

一种简便的视角无关动作识别方法

王策,姬晓飞,李一波

(沈阳航空航天大学 自动化学院,辽宁 沈阳 110136)

摘 要:针对日常生活中人体执行动作时存在视角变化而导致难以识别的问题,提出一种基于视角空间切分的多视角空间隐马尔可夫模型(HMM) 概率融合的视角无关动作识别算法。该方法首先按照人体相对于摄像机的旋转方向将视角空间分割为多个子空间,然后选取兴趣点视频段词袋特征与分区域的光流特征相融合,形成具有一定视角鲁棒性特征对人体运动信息进行描述,并在每个子视角空间下利用 HMM 建立各人体动作的模型,最终通过将多视角空间相应的动作模型似然概率加权融合,实现对未知视角动作的识别。利用多视角 IXMAS 动作识别数据库对该算法进行测试的实验结果表明,该算法实现简单且对未知视角下的动作具有较好识别结果。

关键词:动作识别;视角无关;视角空间切分;兴趣点;光流特征;混合特征;隐马尔可夫模型;似然概率加权融合

中图分类号: TP391.4 **文献标志码:**A **文章编号:**1673-4785(2014)05-0577-07

中文引用格式:王策,姬晓飞,李一波. 一种简便的视角无关动作识别方法[J]. 智能系统学报, 2014, 9(5): 577-583.

英文引用格式:WANG Ce, JI Xiaofei, LI Yibo. Study on a simple view-invariant action recognition method[J]. CAAI Transactions on Intelligent Systems, 2014, 9(5): 577-583.

Study on a simple view-invariant action recognition method

WANG Ce, JI Xiaofei, LI Yibo

(School of Automation, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: It is difficult to recognize the human actions under view changes in daily living. In order to solve this problem, a novel multi-view space hidden Markov model algorithm for view-invariant action recognition based on view space partitioning is proposed in this paper. First, the whole view space is partitioned into multiple sub-view spaces according to the rotation direction of a person relative to camera. Next, a view-robust feature representation by combination of the bag of interest point words in shot length-based video and amplitude histogram of local optical flow is utilized for describing the information of human actions. Thereafter, the human action models in each sub-view space are trained by HMM algorithm. Finally, the unknown view action is recognized via the likelihood probability weighted fusion of the corresponding action models in multi-view space. The experimental results on multi-view action recognition dataset IXMAS demonstrated that the proposed approach is easy to implement and has satisfactory performance for the unknown view action recognition.

Keywords: action recognition; view-invariant; view-space partitioning; interest points; optical flow; mixed feature; hidden Markov model; likelihood probability weighted fusion

基于视频的动作识别研究具有广阔的应用前景^[1-2]。目前大多数动作识别方法都要求运动人体

与摄像机的成像平面平行,或要求摄像机随人体同步运动,即人体相对于摄像机之间的视角固定或在一定范围内运动,然而这样的要求在现实中往往不能得到满足。在实际应用中,人体相对于摄像机的视角通常是任意的、不受约束的,这就要求动作识别

收稿日期:2013-07-30.

基金项目:国家自然科学基金资助项目(61103123).

通信作者:姬晓飞.E-mail:jixiaofei7804@126.com.

或行为分析具有一定的视角无关性^[3]。目前,越来越多的研究者开始投入到了视角无关的人体运动分析的研究之中,并且取得了初步的进展^[4-6]。如 Weinland 等^[7]提出一种 motion history volumes (MHVs) 模版用来表征人体的动作,然后在柱面坐标系中使用傅里叶变换对 MHV 模版进行视角无关的特征描述,最后利用马氏距离和线性判别分析对动作进行识别。此类算法通常实现简单且计算复杂度较低,但往往缺乏考虑运动序列中相邻帧之间的动态特性,因此对于噪声和运动的时间间隔变化较敏感。

Peursum 等^[8]利用分等级的隐马尔可夫模型(hidden Markov modes, HMMs)同时实现动作建模和人体跟踪,该模型分为两层,底层对应于基于3-D姿态的子动作建模,上层对应于动作建模。该方法不仅能够对动作进行识别,还大大地提高了人体跟踪的准确性。此类方法采用3-D信息表示人体姿态信息,自然消除了视角对于模型的影响,但是受到3-D跟踪方法性能影响较大。LYU 等^[9]利用3-D motion capture 数据构造关键姿势,并将各关键姿态数据向多个视角投影得到多视角姿态图像库,采用金字塔匹配核算法快速比较两个人体轮廓的相似性,并利用动作图模型建模动作不同姿态之间的转换约束关系和不同视角之间的转换约束关系对单目视频的人体运动进行识别。此方法需要预先建立多视角下的人体姿态库,并且采用轮廓信息表示,由于采集数据不可能覆盖所有可能视角,因此不能保证识别过程中姿态的最佳匹配,从而影响识别结果。

基于以上分析,本文将具有一定视角鲁棒性的运动特征与典型的状态转移方法有效结合,由特征提取和识别方法两个层次同时解决视角问题,提出了一种基于视角空间切分的多视角空间 HMMs 概率融合的视角无关动作识别算法。该算法因采用视角空间切分思想,在有限个子视角空间下对人体动作建立 HMMs,所以无需大量的特定训练数据。在特征提取阶段由于采用局部兴趣点视频段词袋直方图与光流幅度直方图相融合特征,使得视频动作特征具有一定的视角鲁棒性,同时解决了传统的兴趣点词袋特征无法与具有较强时间建模能力的 HMMs 相结合的问题。

在识别阶段有效地利用了测试序列与多视角空间 HMMs 的似然概率关系,获取了各子视角空间信息,从而根据所获得的视角信息对各子视角空间下动作模型似然概率进行加权融合实现对未知视角动作的识别。

1 特征提取

1.1 兴趣点特征提取

近几年,基于时空兴趣点的特征逐渐得到发展,时空兴趣点代表图像序列中运动改变较为明显的关键位置,围绕兴趣点的各种时空特征均带有丰富的动作信息^[10]。传统的兴趣点词袋特征通常将整个动作视频利用标准单词库进行表征^[11-12],导致其特征缺失时序特性无法与具有较强时间建模能力的 HMMs 相结合。为了充分地利用视频帧间的上下文信息并有效地结合各帧之间的时间特性,文中采用视频段词袋直方图方法对每帧动作图像进行表征。其过程如图1所示:首先,利用2维 Gabor 滤波器对运动区域进行多方向的滤波进行兴趣点检测^[13](文中分别选取从0°、22°、45°、67°、90°进行滤波),寻找视频中那些具有突变性质的视频帧,即有明显运动特性的帧。其次,对这些具有丰富信息的视频帧中的兴趣点利用3-D SIFT 描述算子^[14]进行描述(文中3-D SIFT 描述子选为256维)。最后,利用 Kmeans 聚类算法,对用于训练的动作视频中提取出来的兴趣点3-D SIFT 描述算子进行聚类,建立一个具有 ck 个(文中取 $ck=60$)标准单词的词袋库。当新的动作视频帧输入时,利用其前后 f 帧(文中取 $f=4$)视频段描述算子的 ck 维词袋直方图对其进行表征,即分别计算此视频段内兴趣点描述算子与标准词库内单词的欧氏距离,利用最小距离的判据对视频段内的描述算子进行归类,最后统计输入视频段内标准词库中单词所发生频数,形成当前帧兴趣点视频段词袋直方图描述。该部分特征提取的详细过程可参考前期工作^[15]。

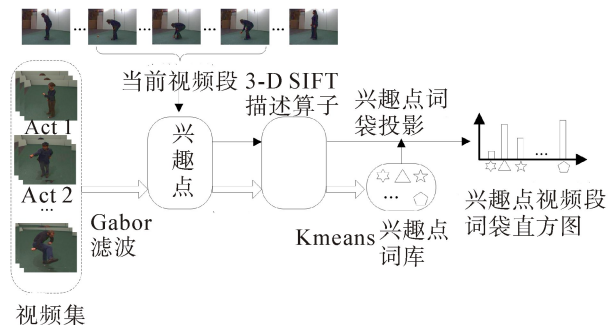


图1 兴趣点视频段词袋特征提取框图

Fig.1 Bag of interest point words feature extraction in shot length-based video

1.2 光流特征提取

为了提高视频段词袋特征对于运动的信息的表述能力,本文在局部特征兴趣点描述算子的基础上,引入了对环境变化有较强适应能力,并且能够同时

表征人体运动信息的光流特征与兴趣点视频段词袋特征进行融合。为了减少光流提取的计算量,本文只在动作视频中具有突变性质的视频帧中提取光流特征,这样的操作能够去掉大量的冗余视频帧。

光流特征提取过程如图 2 所示。

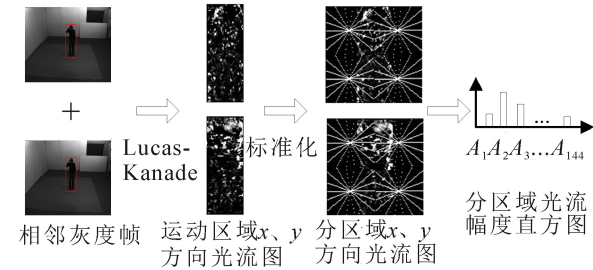


图 2 分区域光流幅度直方图特征提取框图
Fig.2 Grid-based amplitude histogram of optical flow feature extraction

首先,针对相邻两帧的运动区域所对应灰度图像运用 Lucas-Kanade 算法提取动作视频帧的运动区域水平和垂直方向的光流图。其次,鉴于各帧运动区域标准不同,需对提取出的运动区域光流图进行标准化处理,对其利用双线性插值法,使其标准化为 $p \times q$ 维标准大小光流图(文中 $p=q=120$)。通常整体的光流图特征对外界环境的抗扰能力较差,所以本文采用分区域的思想^[16]来进行表述。对所提取的动作视频帧光流图划分为 2×2 子块,且对各子块内部,求取各子块内 18 维的光流幅度径向直方图,即统计各子块内部 18 个扇形区域的光流幅度。如此,每帧动作图像便可由 144 维($2 \times 2 \times 2 \times 18$)光流特征所表征。

1.3 混合特征提取

为了有效地提高识别算法对于复杂环境和视角变化的适应性,本文将上述所得到的局部兴趣点视频段词袋直方图与光流幅度直方图相结合,其结合过程由式(1)给出,该混合特征既可以表述人体运动过程整体的运动信息,又可以很好地提高对于遮挡等问题适应性,同时实验证明该混合特征在一定的视角变化范围内具有一定鲁棒性^[15],最终获得 204 维(60+144)混合特征来表征各动作视频帧。

$$O_t = [O_{Fx} \ O_{Fy} \ O_H] \tag{1}$$

其中

$$O_{Fx} = [O_{Fx_{1,1}} \ O_{Fx_{1,2}} \ \dots \ O_{Fx_{1,18}} \ \dots \ O_{Fx_{4,1}} \ O_{Fx_{4,2}} \ \dots \ O_{Fx_{4,18}}]$$

O_{Fx} 为水平方向光流特征;

$$O_{Fy} = [O_{Fy_{1,1}} \ O_{Fy_{1,2}} \ \dots \ O_{Fy_{1,18}} \ \dots \ O_{Fy_{4,1}} \ O_{Fy_{4,2}} \ \dots \ O_{Fy_{4,18}}]$$

O_{Fy} 为垂直方向光流特征;

$$O_H = [O_{H_1} \ O_{H_2} \ \dots \ O_{H_{60}}]$$

O_H 为兴趣点视频段词袋特征。

2 动作识别

为了实现视角无关的动作识别,本文将运动人体相对于摄像的旋转视角进行平均分割,即将其切分为有限个 V 份子视角空间(文中 $V=8$,即每个子视角空间大小为 45°),并在每个子视角空间下对各类动作建立相应的子 HMM 模型,最终通过将多视角空间相应的动作模型似然概率加权融合对未知视角的动作进行识别。

2.1 子视角空间 HMM 建立

记: $\lambda^{(1,v)}, \lambda^{(2,v)}, \dots, \lambda^{(H,v)}$ 分别为当前给定子视角空间 v 下的 1 号动作,2 号动作, ..., H 号动作的 HMM 模型;这样当给定一个测试动作观察序列 $O = \{o_1^{\text{test}}, o_2^{\text{test}}, \dots, o_T^{\text{test}}\}$ 时,利用前向算法计算这个测试动作观察序列与当前子视角空间 v 中各个动作 HMM 模型的相似度,即:

$$p(O^{\text{test}} | \lambda^{(1,v)}), p(O^{\text{test}} | \lambda^{(2,v)}), \dots, p(O^{\text{test}} | \lambda^{(H,v)}) \tag{2}$$

式中: v 为当前子视角空间, H 为 v 子视角空间下动作模型类别数。

各子视角空间下的各 HMM 动作模型训练问题其实质是对于给定 HMM 模型结构(文中选用 5 个状态节点全连接结构 HMM 模型),调整模型参数对于给定的观察序列使得其对应的模型相似度最大。Baum-Welch 算法能够很好地解决模型的训练问题,但由于 Baum-Welch 算法对于初值依赖较大,存在易达到局部最优问题,所以本文利用 Kmeans 算法为 Baum-Welch 算法进行初值选取。同时为了描述连续状态的发生概率,本文采用高斯概率密度函数进行观察概率的计算,使得模型更加合理。

2.2 多视角空间 HMMs 概率融合识别

为有效利用各子视角空间下的动作模型信息,通过给定的测试动作观察序列与各子视角空间下的动作模型似然概率关系对各子视角空间动作模型进行加权处理,利用各子视角空间下加权后的相应动作模型似然概率和来表征测试观察序列与整个多视角空间下的各个动作模型相似度,最终将测试动作归类为与其相似度最高的动作模型所属类别,即

$$\text{action_class} = \underset{1 \leq h \leq H}{\operatorname{argmax}} \left(\sum_{1 \leq v \leq V} w_v \frac{p(O^{\text{test}} | \lambda^{(h,v)})}{\sum_{1 \leq h \leq H} p(O^{\text{test}} | \lambda^{(h,v)})} \right)$$

式中, w_v 为子视角空间权值(分别取为 $\text{inter_}w_v$ 、 $\text{intra_}w_v$ 、 $\text{mean_}w_v$); v 为当前子视角空间; h 为 v 子

视角空间下当前动作模型类别。

本文主要根据两方面来对各子视角空间的概率权重进行选取:

1) 根据各子视角空间内部各动作模型之间的可分性,若当前子视角空间下的动作模型可分性较好,即加重当前子视角空间下的权值。

2) 根据各子视角空间与给定测试动作所在子视角空间的近似度关系,若当前子视角空间与给定测试动作所在子视角空间近似度较高则加重当前视角空间的权值。本文正是基于以上思想来决定各子视角空间权重大小,具体采用以下4种不同的权值求取方式。

(1) 无视角权值:不考虑各子视角空间下的权重关系,直接将各子视角空间下相应的动作模型似然概率相加,得到多视角空间下的待识别动作与各动作模型相似度,取其最大相似度与之对应的动作模型类别作为待识别动作类别,即:

$$\text{action_class} = \underset{1 \leq h \leq H}{\operatorname{argmax}} \left(\sum_{1 \leq v \leq V} \frac{p(O^{\text{test}} | \lambda^{(h,v)})}{\sum_{1 \leq h \leq H} p(O^{\text{test}} | \lambda^{(h,v)})} \right) \quad (3)$$

(2) 视角之间权值:利用各子视角空间与给定测试动作所在子视角空间的近似度关系进行权值求取,通常各子视角空间与给定的测试动作所在子视角空间近似度越高,其视角空间下各动作模型似然概率和越大,所以,本文选取各子视角空间下的各动作模型似然概率和作为当前子视角空间的概率权值求取的准则函数。式(4)给出了一个子视角空间 v 下的权值求取具体过程。

$$\text{inter_}w_v = \frac{\sum_{1 \leq h \leq H} p(O^{\text{test}} | \lambda^{(h,v)})}{\sum_{1 \leq v \leq V} \sum_{1 \leq h \leq H} p(O^{\text{test}} | \lambda^{(h,v)})} \quad (4)$$

(3) 视角内部权值:利用各子视角空间自身内部各动作模型之间的可分性关系进行权值求取,通常情况下,误识别动作往往发生于执行时相似度较高的两个或几个动作之间,如: wave、scratch head、walk、turn around 等,这样就应该着重分析各子视角空间内部的相似动作,本文利用各子视角空间内部相似动作模型间似然概率差作为当前子视角空间的概率权值求取准则函数,其具体形式由式(5)给出。

$$\text{intra_}w_v = \frac{\max_{1 \leq i \leq H} (p(O^{\text{test}} | \lambda^{(i,v)})) - \frac{1}{N} \sum_{1 \leq n \leq N} p(O^{\text{test}} | \lambda^{(n,v)})}{\max_{1 \leq i \leq H} (p(O^{\text{test}} | \lambda^{(i,v)})) + \frac{1}{N} \sum_{1 \leq n \leq N} p(O^{\text{test}} | \lambda^{(n,v)})} \quad (5)$$

式中:

$$n \in A =$$

$$\{j | \max_{1 \leq i \leq H} (p(O^{\text{test}} | \lambda^{(i,v)})) - p(O^{\text{test}} | \lambda^{(j,v)}) \leq \text{th}, 1 \leq j \leq H, j \neq \underset{1 \leq i \leq H}{\operatorname{argmax}} (p(O^{\text{test}} | \lambda^{(i,v)}))\};$$

th 为相似动作间的阈值。

(4) 混合视角权值:为了更加有效的对未知视角的测试动作进行识别,结合2)、3)给出的视角空间概率权值求取思想,将视角之间权值与视角内部权值相结合,这样可使识别算法同时考虑了各子视角之间的相似关系以及视角内部动作模型的可分性,其具体实现由式(6)给出。

$$\text{mean_}w_v = \frac{\max_{1 \leq i \leq H} (p(O^{\text{test}} | \lambda^{(i,v)})) - \frac{1}{H-1} \sum_{1 \leq j \leq H} p(O^{\text{test}} | \lambda^{(j,v)})}{\sum_{1 \leq v \leq V} \sum_{1 \leq h \leq H} p(O^{\text{test}} | \lambda^{(h,v)})} \quad (6)$$

式中: $j \neq \underset{1 \leq i \leq H}{\operatorname{argmax}} (p(O^{\text{test}} | \lambda^{(i,v)}))$

3 实验与结果与分析

本文所用实验数据均来自于 Inria Xmas Motion Acquisition Sequences (IXMAS) 数据库,其中包含了12个动作、5个摄像机视角,且在每个摄像机视角下每个动作分别由13个人执行3遍,在动作执行过程中,人体相对与摄像机的视角不受约束,该数据库被广泛用于视角无关的人体动作识别。

本文首先对数据库中的各摄像机下的动作数据进行视角聚类,每个摄像机下的动作数据分别将其聚类为3个子视角空间进行实验,在各实验中采用留一法对不同执行人进行循环测试,最终得到实验结果。图3中的每一行展示出了各摄像机下聚类出的子视角空间中一类动作数据样例。

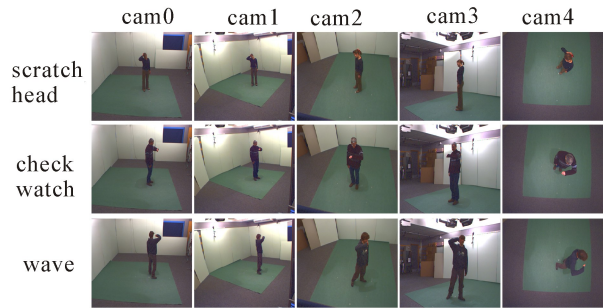


图3 多视角 IXMAS 数据库示例图

Fig.3 Exemplar frames from IXMAS multi-view dataset

3.1 特定摄像机下视角无关动作识别性能验证

在本次实验中,利用 IXMAS 数据库中特定摄像机下的动作数据对算法识别性能的旋转视角变化适

应性进行测试,并且比较了不同的视角概率权值选取方法对于识别性能的影响。如图 3 所示,虽然测试与训练的数据均来自于同一特定摄像机,但人体执行动作时相对于摄像机的旋转视角是不固定的、存在较大的差异,因此通过此实验可对识别算法的旋转视角无关的动作识别性能进行验证。表 1 中分别给出了不同视角空间权值的多视角空间 HMMs 算法在 IXMAS 数据库 5 个摄像机视角下的视角无关动作识别结果。

表 1 特定摄像机下不同视角权值视角无关动作识别结果

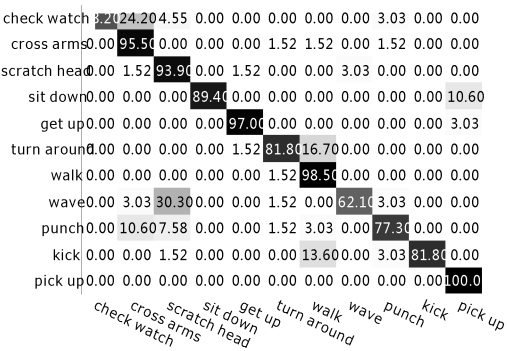
Table 1 The recognition result based on different view space weights under fixed camera parameters

	cam0	cam1	cam2	cam3	cam4	识别率/%
without_ w_v	89.26	87.68	86.84	80.61	78.99	84.68
inter_ w_v	89.26	88.01	87.34	80.61	78.99	84.84
intra_ w_v	89.56	87.51	89.93	86.43	81.55	87.00
mean_ w_v	90.23	88.65	91.15	87.54	82.76	88.07

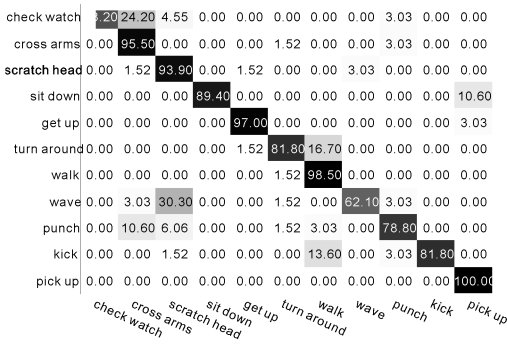
由表 1 实验结果可以看出,对于不使用视角概率权值直接多视角模型概率相加的结果并不是很理想,因为其无法有效地利用其视角间的信息关系,无法分辨与其测试序列所在子视角空间的近似关系。而考虑了各子视角空间关系的 inter_ w_v 权值和视角内部动作可分析的 intra_ w_v 权值算法,在直接加权模型的基础上其识别概率有所提高,这正可以说明测试动作在不同的视角空间下的不同表现是可以通过合适的处理,使得其所提供的信息得到有效利用。mean_ w_v 权值算法由于同时考虑了视角间与视角内整体动作的信息,使得视角无关的动作识别率进一步得到提升。

3.2 不同摄像机参数下视角无关动作识别性能验证

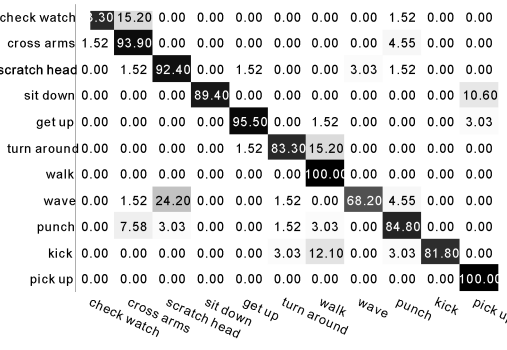
在本次实验中,利用 IXMAS 数据库中不同摄像机参数下的动作数据对算法识别性能的任意视角变化(旋转和俯仰)适应性进行测试。如图 3 所示,在 IXMAS 数据库中,每个摄像机相对于执行人的俯仰角度是不同的,且同一摄像机下的动作数据旋转视角也具有变化性。因此,本次实验利用多个摄像机参数下的动作数据对各子视角空间下的 HMMs 进行训练,测试时利用已训练好的多视角空间 HMMs 对各不同摄像机参数下的动作数据进行识别(测试数据不用于模型训练中),从而对算法的视角无关识别性能进行验证。图 4 给出了利用不同摄像机参数的动作数据对多视角空间 HMMs 识别算法在不同视角空间权值下进行视角无关识别性能验证的混淆概率矩阵图及相应的整体识别率。



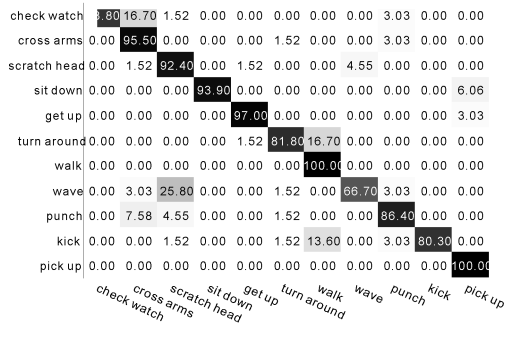
(a) 无视角空间权值算法(识别率 85.95%)



(b) inter- w_v 视角权值算法(识别率 86.09%)



(c) intra- w_v 视角权值算法(识别率 88.43%)



(d) mean- w_v 视角权值算法(识别率 88.43%)

图 4 不同视角空间权值的动作识别混淆概率矩阵图

Fig.4 Confusion matrixes for view-independent action recognition based on different view space weights

从混淆矩阵中可以看出,本文算法对大部分的视角变化动作均能正确识别,其识别率可达 85%,引入了视角空间概率权值后识别性能更是得到提升。从混淆概率矩阵的观察来看,Scratch head、wave 两组动作相互间存在较高误识率,主要是由于这两组动作的相似性过高导致的。

3.3 不同算法识别性能比较

在表 2 中给出了一些近年来同样在 IXMAS 数据库中进行视角无关的动作识别实验结果,将本文算法识别结果与其进行比较,进而验证本文算法对视角变化的适应能力,其各自实验环境下的训练和测试数据均采用不同摄像机参数下的动作数据,识别比较结果如表 2 所示。

表 2 不同算法在 IXMAS 数据库中的动作识别结果比较
Table 2 Comparison of different algorithms forview-independent action recognition in IXMAS dataset

来源	方法	识别率/%
本文	Mixed feature + mutil-view HMMs	88.4
文献[9]	Shape context + PMK-NUP	80.6
文献[17]	SSM descriptor+ SVM	79.6
文献[18]	HOG descriptor + SVM	83.4
文献[19]	BOBW + LWE	82.8

从表 2 的实验结果来看,本文所提出的算法在视角无关动作识别准确率方面要明显优于其他的识别算法。本文算法识别特征获取方面,不用像 LYU 等^[9]那样需要构建 3-D key pose,也不用像 Liu 等^[19]在识别时需要预先知道测试相机的信息;本文特征提取简单,且抗干扰能力强,具有一定时间特性,其在识别时,能够很好地利用上下文的运动信息与 HMM 模型相结合,在识别时利用多视角空间概率图模型思想,对视角空间进行切分处理,并对每个子视角空间下的动作模型似然概率进行加权融合,有效地利用了各子视角空间下的动作模型信息,增加了算法识别的准确率和视角鲁棒性,识别算法简便,计算复杂度低,且实验表明准确率较高。

4 结束语

文中提出了一种基于视角空间切分的多视角空间 HMM 模型概率融合的视角无关动作识别算法,其有效地融合了基于局部信息的兴趣点视频段直方图与基于整体运动信息的分区域光流幅度直方图,使得视频动作特征具有视角的鲁棒性。在识别阶段利用多视角空间构建 HMM 动作模型,根据测试动作所在视角空间与各子视角空间近似度关系以及子

视角空间内部动作可分性关系对子视角空间下的动作模型似然概率进行加权融合,进而对未知视角的测试动作进行识别。在 IMXAS 数据库上的大量测试实验表明,该算法实现简单,且对未知视角下的动作具有较好识别效果。本文算法对相似动作间的识别尚存一定得误识率,下一步的研究重点将是进一步完善特征的表示和识别模型结构,合理利用视角之间的转换信息进一步提高算法的识别率。

参考文献:

[1] WEINLAND D, RONFARD R, BOYER E. A survey of vision-based methods for action representation, segmentation, and recognition [J]. Computer Vision and Image Understanding, 2011, 115(2): 224-241

[2] SEO H, MILANFAR P. Action recognition from one example [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 867-882.

[3] JI X, LIU H. Advances in view-invariant human motion analysis: a review[J]. IEEE Transactions on Systems, Man and Cybernetics, Part C, 2010, 40(1): 13-24.

[4] JUNEJO I, DEXTER E, LAPTEV I. Action recognition from temporal self-similarities[C]//Proceeding of the 10th European Conference on Computer Vision. Marseille, France, 2008:293-306.

[5] LYU F, NEVATIA R. Recognition and segmentation of 3-D human action using HMM and multi-class AdaBoost[C]// Proceeding of the 9th European Conference Computer Vision, Graz, 2006: 359-372

[6] YU S, TAN D, TAN T. Modelling the effect of view angle variation on appearance-based gait recognition [C]//Proceeding of the 8th Asian Conference Computer Vision. Hyderabad, 2006: 807-816.

[7] WEINLAND D, RONFARD R, BOYER E. Free viewpoint action recognition using motion history volumes[J]. Computer Vision and Image Understanding, 2006, 104(2): 249-257.

[8] PEURSUM P, VENKATESH S, WEST G. Tracking as recognition for articulated full body human motion analysis [C]//Proceeding of the 25th IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA, 2007: 1-8.

[9] LYU F, NEVATIA R. Single view human action recognition using key pose matching and Viterbi path searching//Proceeding of the 25th IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA, 2007: 1-8.

[10] DOLLAR P, RABAUD V, COTTRELL G, et al. Behavior recognition via sparse spatio-temporal features[C]// Proceeding of IEEE International Workshop on Visual Surveil-

- lance and Performance Evaluation of Tracking and Surveillance. 2005: 65-72.
- [11] SONG Y, TANG S. A distribution based video representation for human action recognition[C]// Proceeding of the IEEE International Conference on Multimedia and Expo. Suntec, Singapore, 2010: 772-777.
- [12] REN H, THOMAS B. Action recognition using salient neighboring histograms [C]// Proceeding of the 20th IEEE International Conference on Image Processing. Melbourne, Australia, 2013: 2807-2811.
- [13] BREGOMZIO M, GONG S, XIANG T. Recognizing action as clouds of space-time interest points[C]//Proceeding of the 27th IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA, 2009: 1948-1955.
- [14] SCOVANNER P, ALI S, SHAH M. A 3-dimensional sift descriptor and its application to action recognition [C]// Proceeding of the 15th International Conference on Multimedia. Augsburg, Germany, 2007: 357-360.
- [15] 姬晓飞,王策,李一波. 一种基于视角鲁棒性特征的人体动作识别方法研究[C]//第32届中国控制会议论文集. 西安, 中国, 2013: 3877-3881.
- JI Xiaofei, WANG Ce, LI Yibo. An action recognition method based on view-insensitive feature representation [C] // Proceeding of the 32th Chinese Control Conference. Xi'an, China, 2013: 3877-3881.
- [16] TRAN D, SOROKIN A. Human activity recognition with metric learning [C]//Proceeding of the 10th European Conference on Computer Vision. Marseille, France, 2008: 548-561.
- [17] JUNEJO I, DEXTER E, LAPTEV I, et al. Cross-view action recognition from temporal self-similarities [C]//Proceeding of the 10th European Conference on Computer Vision. Marseille, France, 2008: 293-306.
- [18] WEINLAND D, OZUYSAL M, FUA P. Making action recognition robust to occlusions and viewpoint changes [C]// Proceeding of the 11th European Conference on Computer Vision. Heraklion, Greece, 2010: 635-648.
- [19] LIU J, SHAH M, KUIPERS B, et al. Cross-view action recognition via view knowledge transfer [C]//Proceeding of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Colorado, USA, 2011: 3209-3216.
- [20] JIANG Z, ZHE L, LARRY D. Recognizing human actions by learning and matching shape-motion prototype trees[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(3): 533-547.

作者简介:



王策,男,1989年生,硕士研究生,主要研究方向为基于视频的人体运动建模及其分析。



姬晓飞,女,1978年生,副教授,博士,主要研究方向为视频分析与处理、模式识别。承担国家自然科学基金项目、教育部留学回国启动基金项目等多项基金研究工作。发表学术论文30余篇,其中被SCI、EI检索15篇。参与编著英文专著1部。



李一波,男,1963年生,教授,博士生导师,博士,主要研究方向为生物特征识别、图像处理与模式识别、飞行控制、复杂系统、人工智能等。曾获军队级科技进步3等奖1项、省国防工业办公室科技进步2等奖1项。主持和参与近30项科研项目,发表学术论文100余篇,其中被SCI及EI检索30余篇。