

基于时空特征点的非监督姿态建模和 行为识别的算法研究

王传旭¹, 刘 云¹, 厉万庆²

(1. 青岛科技大学信息学院, 山东青岛 266061;

2. School of Computer Science and Software Engineering, University of Wollongong, NSW, Australia 2522)

摘 要: 姿态建模是行为描述和识别的关键环节,提出了基于时空特征点的姿态建模方法.提取样本库中姿态的时空特征点作为底层特征,一个姿态对应一个时空特征点集合;采用非监督分类方法对姿态样本归类,获得典型姿态的聚类结果,每个典型姿态类别采用基于 EM 的高斯混合模型进行建模.计算各典型姿态间的转移概率,建立状态已知的描述人类多种行为的 VMM 模型,实现对行为的描述.针对行为识别问题,提出将汉字分词的二元语法模型应用于行为识别.实验证明该算法具有计算复杂度低、鲁棒性强等优点.

关键词: 行为识别; 姿态建模; 时空特征点; 二元语法模型

中图分类号: TP301.6 **文献标识码:** A **文章编号:** 0372-2112 (2011) 08-1751-06

Research of Unsupervised Posture Modeling and Action Recognition Based on Spatial-Temporal Interesting Points

WANG Chuan-xu¹, LIU Yun¹, LI Wan-qing²

(1. College of Informatics, Qingdao University of Science and Technology, Qingdao, Shandong 266061, China;

2. School of Computer Science and Software Engineering, University of Wollongong, NSW 2522, Australia)

Abstract: Posture modeling is critical for action description and recognition, a posture modeling and action recognition method is proposed in this paper. Spatial Temporal Interesting Points (STIPs) are extracted from learning samples, in fact, one posture consists of a set of STIPs; a unsupervised clustering method is adopted to classify salient postures from these posture samples, then a GMM model is established for each clustering result; transitional probability among salient postures are calculated, and a Visible state Markov Model(VMM) is learnt to describe various actions. Bi_gram method is put forward for action recognition, Extensive experiments are conducted and the results prove its robustness and validity.

Key words: action recognition; posture modeling; spatial-temporal interesting points; bi_gram

1 引言

人类行为理解与识别是当前研究热点,2006年 Moeslund^[1]将行为分为三个层次:即运动单元(motor primitives)、动作(action)和行为(activity).目前对行为的研究大多针对第2层次,本文研究也是针对这个层面的,即对简单行为,如:走、跑、站立、蹲下、挥手等动作的识别.

行为识别算法大致分为模板匹配算法和统计模型算法^[2].模板匹配的优点是计算代价低,但对运动持续时间的变化和噪声比较敏感.统计模型方法最有代表性的是隐马尔可夫模型 HMM^[3~5].Robertson 等人^[6]将人

体行为定义为随机序列,通过对图像特征样本库学习获得 HMM 模型,实现人体动作识别. Venkatesh 等^[7]人提取运动向量,输入训练好的 HMM 进行动作分类. Ahmad 等人^[8]提出不同视点的人体动作识别方法,将人体形状与光流作为特征,PCA 降维后,用多维离散 HMM 对不同视点下的动作进行识别. Bregler^[9,10]基于人体动力学在不同抽象等级的统计分解提出了层次化方法,通过最大化 HMMs 后验概率来完成识别.

上述方法大都依靠分割人体轮廓,易受到背景噪声干扰,使分割的姿态不完整,而导致识别精度差^[11,12].时空特征点 STIPs(Spatial Temporal Interesting Points)作为一种提取底层运动特征的方法,不需要背景建模,近几

年被引起重视^[13,14]. 本文提取 STIPs 点作为底层特征, 利用无监督分类实现姿态建模. 针对行为识别方法, 采用汉字识别的二元语法模型, 提高了识别效果.

2 总体算法描述

整体算法分为模式学习和行为识别两阶段. 在模式学习阶段, 从行为样本数据库中, 通过非监督分类获得典型姿态, 建立姿态间的 VMM 图形, 实现对行为的描述. 如图 1 所示, 以 Walk、Run 和 Side 三种行为为例, 提取了 9 个典型姿态 $S_0 \sim S_8$, 为了更清晰地观察每个行为姿态间的转移, 将它们分别描述.

由图 1 可见, 有些姿态为三种行为共享, 如 S_4 ; 有些姿态则为某种行为独有, 如 S_1 仅为 Run 所有; 一种行为可用图中的一条或多条路径表示. 在行为识别阶段, 根据输入测试视频中姿态出现的顺序, 计算 VMM 模型中的最大概率路径, 进而识别其行为类别. 本文核心内容为典型姿态提取和建模, 以及识别算法的研究, 详述如下.

3 姿态建模

3.1 时空特征点的提取

一种行为的特点通常反映在运动最剧烈的部位, 如行走中脚的抬起、膝关节的弯曲等, 这些运动区域在本文用 STIPs 刻画出来. 典型的 STIPs 点提取有 Ivan^[13] 和 Dollár^[14] 方法, 其中 Ivan 方法是基于时空 Harris 角点的检测, 得到的 STIPs 点多发生在运动突变的区域, 如脚跟抬起、膝关节等部位, 特征点数量较少, 属于稀疏型. 稀疏型特征点容易丢失视频中重要的运动信息, 导致检测实效. 相比之下, Dollár 方法提取的特征点较多,

可以充分捕获运动产生的变化, 属于稠密型特征点. 本文采用 Dollár 方法提取 STIPs 点.

Dollár 认为, 任何运动剧变的人体部位区域都会产生时空特征点, 其运动强度捕获函数 R 定义为:

$$R(x, y, t) = (I * g * h_{ev})^2 + (I * g * h_{od})^2 \quad (1)$$

其中, $*$ 为卷积运算符, I 为视频帧图像, $g(x, y; \sigma)$ 为二维空域高斯滤波器, h_{ev} 和 h_{od} 为一维 Gabor 时域滤波器, 定义分别为:

$$h_{ev}(t; \tau, \omega) = -\cos(2\pi t\omega) e^{-t^2/\tau^2} \quad (2)$$

$$h_{od}(t; \tau, \omega) = -\sin(2\pi t\omega) e^{-t^2/\tau^2} \quad (3)$$

其中, σ 和 τ 分别为空域和时域尺度, 且令 $\omega = 4/\tau$.

Dollár 方法提取的 STIPs 点是对尺度敏感的. 为了克服尺度变化的影响, 通常取 $\sigma \in [2, 9]$, $\tau \in [1, 8]$, 在数值计算中将它们离散化取值, 离散间隔以 $\sqrt{2}$ 为倍数^[15]. 本文 σ 离散取值为 $[2, 3, 4, 6, 8]$, τ 为 $[1, 2, 3, 4, 6]$; 采用尺度组合方式, 即将 τ 和 σ 的这 5 个尺度进行组合, 构成 25 种尺度搭配; 然后, 计算不同尺度下的运动强度函数 R , 求出 R 局部最大值作为不同尺度下的 STIPs 点. 检测阈值的选取本文是以保证每个姿态图像中 STIPs 点数量不少于 15 个为依据而定.

通过对多组视频的研究发现, 相同姿态具有相似分布的时空特征点, 如图 2 所示. 图中三位不同的行人, 尽管背景不同、人体图像尺度大小不同, 但它们的 STIPs 分布具有很好的相似性, 这为姿态聚类提供了基础信息.

3.2 非监督的姿态分类

3.2.1 姿态描述符建立

一个姿态通常可检测出多个不同尺度下的 STIPs 点, 先建立每个 STIPs 点的描述符, 然后再对每个姿态建模.

为了描述不同尺度 STIPs 点区域的时空变化情况, 以每个 STIPs 点为中心选取一个时空立方体, 其尺寸为 STIPs 点尺度值的 6 倍. 根据尺度归一化的时空高斯微分模型, 计算立方体内部每个像素的一阶梯度 (L_x, L_y, L_t) .

$$L_{x y t}^{m n k} = \sigma^{m+n} \tau^k (\partial_{x y t}^{m n k} g) * f \quad (4)$$

这里, g 为三维高斯滤波器 σ 和 τ 为 STIP 点的空间和时间尺度, 即

$$g(x, y, t; \sigma^2, \tau^2) = \frac{1}{\sqrt{(2\pi)^3 \sigma^4 \tau^2}} \exp(-(x^2 + y^2)/2\sigma^2 - t^2/2\tau^2) \quad (5)$$

f 为视频图像, $*$ 为卷积运算符, m, n, k

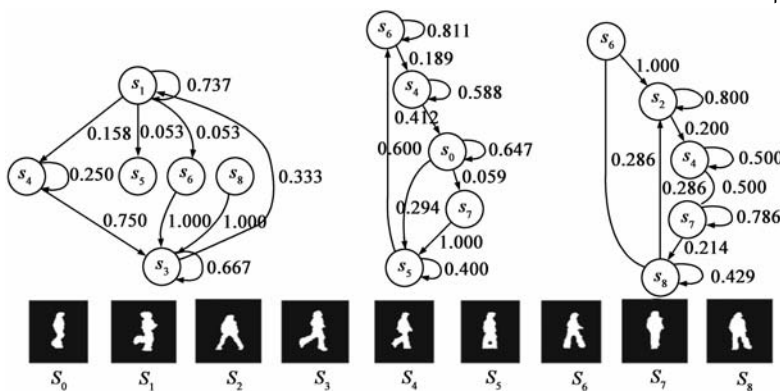


图1 Run(左)、Walk(中)、Side(右)三种行为的姿态转移Markov模型

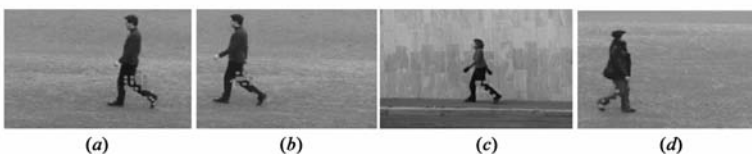


图2 相似的姿态具有相似的STIP点分布

为微分的阶数;另外,为适应不同的视频背景和行人的不同服装灰度变化,将梯度(L_x, L_y, L_t)进行幅度归一化。

每帧中的姿态对应一个 STIPs 点集,该点集作为一个整体共同刻画了当前时刻的姿态时空特征.针对每个姿态统计三个梯度分量的子直方图,并将它们级联作为一个具体姿态的模型。

3.2.2 非监督的姿态分类算法

以 Weizmann 视频库作为学习样本,计算出 J 个具体姿态的描述符.姿态自动分类包括姿态间相似度计算和非监督分类两部分.为了避免相似度计算中的除零溢出问题,选用直方图相交法计算姿态间相似度.设两姿态样本直方图为 q 和 p ,维数为 B ,则两者相似度为:

$$S(p, q) = \sum_{u=1}^B \min\{p^{(u)}, q^{(u)}\} \quad (6)$$

定义差异度为 $d = 1 - S$,这样所有 J 个样本的差异度,可组成一个差异度矩阵 $\mathbf{D} = [d_{ij}]_{J \times J}$, d_{ij} 为 i 与 j 两姿态间的差异度。

差异度矩阵 \mathbf{D} 描述了两姿态间的相似度,关系型特征数据集的分类方法如 NCuts^[16]与 DS^[17],实验中发现它们的相似度传递性能会导致姿态分类的效果不好.传统的关系型模糊分类算法^[18],只适用于欧氏距离关系下的数据集,因此,选用 NERF C-Means 方法^[19]进行姿态的自动分类。

图 3 给出了其中一个聚类结果的部分样本,图中显示姿态中 STIPs 均分布在颠簸的头部、抽动的小腿等运动最强烈区域.时空特征点描述符(L_x, L_y, L_t)中, L_t 分量反映两帧间的运动信息, L_x, L_y 反映空间信息,分类结果反应出该方法的归纳能力和敏锐性,避免了通过传统的背景分割实现姿态分类方法的缺点。

从 Weizmann 样本库聚类的每个典型姿态类别中大概包含 70 到 150 帧样本姿态,需要对每类典型姿态建模.针对有限样本空间上姿态分布的描述问题,采用基于 EM 的高斯混合模型 GMM 进行姿态建模。

4 典型姿态转换的 VMM 模型

为了定量分析 M 种行为中所含有的 ζ 个典型姿态间的相互关系,需要计算典型姿态间的全局转移概率,以及针对每种行为的典型姿态间转移概率,即:

M 个行为表示为: $\Psi = \{\psi_1, \psi_2, \dots, \psi_M\}$

ζ 个姿态表示为: $\Omega = \{\omega_1, \omega_2, \dots, \omega_\zeta\}$

典型姿态全局转移概率矩阵为: $\mathbf{Gp} = [p(\omega_i/\omega_j)]_{i,j=1,\dots,\zeta}$

针对行为 ψ_l 的典型姿态转移概率矩阵为: $\mathbf{Gpl} =$

$[p(\omega_i/\omega_j, \psi_l)]_{i,j=1,\dots,\zeta}^{l=1,\dots,M}$



图3 非监督聚类将具有相似STIPs分布的姿态聚为同一类

根据姿态全局转移概率矩阵,可画出基于 ζ 个状态的描述 M 种行为的状态转移图,它实际上是状态已知的 VMM 模型,即 $Z = (\Omega, \mathbf{Gp}, \mathbf{Gpl}, \Psi)$ 。

5 行为识别算法

行为识别是基于样本学习获取的 VMM 模型实现的,由于一种行为可由 VMM 中的若干个典型姿态按照一定时序路径来描述,因此,行为识别过程实际上就是提取输入视频的姿态序列,匹配 VMM 中是否存在这种典型姿态转移路径,并结合其出现的概率来识别行为的类别。

设输入视频 $F = \{f_1, f_2, \dots, f_n\}$, f_i 为第 $i \in [1, n]$ 帧,行为识别过程如下:

首先,计算每帧图像中姿态的 STIPs 点,并建立每个姿态的描述符,设为 $g = \{g_1, g_2, \dots, g_n\}$ 。

其次,根据式(7)计算每个具体姿态 $g_i, i \in [1, n]$ 所匹配的典型姿态,构成最大似然路径 $S^* = \{S_1^*, \dots, S_q^*, \dots, S_n^*\}, S_q^* \in \Omega$ 。

$$s_q^* = \arg \max_{\omega_r \in \Omega} \sum_{m=1}^K w_m \cdot \eta_m(g_i, \mu_m, \sigma_m) \quad (7)$$

其中, ω_r 为第 r 个典型姿态,该姿态的 GMM 模型中由 K 个高斯子模型, η_m 为第 m 个高斯概率密度函数, μ_m, σ_m 是均值和方差, w_m 是权重,这里 $m \in [1, K]$.当 r 遍历整个典型姿态类别,若第 q 个概率值最大,且大于阈值 δ_1 ,则姿态个体 g_i 与典型姿态 q 匹配。

最后,基于 VMM 计算 S^* 隶属于行为 $\psi_l, l \in [1, 2, \dots, M]$ 的最大似然概率 $L(\psi_h)$,本文采用汉字分词中的一元和二元语法模型计算 $L(\psi_h)$,即:

$$L(\psi_h) = \arg \max_{\psi_l \in \Psi} p(\psi) \prod_{t=1}^n p(s_t^*/\psi) \quad (8)$$

$$L(\psi_h) = \arg \max_{\psi_l \in \Psi} p(\psi) \prod_{t=1}^n p(s_t^*/s_{t-1}^*, \psi) \quad (9)$$

需要说明的是,本文计算 $L(\psi_h)$ 采用二元语法模型,即式(9),但在实验中增加了对一元语法模型识别精度的比较.当 ψ_l 遍历所有 M 个行为,若 $L(\psi_h)$ 最大概率值出现在第 h 行为上,且 $L(\psi_h)$ 大于一定阈值 δ_2 时,则测试视频 F 的行为即判定为 Ψ 中的第 i 种行为.阈值 δ_1 和 δ_2 的确定是通过实验手段确定的,本文取 $\delta_1 = 0.8$ 和 $\delta_2 = 0.65$,可以保证识别正确率高于 95%。

6 实验结果和分析

实验均在 Matlab7.01 环境实现的,微机主频为

2.33GHz, 4G 内存, 模式学习阶段选用 Weizmann 库 10 种行为, 每种行为各有 9 个行为, 共计 90 组视频; 用到的测试数据, 分别为 Weizmann 中 Robust_deform 测试数据和 KTH 视频数据库. 实验分四部分:

(1) LOSO 实验 (leave-one-subject-out), 即将 Weizmann 库中某一个行为人的样本不参加前面的模式学习, 而是用作测试视频, 比较一元和二元语法模型的识别精度与典型姿态个数选择之间关系.

(2) 利用 Robust_deform 视频进行鲁棒性测试和比较.

(3) 选用 KTH 库视频进行鲁棒性测试和比较.

(4) 算法复杂度比较.

6.1 LOSO 实验

将 Weizmann 库中行为 daria 的 10 种行为样本用作测试视频, 其余样本进行模式学习, 改变聚类的典型姿态个数, 分别用一元和二元语法模型进行 LOSO 测试, 结果如图 4.

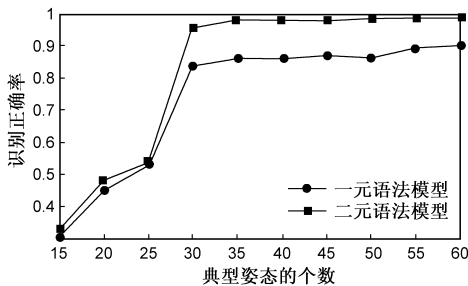


图4 识别精度与典型姿态类别数量之间的关系

实验结果分析 识别正确率随典型姿态数量增加而提高. 当典型姿态数量增至 30 以上时, 一元和二元语法模型的识别率均大幅提升; 当典型姿态数量增至 40 以上时, 二元语法模型识别率 98.2%, 一元语法模型为 88%; 一元语法模型未考虑姿态间的时序逻辑关系, 而二元分类方法则将行为识别看做一个二阶 Markov 过程, 其识别能力优于一元语法模型. 因此, 后续实验仅采用二元语法模型进行测试.

6.2 基于 Robust_deform 测试视频的鲁棒性实验

针对鲁棒性测试, Weizmann 库中有 10 组姿态形变

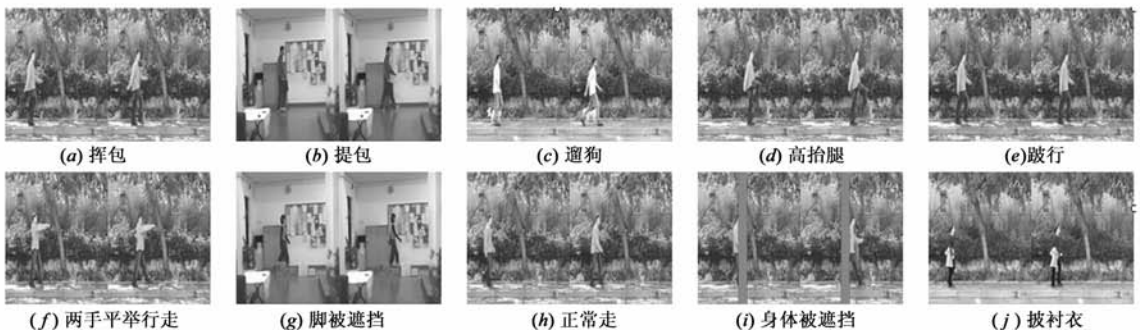


图5 Weizmann库中鲁棒性测试的10组姿态形变的Walk视频

的 Walk 视频, 如图 5 所示. 当典型姿态数量为 40 时, 利用二元语法模型进行识别, 并与同样实验数据下的文献[20]和[21]的结果进行比较, 如表 1 所示.

表 1 基于 Weizmann 库中 Robust_deform 鲁棒性实验比较

行走方式/不同方法	挥包	提包	遛狗	高抬腿	跛行	两手平举行走	脚被箱子遮挡	正常行走	身体被遮挡	腰间缠裹衣服
文献[21]	✓	✓	×	✓	✓	×	✓	✓	✓	✓
文献[20]	✓	✓	✓	✓	✓	×	✓	✓	×	✓
本文方法	×	✓	✓	✓	✓	✓	✓	✓	✓	✓

实验结果及分析 表 1 中结果说明本文方法仅对“行走时挥包”产生误判, 主要原因是该行为中的 STIPs 点不再仅仅出现在脚和膝部, 还分布在大幅度挥动的上肢区域, 与样本学习时的 STIPs 点分布有了很大区别; 对其他的 9 种情况均能正确识别. 而文献[20]和[21]方法分别各有两种情况误判, 表明本文方法优于姿态分割的识别方法.

6.3 基于 KTH 数据库的算法性能测试

为了与文献[20]基于背景分割方法的行为识别结果比较, 本文选取了 36 段 Walk 视频和 32 段 Run 视频, 其中, 每种行为的“常规姿态”和“着装导致姿态形变”的视频各半, 部分截图如图 6 所示. KTH 行为库与 Weizmann 学习样本的主要区别是图像尺寸: Weizmann 库中视频图像大小约为 180×144 , 人体目标约为 61×21 ; 而 KTH 库图像大小约为 160×120 , 人体目标大小约为 79×35 . 第二行图 6(b) 中行为人或身着风衣, 或身着大衣, 导致姿态形变.

针对这 68 个视频鲁棒性测试, 实验结果如表 2, 表中给出了不同典型姿态聚类数量下, 两种方法识别结果比较. 可见, 典型姿态数量增加时, 文献[20]和本文的方法误判数量均会明显降低, 但本文方法在各类情况下均占优.

表 2 基于 KTH 数据库的算法鲁棒性测试

典型姿态数量/不同方法	20	30	45	60
文献[20]的误判数量	11	8	2	2
本文方法的误判数量	9	5	2	1

实验结果分析 Gabor 小波只对人体运动相当的



图6 KTHA数据库行走和跑步行为的部分截图

频率范围最为敏感,可以很大程度上抑制风衣飘动、阴影平移等产生的干扰.图7提取的 STIPs 点分布与图2中的结果基本一致,即相似的姿态具有相似的 STIPs 点分布.该实验结果实际上是进一步证明了基于 STIPs 点进行姿态建模的正确性.而传统人体分割姿态如图8所示,风衣带来的姿态几何形变对行为识别精度影响更大,因而,本文的方法优于文献[20]的实验结果.



图7 KTHA库中相似姿态的STIP点的分布情况

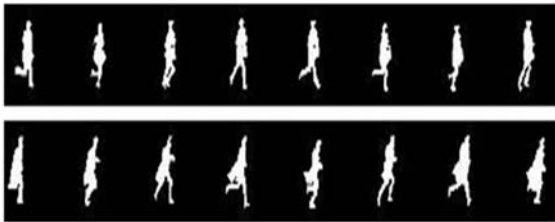


图8 文献[20]针对KTH库中身着风衣跑步视频的人体分割结果

6.4 识别算法复杂度比较

算法复杂度通常包括时间复杂度和空间复杂度,这里以时间复杂度为主比较本文算法和文献[20]的复杂度.选用上一节实验中使用的36段行走和32段跑步视频为测试视频,硬件条件为CPU主频2.33GHz,内存4G.当视频长度 L 由50帧递增1500帧时,统计它们的

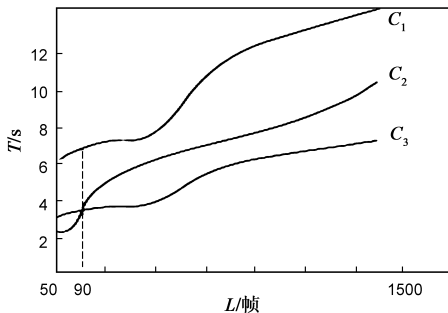


图9 算法时间复杂度比较

运算时间,如图9所示.

实验结果分析 文献[20]算法的复杂度主要取决于分割每帧中姿态的轮廓和建模;本文算法的复杂度主要决定于多尺度的 STIPs 点底层特征提取,它们的运算量都是 $O(L^2)$,但本文算法中的时域卷积和小波相关运算可以借助 FFT 圆周卷积算法,将乘法运算次数降为 $O(L \times \log_2 L)$ 量级.图9结果基本反映两者的算法复杂度关系, C_1 和 C_2 分别本文算法和文献[20]的算法复杂度曲线, C_3 为 FFT 加速后的本文算法.比较 C_2 和 C_3 当视频长度 L 小于 90 帧时,加速效果不明显,但随着视频 L 的增长,本文算法计算耗时将优于文献[20].

7 结论

本文提出了一种姿态建模和行为识别的新方法,较之传统姿态分割实现行为识别的方法具有很好的鲁棒性,主要表现为适应不同的视频背景、具有尺度不变性、对测试视频中的诸多干扰具有良好的抑制能力.本方法的局限性,表现为要求测试视频中摄像机保持不动、算法仅适用于单人行为识别,以及判据阈值的选取仅仅依靠实验经验值确定.作为下一步的工作,拟采用 Geert Willems^[22]方法代替 Dollár 方法计算底层特征,判决阈值的选取采用 ROC (Receiver Operating Characteristic Curve) 曲线确定;此外,进一步研究当行为人与相机间的视角变化较大时,行为识别算法鲁棒性等问题.

参考文献

- [1] T B Moeslund. A survey of advances in vision-based human motion capture and analysis[J]. Computer Vision and Image Understanding, 2006, 104(3): 90 - 261.
- [2] 杜友田,陈峰.基于视觉的人的运动识别综述[J].电子学报, 2007, 35(1): 84 - 90.
Du You-tian, Chen Feng. A survey on the vision-based human motion recognition[J]. Acta Electronica Sinica, 2007, 35(1): 84 - 90. (in Chinese)
- [3] Lawrence R Rabiner. A tutorial on hidden Markov model and selected applications in speech recognition[A]. In Proc ICAS-SP'89[C]. Glasgow: IEEE press, 1989. 257 - 286.
- [4] Michael M, Yin Jason. Application of hidden Markov models to gene prediction in DNA[A]. In Proc. IISR'99[C]. Bethesda: IEEE press, 1999. 40 - 48.
- [5] Yu Shun-zheng, Hisashi Koayashi. An effective forward-backward algorithm for an explicit duration hidden Markov model [J]. IEEE Signal Processing Letters, 2003, 10(1): 11 - 14.
- [6] Robertson, N Reid. A general method for human action recognition in video[J]. Computer Vision and Image Understanding,

- 2006, 104(2): 232 – 248.
- [7] Venkatesh Babu R, et al. Compressed domain action classification using HMM[J]. Pattern Recognition Letters 2002, 23(10): 1203 – 1213.
- [8] Ahmad M, Lee Seong-whan. Human action recognition using multi-view image sequences features[A]. Proc of AFGR'06 [C]. Southampton: UOS press, 2006. 523 – 528.
- [9] Christoph Bregler. Learning and recognizing human dynamics in video sequences[A]. In Proc of CVPR'97[C]. Puerto Rico: IEEE press, 1997. 568 – 574.
- [10] W Lin, M T Sun, et al. Human action recognition for video surveillance[J]. IEEE Transactions, 2008, CSVT-18(8): 1128 – 1139.
- [11] 黎洪松, 许保华. 一种用于视频对象编码的运动模式识别算法[J]. 电子学报, 2007, 35(12): 2324 – 2328.
Li Hong-song, Xu Bao-hua. A motion recognition algorithm for object-based video coding[J]. Acta Electronica Sinica, 2007, 35(12): 2324 – 2328. (in Chinese)
- [12] 魏志强, 纪筱鹏. 基于自适应背景图像更新的运动目标检测方法[J]. 电子学报, 2005, 3(12): 2261 – 2264.
Wei Zhi-qiang, Ji Xiao-peng. A moving object detection method based on self-adaptive updating of background[J]. Acta Electronica Sinica, 2005, 3(12): 2261 – 2264. (in Chinese)
- [13] Ivan Laptev. On space-time interest points[J]. International Journal of Computer Vision, 2005, 64(2/3): 107 – 123.
- [14] P Dollár, V. Rabaud, et al. Behavior recognition via sparse spatio-temporal features[A]. Proc of VSPETS'05[C]. Washington: IEEE Press, 2005. 65 – 72.
- [15] Ivan Laptev, Tony Lindeberg. Interest point detection and scale selection in space-time[J]. Scale Space Methods in Computer Vision Lecture Notes in Computer Science, 2003, 2695(1): 1079 – 1086.
- [16] Jian-bo Shi, Jitendra Malik. Normalized cuts and image segmentation[J]. IEEE Transactions, 2000, PAMI-22(8): 888 – 905.
- [17] Massimiliano Pavan, Marcello Pelillo. Dominant sets and pair wise clustering[J]. IEEE Transactions, 2007, PAMI-29(1): 167 – 172.
- [18] Richard J Hathaway, John W Davenport. Relational duals of the c-Means clustering algorithm[J]. Pattern Recognition, 1989, 22(2): 205 – 212.
- [19] Richard J Hathaway, James C. NERF c-MEANS: Non-Euclidean relational fuzzy clustering[J]. Pattern Recognition, 1994, 27(3): 429 – 437.
- [20] Wan-qing Li, zheng-you Zhang, Zi Liu. Expandable data-driven graphical modeling of human action based on salient postures[J]. IEEE Transactions, 2008, CSVT-18(11): 1499 – 1510.
- [21] 凌志刚, 梁彦, 等. 基于张量子空间学习的人行为识别方法[J]. 中国图像图形学报, 2009, 14(3): 394 – 400.
Ling Zhi-gang, Liang Yan, et al. Human action recognition based on tensor subspace learning[J]. Journal of Image and Graphics. 2009, 14(3): 394 – 400. (in Chinese)
- [22] G Willems, T Tuytelaars. An efficient dense and scale-invariant spatio-temporal interest point detector[A]. Proc of ECCV'08[C]. Berlin: Springer press, 2008. 650 – 663.

作者简介



王传旭 男, 1968 年生于山东邹城, 1990 年与 2000 年获中国石油大学学士和硕士学位, 2007 年获中国海洋大学理学博士学位. 2008 年 9 月至 2009 年 4 月在澳大利亚卧龙岗大学学术访问. 2000 年至今在青岛科技大学信息学院任教, 副教授, 硕士生导师, 研究方向计算机视觉, 发表论文 30 余篇.

E-mail: wangchuanxu_qd@163.com



刘云 男, 1962 年出生, 博士, 教授, 硕士生导师, 青岛科技大学信息学院院长, 主要从事图像处理、模式识别和人机交互技术的研究, 发表论文 50 余篇.

E-mail: lyun-1027@163.com



厉万庆 男, 1964 出生, 澳大利亚卧龙岗大学信息学院任教, 博士, 副教授, 研究方向行为识别和视频事件分析等. 发表学术论文 100 余篇.