

# 一种基于判别随机场模型的联机行为识别方法

黄天羽<sup>1),2)</sup> 石崇德<sup>3)</sup> 李凤霞<sup>2)</sup> 程 成<sup>2)</sup>

<sup>1)</sup>(北京理工大学软件学院 北京 100081)

<sup>2)</sup>(北京理工大学计算机学院智能信息技术北京市重点实验室 北京 100081)

<sup>3)</sup>(北京大学信息管理系 北京 100871)

**摘 要** 提出了一种基于判别随机场模型的联机行为识别方法,将传统的随机场模型和隐藏条件随机场模型的特点相结合,构建一个针对于运动序列帧数据建模的帧-隐藏条件随机场模型,并将该模型应用于数据驱动的行为建模,利用传统条件随机场模型对行为间的运动特性进行建模;通过引入隐藏特征函数,设计有效的特征模板来表示行为中姿态的联系,实现对行为的内在运动特性进行建模.通过对比实验表明,该模型对于联机处理行为序列具有更强的识别能力.

**关键词** 条件判别模型;CRF;Frame-HCRF;行为建模;行为识别

**中图法分类号** TP309 **DOI号**: 10.3724/SP.J.1016.2009.00275

## Discriminative Random Fields for Online Behavior Recognition

HUANG Tian-Yu<sup>1),2)</sup> SHI Chong-De<sup>3)</sup> LI Feng-Xia<sup>2)</sup> CHENG Cheng<sup>2)</sup>

<sup>1)</sup>(School of Software, Beijing Institute of Technology, Beijing 100081)

<sup>2)</sup>(Beijing Laboratory of Intelligent Information Technology, School of Computer Science, Beijing Institute of Technology, Beijing 100081)

<sup>3)</sup>(Department of Information Management, Peking University, Beijing 100871)

**Abstract** This paper proposes an online behavior recognition based on Discriminative Random Fields. In this model, by incorporating CRF and HCRF, a Frame-HCRF was extended to model behaviors for frames of motion data. The motion intrinsic dynamics are captured by CRF structure as well as extrinsic dynamics between different behaviors by hidden feature functions. This model can accommodate motion data online processing with unknown future frames. The experiments show that the proposed model perform over than HMM, CRF and HCRF for human behavior modeling and recognition.

**Keywords** conditional discriminative models; CRF; Frame-HCRF; behavior models; behavior recognition

## 1 引 言

计算机理解人的行为方式是未来计算机技术发展的必然趋势,也是虚拟现实和人机交互领域发展

的必然要求<sup>[1-4]</sup>.三维自然空间中人体行为的建模,是智能人机接口、人机训练系统、智能监测系统和安全系统等应用的主要支撑技术.人体行为复杂、多样、缺少明确的分类结构,同一段运动对于不同的观察者可能被识别为不同的动作:一方面,相似的运动

收稿日期:2008-07-29;最终修改稿收到日期:2008-09-24.本课题得到国家自然科学基金(60773046)、国防基础科研基金(2220061084)、北京市重点学科建设项目资助.黄天羽,女,1979年生,博士,讲师,主要研究方向为计算机图形学、计算机动画、人体运动仿真. E-mail: huangtianyu@bit.edu.cn.石崇德,男,1979年生,博士研究生,主要研究方向为机器翻译、自然语言理解.李凤霞,女,1953年生,教授,主要研究领域为虚拟现实.程 成,男,1966年生,博士,副教授,主要研究方向为人机交互和计算机辅助设计.

在时间线上有不同的放大比例,且具有上下文依赖关系;另一方面,同一段运动通常可以同时被判别属于几个类别,这是因为行为是一些基本运动单元的自然组合,如边跑边招手,边走边握手等.另外,对动作进行分段时往往存在界限模糊或行为重叠等现象.因此,行为识别是一项具有挑战性的工作.

基于产生式模型(Generative Models, GM)的行为建模方法<sup>[5]</sup>,通过观察当前时刻的运动状态往往很难判断该运动所属的类别,最典型的是隐马尔可夫模型(Hidden Markov Model, HMM).产生式模型为了降低计算复杂度,往往假设观察序列是基于条件独立的,这种不确切的假设导致其不能表示时间序列中的依赖关系以及序列间的特征多重重叠.

本文提出一种基于判别式随机场模型(discriminative random field models)的人体行为识别方法,为了改进隐藏条件随机场模型(Hidden Conditional Random Fields, HCRF)<sup>[6]</sup>仅能处理序列数据的问题,将传统的随机场模型(Conditional Random Fields, CRF)和隐藏条件随机场模型的结构特点相结合,构建一个适用于帧数据建模的帧-隐藏条件随机场模型(Frame Hidden Conditional Random Fields, Frame-HCRF),并将该模型应用于数据驱动的联机行为识别,从而提高人体行为的建模能力.

本文第2节介绍用于人体行为建模的相关模型及分析各种方法的优缺点;第3节给出基于判别随机场的行为建模原理和方法;第4节讨论实验方法并分析实验结果,验证本文提出算法的有效性;第5节是结论.

## 2 相关工作

随着交互技术和应用的发展前景越来越广泛,自动行为识别技术已经成为计算机图形学、计算机视觉和模式识别领域研究的热点.

为了研究复杂人体的行为,人们采用非线性系统对行为进行建模.隐马尔可夫模型是实现人体行为建模最常见的模型.为了能够获取复杂运动的层次结构,基于HMM及其扩展模型被广泛应用于2D视频和3D空间<sup>[2,7]</sup>的人体运动建模上,如AHMM<sup>[8]</sup>在HMM之上建立多个语义层研究人体行为运动模式.上述基于HMM的建模方法,其独立性的假设导致产生式模型不能表示像人体运动这种时间序列的上下文依赖关系、重叠性特征.

为了改进产生式模型对时间序列建模所带来的问题,人们开始研究基于判别模型(Discriminative Models, DM)的方法.判别模型最早是根据最大熵马尔可夫模型(Maximum Entropy Markov Models, MEMM)<sup>[9]</sup>的思想建立起来的.针对MEMM的标记偏置问题,提出了CRF<sup>[10]</sup>,该模型通过定义状态序列和观察序列的条件概率来预测最可能的状态序列,从而实现在整个状态空间上的全局归一化分布,将丰富的上下文特征整合到模型中.基于上述原因,CRF模型近五年来逐步受到研究者的关注,并在视频处理和自然语言处理领域<sup>[11]</sup>表现了独特的优势.HCRF通过引入隐藏状态<sup>[6]</sup>对CRF进行了扩展,实现了空间相关性的建模,但是该模型是一种序列-隐藏条件随机场(Sequence-Hidden Conditional Random Fields, Sequence-HCRF),仅能够对已经分段的序列进行识别,即是一种脱机识别模型.正是由于判别模型能够通过属性函数表示大量观察元素之间复杂的相互关联属性和时间依赖关系,Sminchisescu<sup>[1]</sup>和Wang<sup>[3]</sup>开始对人体行为的研究进行了初步尝试.近几年,我国不少学者也开展了人体运动分析相关领域的研究<sup>[12]</sup>,但主要基于产生式模型对人体行为进行建模<sup>[13]</sup>.国内对判别模型的研究还比较少,从国内公开发表的论文来看,判别模型已经开始应用于自然语言处理<sup>[14]</sup>和视频分割<sup>[15]</sup>的研究.

综上所述,行为建模技术得到了人们的广泛研究,但是,对基于判别模型的行为建模和识别方法开展的研究工作还相对较少,这一问题即为本文的主要研究内容.

## 3 基于判别式随机场的行为建模

本文要解决的问题是在联机情况下对给定的运动序列识别出该运动所属的行为类别.基于数据驱动的行为建模方法,运动数据以帧的向量形式存在,多个连续的帧数据组成一个行为,行为与行为之间存在过渡.因此,模型应该具备同时获取行为内在和外在的动态特性.如图1所示,传统的CRF模型建立了类别间的关系,将其用于行为建模能够获取行为间的动态特性;经典的HCRF模型引入隐藏状态表示观察的子结构关系,将其用于行为建模能够获取行为内部的动态特性;本文将CRF和HCRF相结合建立Frame-HCRF,以实现同时对行为内和行为间的动态特性同时建模.

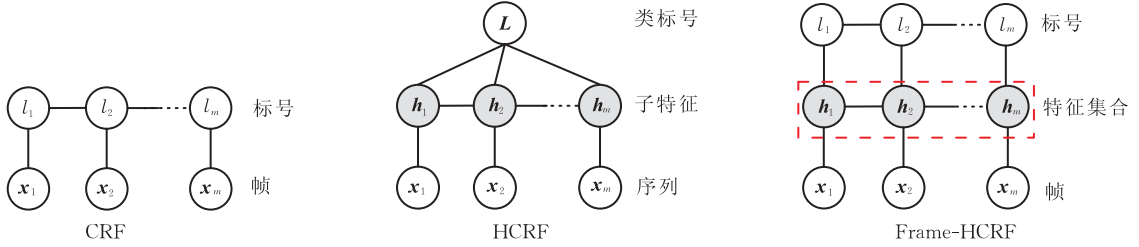


图 1 条件随机场(CRF)、隐藏条件随机场(HCRF)、帧-隐藏条件随机场(Frame-HCRF)图形模型

### 3.1 帧-隐藏条件随机场模型

在图 1 的 CRF、HCRF 和 Frame-HCRF 图形模型中  $x_j$  表示第  $j$  个观察,  $h_j$  是  $x_j$  相应的隐藏状态序列,  $l_j$  是  $x_j$  所属行为类的标号. 图形模型给出了所有观察  $x_j$  的连接关系, 因此上述模型能够表示序列在时间上的依赖关系. 为了实现人体行为的识别, 即学习从观察行为序列  $\mathbf{X} = \{x_1, x_2, \dots, x_m\}$  到分类标号序列  $\mathbf{L} = \{l_1, l_2, \dots, l_m\}$  的映射关系. 与 CRF 和 HCRF 不同的是, 在 Frame-HCRF 中每一个  $l_j$  表示运动序列第  $j$  帧所属的行为标号. 每一帧观察  $x_j$  用一个特征向量  $\phi(x_j) \in \mathbf{R}^d$  表示, 如人体骨骼各个关节在每帧中的旋转角度向量. 对于每一个序列, 定义行为子状态的隐藏变量  $\mathbf{H} = \{h_1, h_2, \dots, h_m\}$ ,  $h_m$  表示一个行为特征模板. 根据上述定义的条件模型为

$$p(\mathbf{L}|\mathbf{X}, \theta) = \sum_{\mathbf{H}} P(\mathbf{L}|\mathbf{H}, \mathbf{X}, \theta) P(\mathbf{H}|\mathbf{X}, \theta) \quad (1)$$

其中  $\theta$  是模型中的参数.

为了降低 Frame-HCRF 模型训练和推断复杂度, 假定与各类标号相关的隐藏状态集合不相交, 即每一个  $h_j$  仅表示与类标号  $l_j$  相关的隐藏状态集合  $H_{l_j}$  的一个成员, 则对于任意  $h_j \notin H_{l_j}$  都有  $P(\mathbf{L}|\mathbf{H}, \mathbf{X}, \theta) = 0$ , 模型可以表示为

$$p(\mathbf{L}|\mathbf{X}, \theta) = \sum_{\mathbf{H}} P(\mathbf{H}|\mathbf{X}, \theta) = \frac{1}{Z(\mathbf{X}, \theta)} \exp\left(\sum_k \theta_k \cdot \phi_k(\mathbf{H}, \mathbf{X})\right) \quad (2)$$

其中  $Z(\mathbf{X}, \theta)$  是分配函数, 它保证模型具有归一化概率

$$Z(\mathbf{X}, \theta) = \sum_{\mathbf{H}} \exp\left(\sum_k \theta_k \cdot \phi_k(\mathbf{H}, \mathbf{X})\right) \quad (3)$$

$\phi(\mathbf{H}, \mathbf{X})$  在本文中被定义为

$$\phi(\mathbf{H}, \mathbf{X}) = \sum_{j=1}^m f_k \quad (4)$$

其中  $f$  为模型中的特征函数, 可表示一个隐藏变量的状态函数  $f_k(h_j, \mathbf{X}, j)$ , 或一对隐藏变量关系的状态函数  $f_k(h_{j-1}, h_j, \mathbf{X}, j)$ . 在该模型中,  $f_k(h_{j-1}, h_j, \mathbf{X}, j)$  定义了一对隐藏状态的过渡函数

$$f_k(h_{j-1}, h_j, \mathbf{X}, j) = \begin{cases} 1, & h_{j-1}, h_j \text{ 是隐藏状态对} \\ 0, & \text{其它} \end{cases} \quad (5)$$

从式(5)可以看出, 如果与  $\theta_k$  相关的过渡函数作用于同一个隐藏状态子集中, 则表示行为内的内在动态特性; 如果作用于不同的隐藏状态集中, 则表示行为间的外在动态特性.

### 3.2 参数训练和行为识别

假设训练集合包含  $n$  个标号序列  $(x_i, l_i)$ , 根据 CRF 模型定义, 通过如下目标函数学习参数  $\theta^*$ :

$$L(\theta) = \sum_{i=1}^n \log P(l_i | x_i, \theta) - \frac{1}{2\sigma^2} \|\theta\|^2 \quad (6)$$

我们使用梯度计算参数最优值  $\theta^* = \arg \max_{\theta} L(\theta)$ . 在标注好行为标号的运动序列上训练该模型, 生成的分类器可以直接用来联机识别行为, 而不需要对运动数据进行预分段. 由于假定每一类标号具有不相交的隐藏状态集, 在有效改善识别效果的同时, 大大简化了训练过程.

在行为识别过程中, 对于一个行为序列  $\mathbf{X}$ , 其标号序列的最大估计为  $l^* = \arg \max P(\mathbf{L}|\mathbf{X}, \theta^*)$ , 其中  $\theta^*$  是从训练集中得到的参数. 对基于 Frame-HCRF 的行为模型, 其某一姿态  $g$  的概率等于边缘概率  $P(y_j = g | \mathbf{x}, \theta^*)$ , 该概率等价于相关隐藏状态子集  $H_g$  边缘概率的总和:

$$P(l_j = g | \mathbf{x}, \theta^*) = \sum_{\forall h_j \in H_g} P(\mathbf{H} | \mathbf{X}, \theta^*) \quad (7)$$

通过计算 Viterbi 路径, 使用最大边缘概率的方法可以估计行为序列的标号.

### 3.3 基于几何姿态的特征模板构造

条件随机场模型中一个重要的问题是为模型选择合适的特征集合, 从而用简单的特征表示复杂的行为模式. 运动数据记录的是骨骼在每一个时刻所处的全局坐标系和各个关节局部坐标系的旋转角度. 根据这些数据, 可以计算出人体各个关节在三维空间中的位置、骨骼躯干呈现的姿态、骨骼肢体的空间几何关系. 基于以上考虑, 本文定义了模型的特征模板. 通过一组向量(为简化起见忽略对运动影响较小的关节如指关节等)直观地描述人体姿态的几何特征.  $\mathbf{v}_{\text{body}}$  称为特征向量  $\mathbf{v}_{\text{up\_body}}$  和  $\mathbf{v}_{\text{down\_body}}$  的特征合

并, 所有特征被划分为两组特征集合:  $f_{\text{body}} = \{v_{\text{body}} = \langle v_{\text{up\_body}}, v_{\text{down\_body}} \rangle\}$  描述了人体躯干的姿态,  $f_{\text{limb}} = \{v_{\text{leg}} = \langle v_{\text{l\_up\_leg}}, v_{\text{l\_down\_leg}} \rangle, v_{\text{rleg}} = \langle v_{\text{r\_up\_leg}}, v_{\text{r\_down\_leg}} \rangle, v_{\text{larm}} = \langle v_{\text{l\_up\_arm}}, v_{\text{l\_down\_arm}} \rangle, v_{\text{rarm}} = \langle v_{\text{r\_up\_arm}}, v_{\text{r\_down\_arm}} \rangle\}$  描述四肢的姿态, 如图 2 所示. 定义  $L_{\text{upper}}$  和  $L_{\text{lower}}$  分别描述骨骼的上身和下身运动的参考平面.

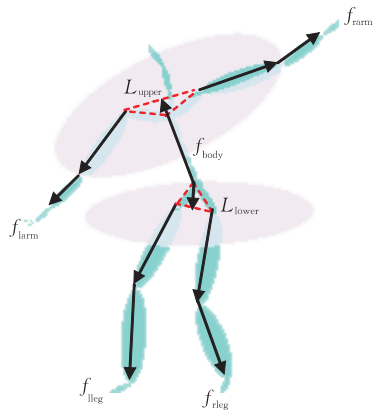


图 2 特征集及参考平面示意图

我们通过确定各个特征向量与所在局部坐标系的几何位置关系来决定骨骼的姿态. 下面以下身骨骼为例分析下肢的运动特征模板, 下身的方向向量与下肢特征向量的关系如图 3 所示. 首先以  $C_{\text{down}}$  为原点,  $v_{\text{down}}$  为 Y 轴, 参考平面  $L_{\text{Lower}}$  为 OXZ 平面, 建立上身局部坐标系, 则上肢与上身躯干的几何关系可以用它们的向量元组  $\langle v_1, v_2 \rangle$  来描述, 也称作特征模板. 如元组  $\langle v_{\text{r\_up\_leg}}, v_{\text{r\_down\_leg}} \rangle$  描述右下肢的运动姿态.

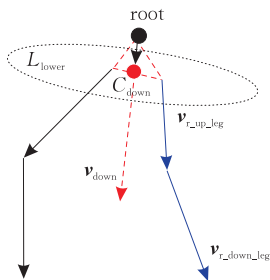


图 3 下身特征向量示意图

为了简化 Frame-HCRF 模型计算, 将行为特征以编码的形式提供给隐藏状态的输入. 定义一个关于元组  $\langle v_1, v_2 \rangle$  的编码函数  $f$ , 即

$$f(v_1, v_2) := \begin{cases} 1, & \theta > \theta_{\text{threshold}} \\ 0, & \theta \leq \theta_{\text{threshold}} \end{cases} \quad (8)$$

其中,  $\theta_{\text{threshold}}$  为控制特征向量间夹角的阈值. 这样, 使用编码函数作为过滤器处理运动数据流, 可以将复杂的运动数据转换成一系列简单的特征编码流:  $f \circ M \rightarrow \{0, 1\}^N$ .

## 4 实验与结果分析

实验使用 CRF++ 工具包对算法进行改进, 实现了基于判别随机场的人体行为模型, 并在 HP Workstation Core 2 Duo T9300 和 4GB Memory 上对本文算法进行验证. 实验所用的运动数据全部来自于 CMU 运动数据库, 用于识别的行为主要包括如下几类: 行走 (Walk Around, WA)、弯腰拾捡 (Bending and Pick, BP) 和跑后跳跃 (Run and Jump, RJ). 为了减少训练复杂度, 又不失行为特性, 对运动捕捉数据按照 10 帧/秒进行稀疏采样, 训练数据集包含约 20000 帧.

在实验中对基于 HMM 产生式模型 (GM) 和 CRF、HCRF、Frame-HCRF 判别式模型 (DM) 的行为建模方法, 以验证 Frame-HCRF 模型具有更好的行为建模和识别能力. 下面给出各模型的训练和识别方法, 根据训练数据的误差最小化原则设置模型中的隐藏状态的数目.

(1) HMM 模型 (GM). 为每一类分割好的行为序列训练一个 HMM 模型; 在识别过程中, 每一个测试序列要求在每一类 HMM 模型上进行独立测试.

(2) CRF 模型 (DM). 训练一个 CRF 链模型, 每一个行为对应一个相应的类状态; 在识别过程中, 对每一类行为标号计算边缘概率.

(3) HCRF 模型 (DM). 在所有的行为类上训练一个 HCRF 模型, 由于 HCRF 不能对行为间的动态特性进行建模, 与 HMM 类似在分割好的行为序列集上进行训练; 在识别过程中, 具有最大可能性的类标号作为识别的姿态.

(4) Frame-HCRF 模型 (DM). 使用式 (6) 目标函数训练该模型; 在识别过程中, 使用式 (7) 计算最大边缘概率.

实验采用  $K$ -折交叉验证的方法, 即每次将  $K$  个序列作为测试序列, 其余序列用作训练序列. 每一组实验在不同的窗口尺寸  $w$  下进行. 窗口尺寸为 0 ( $w=0$ ) 表示将当前帧人体各关节的速度值作为模型输入, 窗口尺寸为 2 表示将包括当前帧在内的前后共 5 帧数据用来作为模型的输入.

图 4 给出了 4 个模型在窗口尺寸为 0 时对 3 种不同行为 WA、BP、RJ 的识别 ROC 曲线, 从图中可以看出基于 Frame-HCRF 模型具有更好的行为建模能力. 实验对 4 个模型在不同窗口尺寸下识别不

同行为的相等错误率进行了统计,为了量化实验数据的一般性,为每一组实验计算了  $p$  值. 表 1 给出了识别 BP 行为的数据统计. 从表中可以看出当  $w=0$  时,4 个模型的识别效果由高到低是: Frame-HCRF、HCRF、CRF 和 HMM,且此时 3 个模型的

$p$  值都较小,则说明当  $w=0$  时实验的统计性较好;而当  $w=9$  时, HCRF 的识别效果好于 Frame-HCRF,而此时的  $p$  值为 0.4515,说明此时的实验统计性较差.

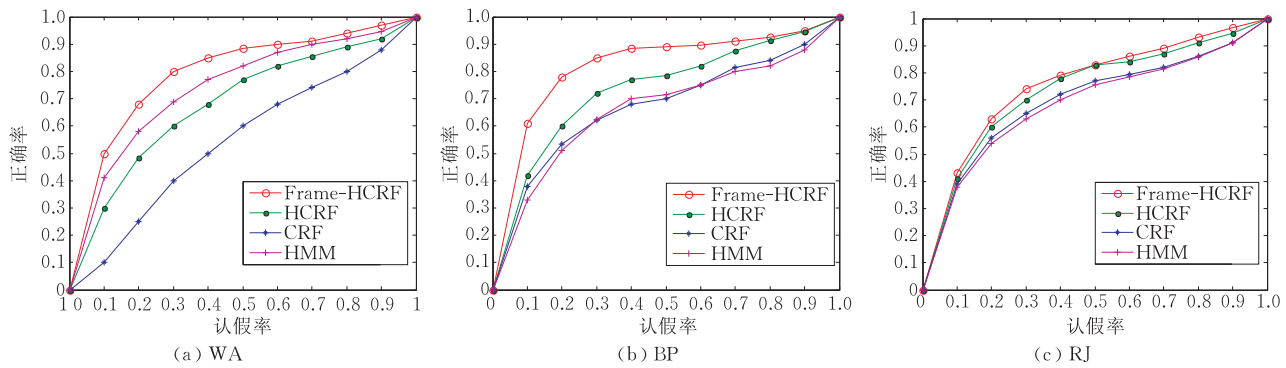


图 4 使用不同模型对 3 种行为识别的 ROC 曲线( $w=0$ )

表 1 行为 BP 在不同窗口尺寸下的识别相等错误率/%

Model	HMM	CRF	HCRF	Frame-HCRF
$w=0$	65.2( $p=0.0465$ )	66.8( $p=0.1281$ )	68.0( $p=0.0245$ )	82.7
$w=2$	63.0( $p=0.1107$ )	66.5( $p=0.0132$ )	72.2( $p=0.0691$ )	80.6
$w=5$	68.7( $p=0.1326$ )	67.8( $p=0.3231$ )	76.5( $p=0.1263$ )	79.3
$w=9$	72.0( $p=0.1647$ )	70.3( $p=0.2673$ )	79.2( $p=0.4515$ )	75.1

图 5 给出不同模型在不同窗口尺寸下的相等错误率统计图. 从图中可以看出,即使在窗口尺寸很小的情况下,Frame-HCRF 的识别性能也表现得比较出色,这是因为 Frame-HCRF 模型能够同时对行为的外在动态特性和内在动态特性进行建模. 当仅以当前帧关节的速度作为模型输入时,Frame-HCRF 模型能够处理的输入向量尺寸要比其他模型大

10~20 倍,这就使得 Frame-HCRF 模型处理联机的数据序列识别成为可能. 实验结果表明 Frame-HCRF 与其他几个模型相比,在不需要未来观测数据或需要小窗口尺寸数据的情况下,其识别准确性较高. 因此,基于 Frame-HCRF 的行为模型能够更好地执行联机行为识别.

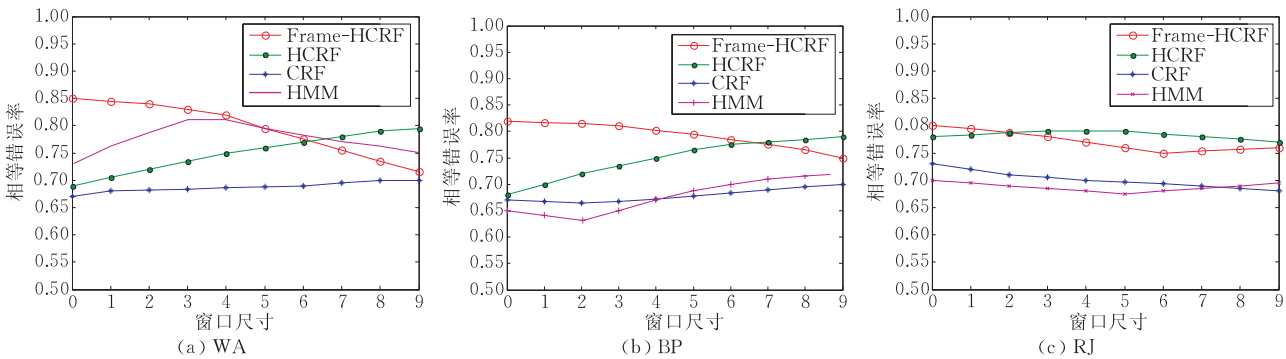


图 5 不同模型在不同窗口尺寸下对 3 种行为识别的相等错误率统计图

5 结 论

本文提出了基于帧的隐藏判别随机场行为模型框架,获取行为内在的动态特性和外在动态联系,使

其能够同时处理模型的分段和标号识别. 该模型引入了基于帧的隐藏状态变量,该隐藏状态变量可以对子结构行为特征进行建模和学习行为类标号之间的动态特性. 本文通过对比 HMM,CRF 和 HCRF 模型对未分段的行为序列进行识别,验证了本文提出的模型具有更好的建模能力和联机识别能力. 本

文的工作为人机交互领域和计算机动画领域提供了一种有效的行为识别方法,然而许多极具挑战性的问题远没有得到很好的解决,比如,高效、准确地分析地进行提取行为特征,提高模型训练的速度,更加精确鲁棒地进行运动识别等,这些问题都将是我们今后工作的重点.

**致 谢** 感谢爱丁堡大学 Andrew Smith 博士在建模过程中耐心的指导与答疑;论文实验中所用的运动捕捉数据全部来自于 mocap. cs. cmu. edu, 该数据库由 NSF EIA-0196217 项目创建!

## 参 考 文 献

- [1] Sminchisescu C, Kanaujia A, L Zhiguo, Metaxas D. Conditional models for contextual human motion recognition//Proceedings of the International Conference on Computer Vision. Beijing, China, 2005; 1808-1815
- [2] Sminchisescu C, Jepson A. Variational mixture smoothing for non-linear dynamical systems//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington D C, 2004, (2): 608-615
- [3] Wang S B, Quattoni A, Morency L et al. Hidden conditional random fields for gesture recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, USA, 2006; 521-1527
- [4] Kang E, Kevin Y, Panne M. SIMBICON: Simple biped locomotion control. ACM Transactions on Graph, 2007, 26 (3): 105
- [5] Aggarwal J K, Cai Q. Human motion analysis: A review. Computer Vision and Image Understanding, 1999, 73 (3): 428-440
- [6] Quattoni A, Collins M, Darrell T. Conditional random fields for object recognition//Proceedings of the Neural Information Processing Systems (NIPS). Vancouver, Canada, 2004; 1097-1104
- [7] Gong S, Xing T. Recognition of group activities using dynamic probabilistic networks//Proceedings of the International Conference on Computer Vision. Nice, France, 2003; 742-749
- [8] Sarah O, Victoria M, Sridhar M. Learning hierarchical models of activity//Proceedings of the IEEE/RSJ International Conference on Robots and Systems. Sendai, Japan, 2004; 891-896
- [9] McCallum A, Freitag D, Pereira F. Maximum entropy Markov models for information extraction and segmentation//Proceedings of the Annual International Conference on Machine Learning. California, USA, 2000; 591-598
- [10] Lafferty J, McCallum A, Pereira F. A conditional random fields: Probabilistic models for segmenting and labeling sequence data//Proceedings of the Annual International Conference on Machine Learning. San Francisco, USA, 2001; 282-289
- [11] Tang J, Hong M, Li J, Liang B. Tree-structured conditional random fields for semantic annotation//Proceedings of the International Semantic Web Conference. GA, USA, 2006; 640-653
- [12] Zhao Guo-Ying, Li Zhen-Bo, Deng Yu, Li Hua. Human motion recognition and simulation based on retrieval. Journal of Computer Research and Development, 2006, 43(2): 368-374 (in Chinese)  
(赵国英, 李振波, 邓宇, 李华. 基于检索的人体运动识别和模拟. 计算机研究与发展, 2006, 43(2): 368-374)
- [13] Li He-Ping, Hu Zhan-Yi, Wu Yi-Hong, Wu Fu-Zhao. Behavior modeling and abnormality detection based on semi-supervised learning method. Journal of Software, 2007, 18 (3): 527-537 (in Chinese)  
(李和平, 胡占义, 吴毅红, 吴福朝. 基于半监督学习的行为建模与异常检测. 软件学报, 2007, 18(3): 527-537)
- [14] Zhou Jun-Sheng, Dai Xin-Yu, Yin Cun-Yan, Chen Jia-Jun. Automatic recognition of Chinese organization name based on cascaded conditional random fields. Acta Electronica Sinica, 2006, 34(5): 804-809 (in Chinese)  
(周俊生, 戴新宇, 尹存燕, 陈家骏. 基于层叠条件随机场模型的中文机构名自动识别. 电子学报, 2006, 34(5): 804-809)
- [15] Chu Yi-Ping, Zhang Yin, Ye Xiu-Zi, Zhang San-Yan. Adaptive video segmentation algorithm using hidden conditional random fields. Acta Automatica Sinica, 2007, 33(12): 1252-1258 (in Chinese)  
(褚一平, 张引, 叶修梓, 张三元. 基于隐条件随机场的自适应视频分割算法. 自动化学报, 2007, 33(12): 1252-1258)



**HUANG Tian-Yu**, born in 1979, Ph. D., lecturer. Her research interests include computer graphics, computer animation, human motion simulation.

**SHI Chong-De**, born in 1979, Ph. D. candidate. His research interests include machine translation and natural language processing.

**CHENG Cheng**, born in 1966, Ph. D., associate professor. His research interests include human computer interaction, computer aided design and virtual manufacturing.

**LI Feng-Xia**, born in 1953, professor. Her research interests include virtual reality and computer simulation.

Background

This work is partly supported by the National Natural Science Foundation of China under grant Nos. 60773046, the Advanced Research of General Armament Department Foundation of China under Grant Nos. 2220061084, and Beijing Key Discipline Program. Human behavior modeling is the main supporting technique in applications of intelligent Human-Machine interface, Human-Machine training system, intelligent monitoring system, security system etc. It is an important issue belonging to the field of virtual reality, computer graphics and human interaction. By the statistic of recent five-year Siggraph conferences, the study of data-driven virtual human motion has become popular. Most of the research focused on the generative methods, such as hidden-markov models, non-linear dynamics system and their extensions.

Caused by the conditional independence assumption and markov distribution assumption, those models can't represent the contextual dependences and feature overlapping of time series data such as human motions. Discriminative Models is the most effective approach to solve the dependencies and overlapping problem in machine translation. This paper proposes an online behavior recognition based on Discriminative Random Fields. A Frame-HCRF was extended to model behaviors for frames of motion data. The behavior intrinsic dynamics were captured by CRF structure, and the extrinsic dynamics by hidden feature functions. This model can accommodate motion data online processing with unknown further frames. The experiments prove that the proposed model for human behavior modeling and recognition.