地提高学习器相互学习能力.
- (4) 通过理论分析和实验,表明基于多学习器协同训练算法的有效性.

1 相关工作介绍

关于半监督学习已涌现出许多方法,并且都取得了较好的学习效果.半监督学习主要分为以下几类:自训练算法(self-training)^[2,3]、基于图的半监督学习方法(SSL with graphs)^[4,5]、基于生成式模型的半监督学习方法(SSL with generative models)^[6-9]、半监督支持向量机(semi-supervised support vector machines,简称 S3VMs)^[10,11]和基于视图的半监督学习方法(SSL with views)^[12]等.

目前,将半监督学习应用于行为识别领域的研究还不是很多,我们可以很容易地从公共视频中获取大量未标记的样本,这为开展这方面的研究提供了可能.Guan 等人^[13]最早开始了这方面的工作,他们提出了 En-Co-Training 方法.该方法使用 3 种不同的学习器,分别是决策树、朴素的贝叶斯和 K 近邻学习器,作为 3 种不同的视图.实验表明,该方法可以有效地利用未标记样本提升学习系统的泛化能力.但是该方法没有考虑到视图的独立性和冗余性,同时,实验也没有与目前的最优方法进行比较.Gupta 等人^[14]提出了基于 Co-training 的行为识别方法.他们从视频数据中提取视觉和纹理两种特征作为视图,取得了较好的学习效果.Zhang 等人^[15]提出了基于 Co-EM 算法半监督学习算法,以对行为进行分类.为了克服特征高维特性,WMDA(weighted multiple discriminant analysis)技术用来将高维数据映射到低维子空间中,同时,将 Gaussian Mixture Models(GMM)用来整合子空间模型.为了解决标准协同训练中对视图独立性与冗余性的要求,Liu 等人^[16]提出了两种置信度评价方法:Inter-view 置信度和 Intra-view 置信度,分别动态地融入到半监督学习过程中,并采用互信息来量化 Inter-

view 不确定性,同时也可以用来度量不同视图间的独立性.Intra-view 置信度可以用来度量整个有标记样本和未标记样本之间的整体不一致性.Wang 等人^[17]提出了一种新的基于多视觉特征的半监督学习算法,用于行为识别.该方法可以从少量有标记训练视频中同时提取出多种不同的特征,并可以利用标记样本和未标记样本的分布特性来提升学习器的泛化能力.

2 系统框架

本节将介绍基于 MCM 的行为识别系统.该系统采用 Kinect^[18]作为系统传感器,图 1 为整个行为识别系统的示意图.系统可以通过示教者或是用户采集有标记的人体动作帧,值得注意的是,其中有标记样本的类别是通过 Kinect 麦克风,经自动语音识别系统自动添加上的.同时,系统还可以得到大量未标记的视频动作帧,通过收集到的少量有标记样本和大量未标记样本,经过系统的核心部分半监督学习模块,从而得到分类模型.

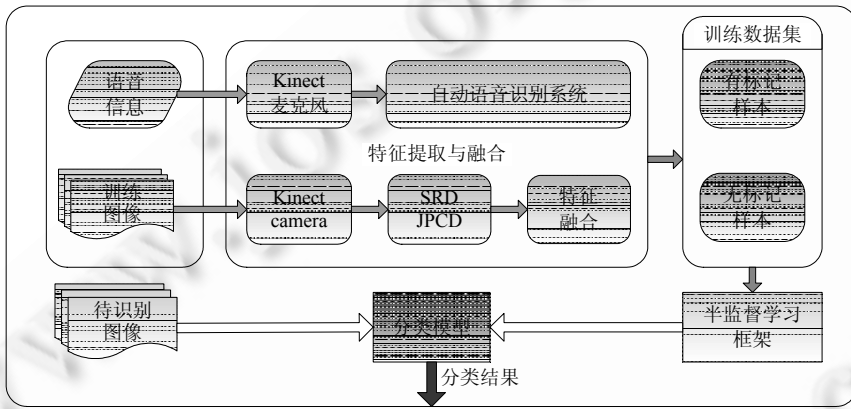


Fig.1 Overview of the action recognition system

图 1 行为识别系统示意图

图 2 给出行为识别系统中半监督学习算法框架.训练数据集中包括两部分数据:一部分是少量有标记的训练样本,一部分是大量的无标记样本.算法在此数据集学习产生分类模型.该方法首先通过基于 Q 统计量的学习器差异性度量选择算法来从初始学习器集合中挑取 M 个学习器,作为协同训练中基学习器集.初始学习器采用支持向量机,并通过核函数与核参数的不同设置来训练始初学习器集合.在协同训练阶段,采用了基于成员分类器的标记近邻的置信度计算公式来评估未标记样本的标记置信度.添加到每一个成员分类器的有标记训练集中的未标记样本,是由除了这个分类器以外的其他成员分类器进行评估.在预测阶段,是通过对所有的成员分类器进行集成进行决策输出.

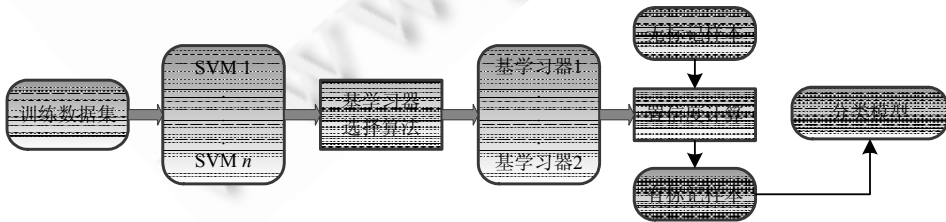


Fig.2 Multi-Learner co-training model

图 2 多学习器协同训练模型

3 多学习器协同训练模型

3.1 基于 Q 统计量的学习器选择算法

为了构造 MCM 模型,首先生成一组基学习器,这些基学习器是通过初始学习器筛选得到的.初始学习器采用支持向量机,并通过采用不同的核函数和核参数的不同生成.那么如何来筛选这些初始学习器呢?这就需要考虑到学习器之间差异性,如果两个学习器是一样的,对于协同训练来说则是无意义的.在这里,引用了集成学习中对于学习器差异性的描述.目前,对于学习器差异性的标准定义,还没有一个统一的说法.在文献[19]中,学习器之间的差异性被定义为:对于新的数据样本,各个分类器做出不同错误的趋势.

假设 h_i 和 h_j 为两个不同的学习器,那么它们的关系可用表 1 的方法来表示.

Table 1 Learner relational table of classification of the samples

表 1 学习器分类样本关系表

	h_i correct(1)	h_j wrong(0)
h_i correct(1)	N^{11}	N^{10}
h_i wrong(0)	N^{01}	N^{00}

表 1 中, N^{11} 代表学习器 h_i 和 h_j 同时分类正确的样本数量, N^{10} 表示 h_j 分类正确、 h_i 分类错误的样本数量,其余以此类推.本文使用 Q 统计量(the Q statistic)^[20]作为学习器差异性的度量方法:

$$Q_{ij}(h_i, h_j) = \frac{N_{11}N_{00} - N_{01}N_{10}}{N_{11}N_{00} + N_{01}N_{10}} \quad (1)$$

Q 的取值范围为 $[-1, 1]$,当 $Q=0$ 时,说明两个学习器独立,差异性最大.

在此基础上提出了基于 Q 统计量的学习器选择算法 QSCSA,用以构造协同训练中基学习器.成员基分类器之间的差异性对协同训练系统的性能非常重要,同时,成员基分类器的平均准确率也很重要.为此,本文设计一种 QSCSA 算法,能够取得两者之间较好的平衡,挑选出的成员基分类器既有较高的精度,同时又能保持较大的差异度.其具体算法如算法 1 所示.

算法 1. 基于 Q 统计的学习器选择算法.

输入:训练集 S ,测试集 T ,初始分类器集 $\{h_1, h_2, \dots, h_N\}$,迭代次数 M .

过程:

1. **for** $i, j=1, 2, \dots, N$

2. 使用测试集来测试分类器 h_i 的分类结果识别率 p_i ,识别率定义为 $p_i = \frac{TP}{TP + FP}$

3. $p_i \leftarrow \text{Reg_rate}(T, h_i)$

4. **end for**

6. 按照识别率 p_i 的大小,对 $h = \{h_1, h_2, \dots, h_M\}$ 进行排序

7. 挑选出排在前 K 的分类器 $A = \{h'_1, h'_2, \dots, h'_N\}, 1 < K < N$

8. **for each** $h'_i, h'_j \in A$

9. 计算两个学习器之间的差异性: $Q_{i,j}(h_i, h_j) = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}}$

10. **end for**

11. 对计算出来的所有学习器的差异性值 $\{Q_{i,j}^{(1)}, Q_{i,j}^{(2)}, \dots, Q_{i,j}^{((N-1) \times (N-1))}\}$ 进行排序,选出所有差异性最大的 M 个分类器

12. 输出: M 个基分类器集合 $H = \{BL_1, BL_2, \dots, BL_M\}$

3.2 MCM方法

在使用 QSCSA 选择出的基分类器集 $H = \{BL_1, BL_2, \dots, BL_M\}$ 来标记未标记样本后,需要进一步评估其置信度.

为了给每一个成员学习器的训练集中加入未标记训练样本,本文给出了基于成员学习器集的标记近邻置信度计算公式来挑选置信度较高的未标记样本,将其加入到每一个学习器的训练集中.加入到每一个成员分类器 BL_i 的训练集 L_i 中的未标记训练样本是通过 BL_m 成员分类器集 $H_m = \{BL_1, BL_2, \dots, BL_{M-1} | BL_M \in H\}$ 进行评估的,其中, H_m 是不包括 BL_m 的其他分类器.

给未标记样本进行分类时,通常遵守半监督学习中的聚类假设^[1,21]:

- (1) 数据来自簇中;
- (2) 同一簇中的样本,更可能有相同的类别标记.

根据这个假设,未标记样本 x_i^u 的置信度可以通过计算与它最相似的 K 个标记近邻的样本的分类一致性来度量.

令 $S(x_i^u, x_j^l)$ 表示未标记样本 x_i^u 与它的标记近邻 x_j^l 的相似度,这里, $S(x_i^u, x_j^l)$ 采用余弦相似性度量(cosine similarity).

给定两个样本 x_i 和 x_j ,则余弦相似性度量可以定义为

$$S(x_i, x_j) = \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \quad (2)$$

加入到成员分类器 BL_m 的训练集 L_m 的未标记样本 x_i^u , 被成员分类器集 H_m 标记为某个类 C , 的置信度,由 x_i^u 与它的 K 个标记近邻的标记一致性决定,用 $Conf(x_i^u)$ 表示,计算如下:

$$Conf(x_i^u) = \sum_{\substack{q=1 \\ q \neq m}}^{M-1} \sum_j^k S(x_i^u, x_j^l) \times constistecny(BL_q(x_i^u), y_{x_j^l}) \quad (3)$$

其中,

$$constistecny(BL_m(x_i^u), y_{x_j^l}) = \begin{cases} 1, & BL_m(x_i^u) = y_{x_j^l}, x_j^l \in \phi_i^l \\ -1, & BL_m(x_i^u) \neq y_{x_j^l} \end{cases} \quad (4)$$

其中,如果 BL_m 对 x_i^u 和 x_j^l 分类相同,那么 $constistecny(BL_q(x_i^u), y_{x_j^l}) = 1$; 否则, $constistecny(BL_q(x_i^u), y_{x_j^l}) = -1$. 其中, $Conf(x_i^u)$ 越大,表明样本 x_i^u 与最相似的 K 个已标记近邻的分类结果越一致, x_i^u 的置信度就越高;反之, $Conf(x_i^u)$ 越小,表明样本 x_i^u 与最相似的 K 个已标记近邻的分类结果越不一致, x_i^u 的置信度就越低. x_i^u 标记由成员分类器集 H_m 多数投票给出:

$$\hat{y}_u = \arg \max_{1 \leq c \leq C} H_m(x_u) \quad (5)$$

具体 MCM 算法如下:

算法 2. 多分类器协同学习模型算法.

输入:有标记训练集 L , 无标记训练集 U , 迭代次数 T , 集成规模 M , 类别数 C , 相似度计算 *Similarity*, 欧式距离测度 *dis*, 近邻算法 *Neighbors*, 近邻数 k , 基学习器算法 *SVM*, 核函数 *kernel*, 核参数 *para*.

训练部分:

1. **for** $i=1, 2, \dots, N$
2. 构建初始学习器集合
3. $h_i \leftarrow SVM(L, kernel_i, para_i)$
4. **end for**
5. 对学习器进行筛选
6. $\{BL_1, BL_2, \dots, BL_M\} \leftarrow CSAQ\{h_1, h_2, \dots, h_N\}$
7. $L_i = L (i=1, \dots, M)$
8. **for** $t=1, 2, \dots, T$

9. **for** $m=1,2,\dots,M$
 10. 从无标记训练集 U 中随机抽取出一个大小为 s 的子集 $U'_{t,m}$
 11. **for** 每一个 $x_i^u \in U'_{t,m}$ **do**
 12. $\Phi_i^l \leftarrow \text{Neighbors}(x_i^u, L, k, dis_i), s(x_i^u, x_j^l) \leftarrow \text{Similarity}(x_i^u, x_j^l, dis)$
其中, $x_j^l \in \Phi_i^l$
 13. **end for**
 14. 计算置信度 $\text{Conf}(x_i^u)$
 15. $\text{Conf}(x_i^u) \leftarrow \sum_{\substack{q=1 \\ q \neq m}}^{M-1} \sum_j^k S(x_i^u, x_j^q) * \text{constistecny}(BL_m(x_i^u), y_{x_j^q})$
其中, $\text{constistecny}(BL_m(x_i^u), y_{x_j^q}) = \begin{cases} 1, & BL_m(x_i^u) = y_{x_j^q} \\ -1, & BL_m(x_i^u) \neq y_{x_j^q} \end{cases}, x_j^q \in \Phi_i^l$
 16. 按照 $\text{Conf}(x_i^u)$ 的大小排序 $\{x_{i_1}^u, x_{i_2}^u, \dots, x_{i_s}^u\} \leftarrow \{x_i^u \mid x_i^u \in U'_t\}$
 17. 挑取 top_N 个 x_i^u 和它的预测标记 $\hat{y}_i^u = \arg \max_{1 \leq c \leq C} H_m(x_u)$ 保存在训练集 $L'_{t,m}$, 同时, 从 $U'_{t,m}$ 删除 x_i^u
 18. 更新有标记训练集 $L_m = L_m + L'_{t,m}$
 19. 更新学习器 $BL_m \leftarrow \text{SVM}(L_m, \text{kernel_type}, \text{kernel_parameters})$
 20. 随机挑选从 U 个 top_N 个样本更新子集 $U'_{t,m}$
 21. **end for**
 22. **end for**
- 预测部分:
23. **return** $BL^*(x) = \frac{1}{M} \sum_{i=1}^M BL_i(x)$

4 实验结果与分析

4.1 数据集

为了验证本文提出的基于多分类器协同训练模型算法识别的效果,我们首先构建了基于 Kinect 的室内行为识别数据库.如图 3 所示,该数据库包括 9 个表演者执行 10 种简单动作,包括的动作有 walk,sit,side walk,run,pick up,jump,hand wave,hand clap,box 和 bend.在此数据集上,分别提取两种行为特征描述子来表征动作,分别为基于区域的描述子和基于关节的极坐标的描述子.

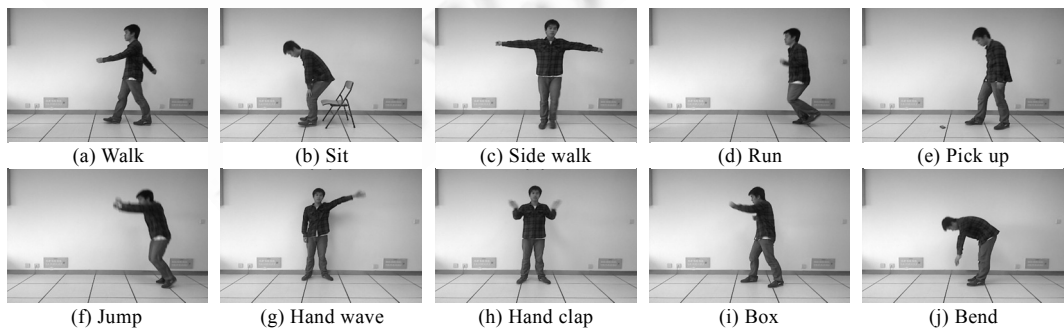


Fig.3 Ten sample actions from Kinect human action classification database

图 3 Kinect 人体动作分类数据集中十个样例动作

4.1.1 人体区域统计量特征

基于区域描述子对人体区域的几何属性进行特征化.该方法将人体区域形状看成一个整体,这样可以有效地利用区域内的所有像素,因而受噪声和形状变化的影响较小.基于区域的特征表示方法可以分为全局与局部两种.为了得到有效的人体区域,需要对人体二值图像序列进行图像连通预处理:首先,采用八向连通填充算法对人体二值图像序列进行图像连通处理,主要是填充图像中的空洞,使得人体行为二值图像中只有一块像素为1的区域.使用面积(area)、周长(perimeter)、包含区域的最小矩形、圆形成度(roundness)、形状复杂性(shape complexity)、重心坐标(center of gravity)、偏心率(eccentricity)、球状性(sphericity)和圆形性这9种不同的形状参数构成人体区统计特征向量来描述人体模式,这9个统计量信息特征进行规一化后得到基于形状的区域特征(shape-based region descriptors,简称SRD),用来描述人体行为特征,即

$$SRD=\{SRD_1,SRD_2,\dots,SRD_9\} \quad (6)$$

4.1.2 人体关节极坐标特征

通过实验发现,人类视觉系统可以在很短的时间内通过运动光照显示设备判别出人体动作.基于这样的视觉认知过程,可以将人体重要骨骼关节点放在一个坐标系下来进行描述,以此作为动作的表征形式.因此,采用微软公司的Kinect室内深度图像传感器来实时捕捉人体重要部位和关节点坐标,并采用极坐标的形式来描述这些重要部分与关节点的坐标位置.这些重要部位和关节有11个,从上到下依次分别是:人体躯干中心点、头部中心点、左肩关节点、左边肘部关节点、左手腕关节点、右肩关节点、右边肘部关节点、右手腕节关节点、左膝盖关节点、左踝关节点、右膝盖关节点和右踝关节点.采用人体中心点作为极点 P ,这样,其余10个关节点在图像坐标系下的坐标就可以转换为极坐标表示形式,如公式(7)所示:

$$(r_i, \phi_i) = \begin{cases} r_i = \sqrt{(y_i - y_c)^2 + (x_i - x_c)^2} \\ \phi_i = \arctan\left(\frac{y_i - y_c}{x_i - x_c}\right) \end{cases}, i = 1, \dots, 11 \quad (7)$$

其中, (x_c, y_c) 和 (x_i, y_i) 分别是人体中心点和其他关节点在图像直角坐标系下的坐标,如图4所示.

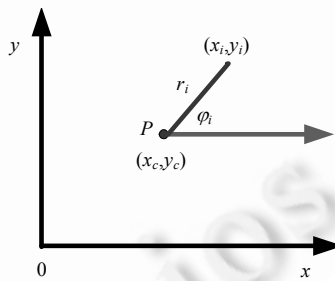


Fig.4 Coordinate transformation of human skeleton joints

图4 人体骨骼关节点坐标变换

在得到关节点极坐标后,需要进一步对其进行归一化处理.为了保存数据落在区间[0,1]之间,采用极小极大法,具体如公式(8)和公式(9)所示.

$$r'_i = \frac{r_i - \min_r}{\max_r - \min_r} \quad (8)$$

$$\phi'_i = \frac{\phi_i - \min_phi}{\max_phi - \min_phi} \quad (9)$$

其中, \min_r 和 \max_r 分别是原始 r 中的最小值和最大值; \min_phi 和 \max_phi 分别是原始 phi 中的最小值和最大值.经过规一化处理后,就可以得到具有平移变换、尺度变换和旋转变换不变性的基于关节点的行为描述子(joints-based polar coordinates descriptors,简称JPCD),即

$$JPCD=\{JPCD_1,JPCD_2,\dots,JPCD_{22}\} \quad (10)$$

4.2 方法评估

采用 10 折交叉验证的方法评估本文提出的半监督学习算法与监督学习算法的比较.图 5 给出混合特征 JPCD+SRD 下的监督学习与半监督学习的识别结果混淆矩阵,其中,第(i,j)元素表示第 i 类行为被分类为第 j 行为的比例.从图 5 中可以看到,在 JPCD+SRD 的数据集上,本文的方法取得了比传统的监督学习算法更高的识别率.从而表明,可以有效利用无标记样本提升学习器性能,达到较好的识别效果.

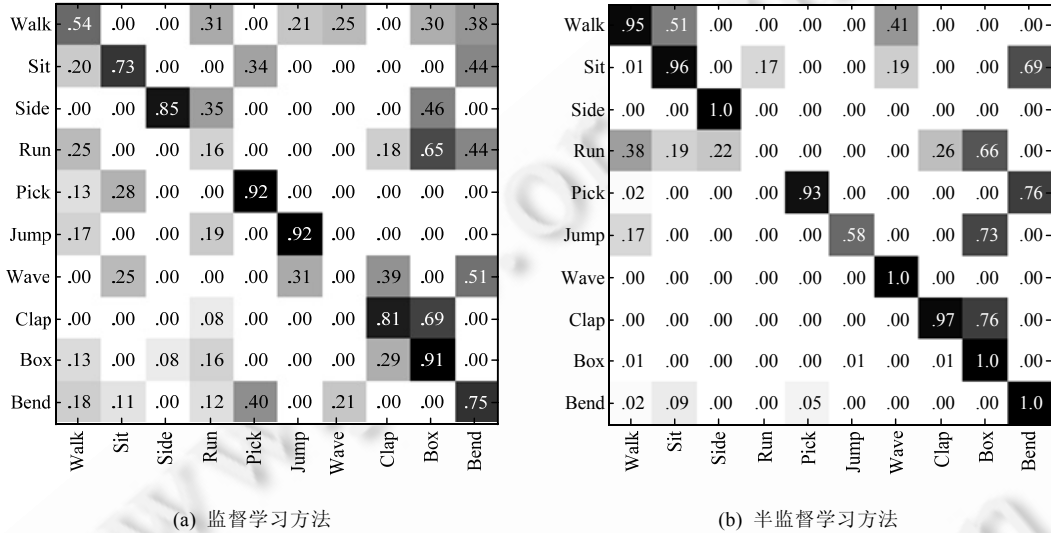


Fig.5 Confusion matrix of the proposed action recognition using the mixed features SRD+JPCD

图 5 行为识别混淆矩阵,采用 JPCD+SRD 特征数据集

图 6 和图 7 给出了单特征数据集下的实验结果.从中我们可以发现,半监督学习方法都取得优于监督学习算法的识别率.

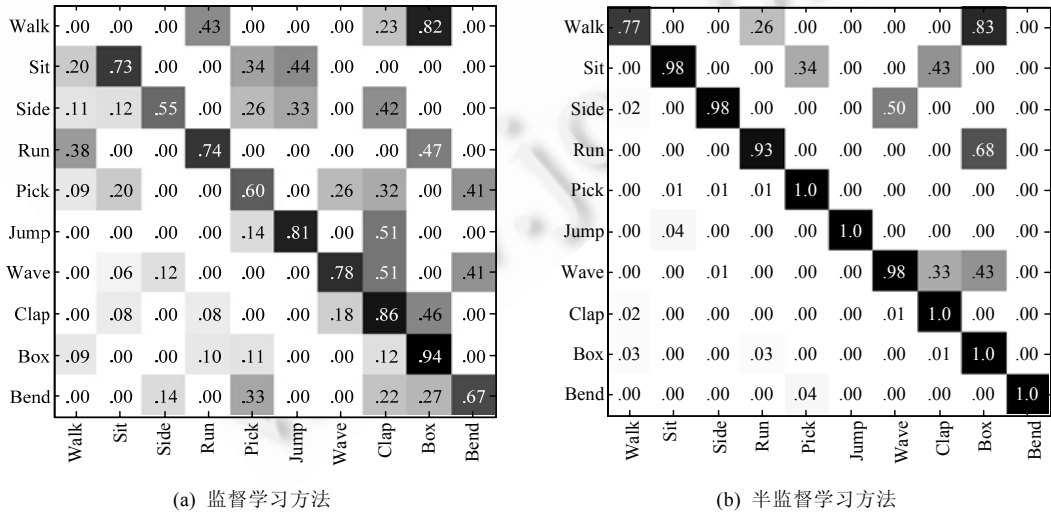


Fig.6 Confusion matrix of the proposed action recognition using the features SRD

图 6 行为识别混淆矩阵,采用 SRD 特征数据集

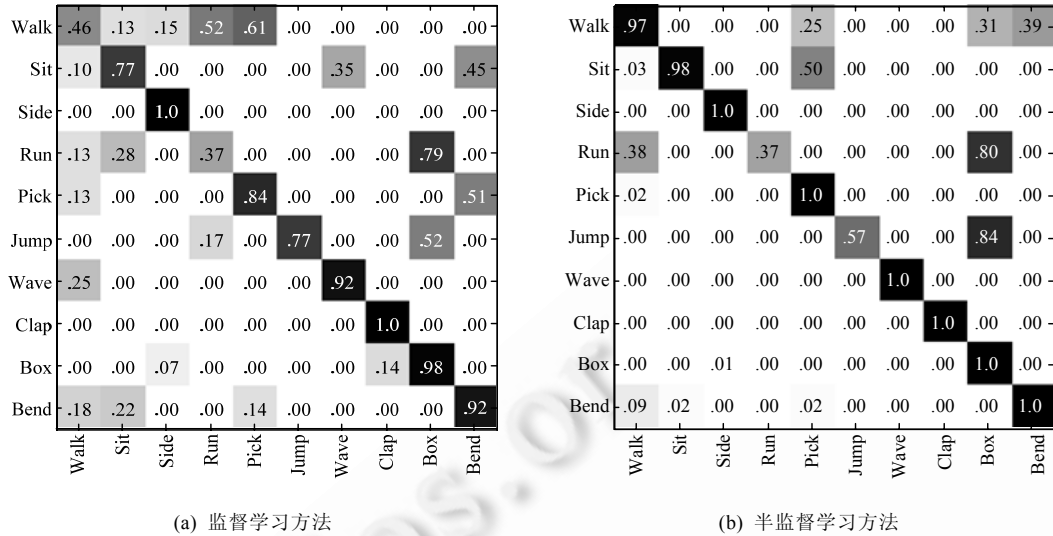


Fig.7 Confusion matrix of the proposed action recognition using features JPCD

图 7 行为识别混淆矩阵,采用 JPCD 特征数据集

表 2 是不同特征和特征组合数据集下,监督学习算法与本文的半监督学习算法的比较.其中,LABELED 代表监督学习方法,MCM 代表本文提出的算法.从中我们可以看到,在 3 个评测标准下,MCM 算法都取得了比较好的学习效果.

Table 2 Recognition accuracy of supervised learning & semi-supervised learning strategies

表 2 监督学习与半监督学习策略下的识别精度

特征描述子	Precision		Recall		F-Measure	
	LABELED	MCM	LABELED	MCM	LABELED	MCM
JPCD+SRD	0.53	0.93	0.53	0.94	0.52	0.94
SRD	0.53	0.96	0.52	0.96	0.52	0.96
JPCD	0.70	0.97	0.71	0.97	0.70	0.96
Avg.	0.62	0.94	0.62	0.95	0.62	0.95

图 8 和图 9 分别给出了在不同特征数据集下,无标记训练样本个数以及有标记训练样本个数对学习器精度的影响.从图 8 中可以看到,在初始有标记训练样本为 100,700 和 1 300 的情况下,伴随着无标记训练样本的增加,学习器的精度均得到了提升;其中,当有标记样本为 100 时,学习器的精度提升得最大.因此对于协同训练来说,选择合适的有标记和无标记样本的比例很重要.

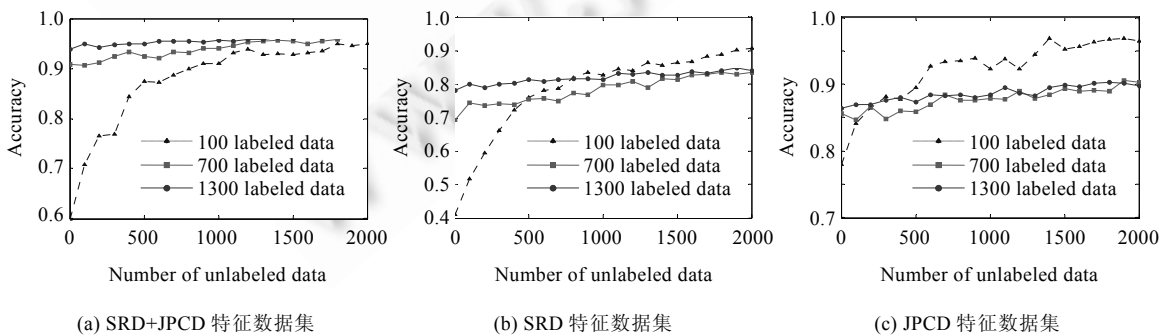


Fig.8 Evaluation of our MCM method with different number of unlabeled data

图 8 不同的无标记样本个数对 MCM 的影响

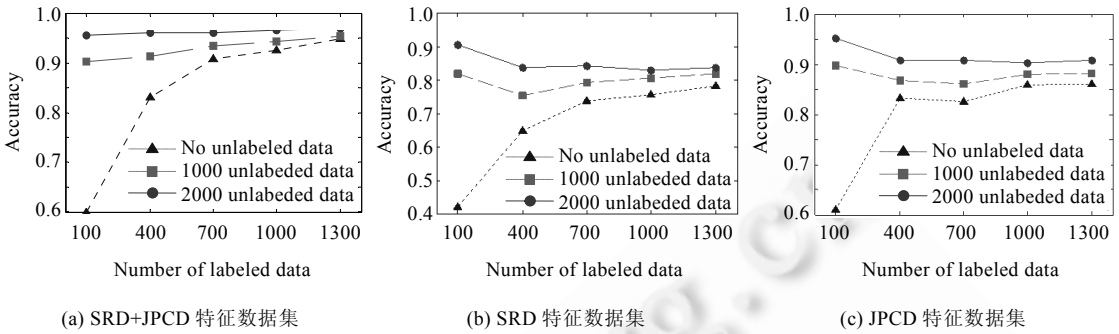


Fig.9 Evaluation of our MCM method with different number of labeled data
图 9 不同的有标记样本个数对 MCM 的影响

4.3 与其余方法的对比

表 3 给出了 MCM 方法与其他一些监督学习方法的比较.这是一个相对的比较方法:第一,由于每种方法的具体实验设置参数均不同;第二,各种方法使用行为特征也不相同.总体上说,本文方法取得了一个较好的识别结果.

Table 3 Comparison of the action recognition performance (%) between the proposed method and the other approaches

表 3 本文方法与其他方法的动作识别精度(%)比较

方法 识别率(%)	Dollar, et al. ^[22]	Niebles, et al. ^[23]	Schuldt, et al. ^[24]	Liu, et al. ^[16]	Ours
	81.2	83.3	71.7	93.8	97.0

为了进一步评估本文方法与其他半监督学习算法,我们设置了 3 种方法与其进行对比.

- SELF1 是一种自训练的半监督学习算法,采用了 SVM 作为学习器.
- SELF2 也是一种自训练的半监督学习算法,采用了 RF 作为学习器.
- FAKE-CO 是一种协同训练算法.该算法采用随机的方法对数据特征进行分割,产生两个视图;然后,应用标准的协同训练方法进行训练.

从表 4 可以看到,本文方法在各个特征数据集上都取得了优于另外 3 种方法的学习效果.图 10 是不同迭代次数下,MCM 方法与其他 3 种半监督方法的比较.

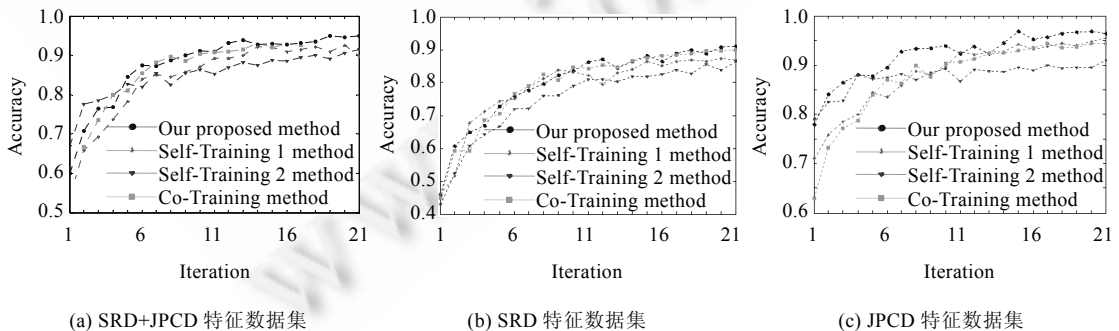


Fig.10 Comparative results for our proposed method MCM with the other semi-supervised learning method in different iterations

图 10 不同迭代次数下,MCM 与其他半监督学习方法的比较结果

Table 4 Comparative results for MCM with the other semi-supervised learning method (5% labeled rate)**表 4** MCM 方法与其他半监督学习方法的比较结果(有标记率 5%)

特征数据集	LABELED	MCM	SELF1	SELF2	FAKE-CO
JPCD+SRD	0.60	0.95	0.89	0.91	0.95
SRD	0.46	0.91	0.87	0.86	0.90
JPCD	0.78	0.96	0.95	0.91	0.94

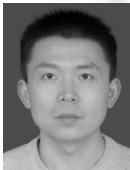
5 结 论

本文提出了一种基于多学习器协同训练模型的人体行为识别方法 MCM 分类模型.这是一种基于半监督学习框架的识别算法.该方法通过基于 Q 统计量的学习器差异性度量选择算法,挑取一组学习器作为协同训练中基学习器,在学习过程中,这组学习器对未标记样本进行标记;然后,采用基于多学习器的标记近邻置信度计算公式来进一步评估这些样本的置信度,选取一定比例置信度较高的未标记样本加入到已标记的训练样本集.实验结果表明,MCM 模型有效地利用大量无标记训练样本提升了整个行为识别系统的泛化能力,弥补了以往基于监督学习系统的行为分类模型的不足.在未来的工作中,进一步优化 MCM 模型中基分类器的权重是一个值得研究的问题.

References:

- [1] Chapelle O, Schölkopf B, Zien A. *Semi-Supervised Learning*. Cambridge: MIT Press, 2006. 1–12.
- [2] Nigam K, Ghani R. Analyzing the effectiveness and applicability of co-training. In: Proc. of the 9th IEEE Int'l Conf. on Information and Knowledge Management. ACM Press, 2000. 86–93. [doi: 10.1145/354756.354805]
- [3] Inoue M, Ueda N. Exploitation of unlabeled sequences in hidden Markov models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2003,25(12):1570–1581. [doi: 10.1109/tpami.2003.1251150]
- [4] Blum A, Chawla S. Learning from labeled and unlabeled data using graph mincuts. In: Proc. of the 18th IEEE Int'l Conf. on Machine Learning. Morgan Kaufmann Publishers, 2001. 19–26. [doi: 10.1.1.19.3957]
- [5] Zhu X, Ghahramani Z, Lafferty J. Semi-Supervised learning using Gaussian fields and harmonic functions. In: Proc. of the 20th IEEE Int'l Conf. on Machine Learning. Washington: AAAI Press, 2003. 912–919. [doi: 10.1.1.5.68]
- [6] Belkin M, Niyogi P. Semi-Supervised learning on riemannian manifolds. *Machine Learning*, 2004,56(1-3):209–239. [doi: 10.1023/b:mach.0000033120.25363.1e]
- [7] Zhou D, Bousquet O, Lal TN, Weston J, Schölkopf B. Learning with local and global consistency. In: Proc. of the Advances in Neural Information Processing Systems 16. Cambridge: MIT Press, 2004. 321–328.
- [8] Shahshahani BM, Landgrebe DA. The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. *IEEE Trans. on Geoscience and Remote Sensing*, 1994,32(5):1087–1095. [doi: 10.1109/36.312897]
- [9] Nigam K, McCallum AK, Thrun S, Mitchell T. Text classification from labeled and unlabeled documents using EM. *Machine Learning*, 2000,39(2-3):103–134. [doi: 10.1023/A:1007692713085]
- [10] Gao J, Xiang L. Local preserving projections and within-class scatter based semi-supervised support vector machines. In: Proc. of the 3th IEEE Int'l Conf. on Computer Science and Information Technology. IEEE Press, 2010. 267–270. [doi: 10.1109/iccscit.2010.5563589]
- [11] Zhu DZ, Wang X, Chen H, Wu R. Semi-Supervised support vector machines regression. In: Proc. of the 9th IEEE Int'l Conf. on Industrial Electronics and Applications. IEEE Press, 2014. 2015–2018. [doi: 10.1109/iciea.2014.6931500]
- [12] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training. In: Proc. of the 11th Annual Conf. on Computational Learning Theory. ACM Press, 1998. 92–100. [doi: 10.1145/279943.279962]
- [13] Guan DH, Yuan WW, Lee YK, Gavrilov A, Lee SY. Activity recognition based on semi-supervised learning. In: Proc. of the 13th IEEE Int'l Conf. on Embedded and Real-Time Computing Systems and Applications. IEEE Press, 2007. 469–475. [doi: 10.1109/rtsa.2007.17]

- [14] Gupta S, Kim J, Grauman K, Mooney R. Watch, listen & learn: Co-Training on captioned images and videos. Machine Learning and Knowledge Discovery in Databases. LNCS 5211, Berlin, Heidelberg: Springer-Verlag, 2008. 457–472. [doi: 10.1007/978-3-540-87479-9_48]
- [15] Zhang TZ, Liu S, Xu CS, Lu HQ. Boosted multi-class semi-supervised learning for human action recognition. Pattern Recognition, 2011,44(10-11):2334–2342. [doi: 10.1016/j.patcog.2010.06.018]
- [16] Liu C, Yuen PC. A boosted co-training algorithm for human action recognition. IEEE Trans. on Circuits and Systems for Video Technology, 2011,21(9):1203–1213. [doi: 10.1109/tcsvt.2011.2130270]
- [17] Wang S, Ma ZG, Yang Y, Li X, Pang CY, Hauptmann AG. Semi-Supervised multiple feature analysis for action recognition. IEEE Trans. on Multimedia, 2014, 16(2):289–298. [doi: 10.1109/tmm.2013.2293060]
- [18] <http://www.xbox.com/en-US/kinect>
- [19] Kuncheva LI, Whitaker CJ. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Machine Learning, 2003,51(2):181–207. [doi: 10.1023/A:1022859003006]
- [20] Kunchva LI, Whitaker CJ. Ten measures of diversity in classifier ensembles: Limits for two classifiers. In: Proc. of the IEEE Workshop on Intelligent Sensor Processing. IET Press, 2001. 10/1–10/10. [doi: 10.1049/ic:20010105]
- [21] Mallapragada PK, Jin R, Jain AK, Liu Y. SemiBoost: Boosting for semi-supervised learning. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009,31(11):2000–2014. [doi: 10.1109/tpami.2008.235]
- [22] Dollár P, Rabaud V, Cottrell G, Belongie S. Behavior recognition via sparse spatio-temporal features. In: Proc. of the IEEE Int'l Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. IEEE Press, 2005. 65–72. [doi: 10.1109/vspets.2005.1570899]
- [23] Niebles JC, Wang HC, Li FF. Unsupervised learning of human action categories using spatial-temporal words. Int'l Journal of Computer Vision, 2008,79(3):299–318. [doi: 10.1007/s11263-007-0122-4]
- [24] Schudt C, Laptev I, Caputo B. Recognizing human actions: A local SVM approach. In: Proc. of the 17th IEEE Int'l Conf. on Pattern Recognition. IEEE Press, 2004. 32–36. [doi: 10.1109/icpr.2004.1334462]



唐超(1977—),男,安徽合肥人,博士,讲师,CCF 会员,主要研究领域为机器学习,模式识别,计算机视觉,人体行为识别.



李国斌(1981—),男,实验师,主要研究领域为机器视觉,嵌入式系统.



王文剑(1968—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为机器学习,数据挖掘,计算智能.



曹峰(1980—),男,博士,讲师,CCF 会员,主要研究领域为机器学习,粗糙集理论.



李伟(1979—),男,博士,副教授,CCF 会员,主要研究领域为人工智能,计算机图形学,人机交互.