

doi: 10.11830/ISSN.1000-5013.201601011



# 结合金字塔模型和随机森林的 运动捕获序列语义标注

徐 濛, 彭淑娟, 柳欣

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

**摘要:** 针对原始运动捕获数据结构复杂、语义模糊的问题, 提出一种结合金字塔模型和随机森林的运动捕获序列语义标注方法. 首先, 利用概率主成分分析将运动序列划分为具有特定语义的运动片段. 然后, 将运动片段的欧拉角数据转换为人体各个关节点的三维空间位置坐标数据, 统一骨骼长度, 提取运动数据的 2 种互补性几何特征, 并分别归一化. 再次, 运用傅里叶时间金字塔模型构建运动片段完整的时空特征. 最后, 利用已训练的随机森林分类器对各个运动片段进行标注. 结果表明: 该方法能够对具有不同语义的复杂运动序列进行有效标注, 且可用于不同表演者, 具有一定的实用性和通用性.

**关键词:** 语义标注; 概率主成分分析; 傅里叶时间金字塔; 随机森林

**中图分类号:** TP 391

**文献标志码:** A

**文章编号:** 1000-5013(2017)06-0848-06

## Motion Capture Sequence Semantic Annotation Via Pyramid Model and Random Forests

XU Meng, PENG Shujuan, LIU Xin

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

**Abstract:** According to the complexity and semantic ambiguity within original motion capture data, we presents an effective motion capture sequence semantic approach via pyramid model and random forests. Firstly, utilize probabilistic principal component analysis to segment motion sequences into several motion clips with certain semantics. Then, the Euler angler data of each motion clip are transformed into three-dimensional space coordinates of each human joint, and the bone lengths are unified. Subsequently, two complementary features are extracted and normalized. Accordingly, the Fourier temporal pyramid model is adopted to represent the spatiotemporal characteristics of motion clips. Finally, the trained random forests classifier is employed to label each motion clip. The proposed approach is able to well annotate complex motion sequences effectively and can be applied to different performers. The experimental results show that it has certain practicality and generality.

**Keywords:** semantic annotation; probabilistic principal component analysis; fourier temporal pyramid; random forests

随着数字媒体技术的发展及光学运动捕获设备的广泛应用, 人们可以方便地获得大量的运动捕获

**收稿日期:** 2015-01-06

**通信作者:** 彭淑娟(1982-), 女, 讲师, 博士, 主要从事计算机视觉与计算机动画的研究. E-mail: pshujuan@hqu.edu.cn.

**基金项目:** 国家自然科学基金资助项目(61202298, 61300138); 福建省自然科学基金资助项目(2014J01239, 2015J01656); 华侨大学高层次人才科研启动项目(14BS207); 华侨大学中青年科研提升计划(ZQN-PY309); 华侨大学研究生科研创新能力培育计划资助项目(1400414009)

数据.这些运动捕获数据被广泛地应用于三维动画、影视制作、游戏、教育等领域.同时,大规模的运动捕获数据库也应运而生,动画师从这些数据库中快速、高效地获取满足需求的运动数据片段用于动画创作.有效地对运动数据进行分类,并进行有组织的数据库管理,首先,需要对一段未知的运动片段进行识别和标注.近年来,关于运动识别和标注已经有大量的研究成果<sup>[1-7]</sup>.现阶段大多数方法仅对特定语义的运动片段进行分类识别,而在运动捕获过程中,为了保持动作的连贯性,捕获到的数据通常包含多个连续的动作.传统的手动分割及标注方法往往需要花费大量的人力和时间.因此,有效地对一段复杂运动序列进行自动语义标注,对于后续的存储、编辑和重用具有极其重要的意义.鉴于此,本文提出一种针对复杂运动捕获序列的自动语义划分及标注方法.

## 1 运动捕获序列语义标注方法

结合金字塔模型和随机森林进行运动捕获序列语义标注,其流程如图1所示.流程主要分为训练阶段和测试阶段.在训练阶段,对训练集中的运动语义片段进行统一骨架预处理,提取2种互补性特征并分别归一化,构建傅里叶时间金字塔模型,对时间序列进行对齐,得到完整的时空特征向量.在测试阶段,运用概率主成分分析方法(PP-

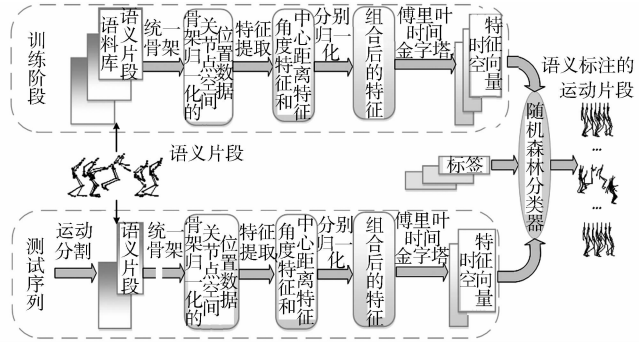


图1 文中方法流程图

Fig.1 Flow chart of our method

CA)将一段具有多个语义的运动序列进行语义划分,得到具有特定语义的子片段集合,通过与训练阶段相同的处理,得到各个语义片段的特征向量,根据训练阶段得到的随机森林模型,对语义片段进行标注.

### 1.1 人体运动数据表示形式

采用HDM05<sup>[8]</sup>数据库中标准的ASF/AMC文件格式.其中,ASF文件表示人体骨架模型.该模型有31个关节点,各个关节点采用树形层次结构组织.人体在空间中的位置由各个关节点的自由度(DOF)决定,并记录在AMC文件中,可用矩阵 $S=[s_1, s_2, \dots, s_m]^T$ 表示,  $s_i \in \mathbf{R}^{62}$ ,  $m$ 为运动序列帧数.

### 1.2 运动序列分割

对于原始运动序列,采用PPCA<sup>[9]</sup>方法将其划分为不同语义的运动片段.假定一段运动序列为 $S=[s_1, s_2, \dots, s_m]^T$ ,  $S \in \mathbf{R}^{m \times 62}$ ,  $m$ 为运动序列帧数,具体分割过程有以下6个步骤.

**步骤1** 计算运动序列的均值,即 $\bar{S} = \frac{1}{m} \sum_{i=1}^m s_i$ ,并构造矩阵 $D=[d_1, \dots, d_m]$ .其中, $d_i = s_i - \bar{S}$ .

**步骤2** 奇异值分解矩阵 $D$ ,即 $D=U\Sigma V^T$ .其中, $U, V$ 为单位正交矩阵; $\Sigma$ 为由元素 $\sigma_i$ 组成的非负递减对角矩阵.

**步骤3** 将 $\Sigma$ 前 $r$ 个向量 $V_r=(v_1, v_2, \dots, v_r)$ 作为基向量,剩余 $(n-r)$ 维奇异值的均方为 $\sigma^2 = \frac{1}{62-r} \sum_{i=r+1}^m \sigma_i^2$ .

**步骤4** 对整个运动序列进行高斯建模,均值为 $\bar{S}$ ,协方差矩阵为

$$C = \frac{1}{m-1} (WW^T + \sigma^2 I)^{1/2} = \frac{1}{m-1} V \Sigma^* V^T.$$

上式中: $W=V_r(\Sigma_r^2 - \sigma^2 I)^{1/2}$ ,  $V_r$ 为矩阵 $V$ 的前 $r$ 列,  $\Sigma_r$ 为矩阵 $\Sigma$ 的左上方的 $r \times r$ 的块矩阵; $\Sigma^*$ 为用 $\sigma$ 取代丢弃的奇异值得到的新的矩阵.

**步骤5** 求运动序列的马氏距离,即 $M = \frac{1}{T} (d_i - \bar{d}) T C^{-1} (d_i - \bar{d})$ .

**步骤6** 根据跳变原则,对步骤5获得的马氏距离曲线进行分割,得到运动序列的子运动片段 $S_i \in \mathbf{R}^{m_i \times 62}$ .其中, $m=[m_1, m_2, \dots, m_{\text{cut}}]$ ,  $m_i$ 为第 $i$ 段运动序列的帧数, cut为分段数.

### 1.3 语义片段预处理

由于AMC格式存储的运动捕获数据是由欧拉角表示的,首先,将其转换为笛卡尔坐标系下各个关节点的空间位置信息,数学描述为 $P=[p_1, p_2, \dots, p_m]^T$ ,  $p_i \in \mathbf{R}^{3 \times j}$ ,  $j$ 为关节点数目,  $m$ 为运动序列帧数.

为了便于比较不同表演者之间运动姿态,采用 Shum 等<sup>[10]</sup>提出的方法,将运动数据用统一的骨架结构表示.假定  $p^a, p^b$  分别表示关节点  $a, b$  在空间的位置,且  $b$  是  $a$  的父节点,则  $a, b$  关节点之间的单位向量为  $u^a = \frac{p^a - p^b}{|p^a - p^b|}$ .节点  $a$  新的位置坐标为

$$p_{\text{new}}^a = p_{\text{new}}^a + u^a \times D(a, b).$$

上式中:  $p_{\text{new}}^b$  为节点  $b$  新的位置坐标;  $D(a, b)$  是节点  $a$  和  $b$  之间骨骼的标准长度.由此可见,关节点新的位置坐标由其父节点决定.因此,计算时往往从根节点开始计算.

基于全局坐标的关节点位置数据包含人体的位置、朝向和运动路径等信息,而部分信息在姿态比较过程中存在冗余.此外,数值特征不能解决逻辑相似和运动相似的语义鸿沟,即逻辑相似的运动,数值上不一定相似.因此,提取角度特征<sup>[11]</sup>和中心距离特征<sup>[12]</sup>,以组合表示每一帧的运动姿态,组合后的特征向量为  $g = [g_1, g_2, \dots, g_{20}]^T$ .二者之间存在互补关系,能更全面地表现运动细节信息.但由于 2 种特征从不同角度提取信息且量纲不同,因此,分别进行归一化,得到特征  $g^* = [g_1^*, g_2^*, \dots, g_{20}^*]^T$ .

1.4 运动序列标注

1.4.1 傅里叶时间金字塔 由于采样速度或表演风格的不同,单个运动片段在时间上往往是未对齐的.参考文献[13]的方法,利用傅里叶时间金字塔模型,将各个运动片段进行对齐.为了获取一段运动片段的时间信息和全局傅里叶系数,递归地将运动片段划分为金字塔,对所有的划分片段使用短时傅里叶变换(STFT).序列最终可以表示为所有划分片段的短时傅里叶变换系数的时序组合,如图 2 所示.

对于特征向量  $g^* = [g_1^*, g_2^*, \dots, g_{20}^*]^T$ ,每一个元素  $g_i^*$  可看作时间函数,记为  $g_i^*(t)$ .对于每层金字塔时间片段,对其中的元素  $g_i^*(t)$  采用短时傅里叶变换,即  $\text{STFT}_{g_i^*}^u(t', u) = \int_l [g_i^*(t) \cdot W(t - t')] \cdot \exp(-j2\pi ut) dt$ .其中,  $W(t) = (1/\sqrt{2\pi}) \exp(-t^2/2)$  为高斯窗函数,获取其傅里叶系数,将低频系数作为特征.则特征  $g^*$  转化为每层傅里叶金字塔的低频部分的组合,即  $x \in \mathbb{R}^N$ ,  $N$  为特征维度.

1.4.2 随机森林分类 随机森林<sup>[14]</sup>(RF)旨在将多个弱分类器融合为一个强分类器,能够很好地解决多分类问题.同时,RF 算法将 Bagging 和随机选择分裂特征结合,能够有效处理高维数据且不会导致过拟合.因此,采用随机森林分类器对待测运动片段进行分类.

随机森林分类过程,如图 3 所示.随机森林是决策树  $\{h(\mathbf{X}, \theta_k)\}$  的集合,  $\mathbf{X}$  为输入向量,  $\theta_k$  是独立同分布的随机向量,决定单棵树的生长.假定训练集为  $T = \{(x_i, y_i)\}$ ,  $x_i$  为特征向量,  $x_i \in \mathbb{R}^N$ ,  $y_i$  为类别标签,  $y_i \in \mathbb{R}$ ,  $i = 1, 2, \dots, n$ ,待测样本  $x_t \in \mathbb{R}^N$ .具体有以下 3 个步骤.

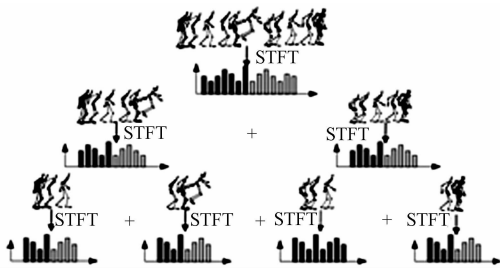


图 2 傅里叶时间金字塔

Fig. 2 Fourier temporal pyramid

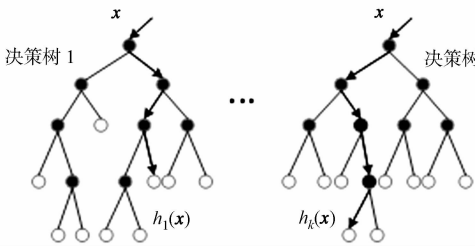


图 3 随机森林分类器

Fig. 3 Random forests classifier

步骤 1 采用给定权重的方法建立随机向量模型  $\theta$ .

步骤 2 构建随机森林分类器. 1) 对原始训练集  $T$  进行 Bootstrap 抽样,生成训练集  $T_i$ . 2) 使用  $T_i$  通过 CART 算法生成一棵不减枝的决策树  $h_i$ . a) 从  $N$  个特征中随机选取  $N_{\text{try}}$  个特征; b) 在树的每个节点上,依据 Gini 指标从  $N_{\text{try}}$  个特征选取最佳分裂特征作为节点; c) 分裂直至树生长到最大. 3) 循环 1), 2) 步,直至建立  $k$  棵决策树,树的集合为  $\{h_i\}, i = 1, 2, \dots, k$ .

步骤 3 利用随机森林分类.通过得到的决策树集合投票决定待测样本类别  $f(x_t) = \text{majority vote } \{h_i(x_t)\}_{i=1}^{N_{\text{try}}}$ .其中,  $h_i(x_t)$  为决策树  $h_i$  对  $x_t$  的分类.随机森林泛化误差的上界为

$$\text{PE}^* \leq \bar{\rho}(1 - s^2)/s^2.$$

上式中:  $\bar{\rho}$  为决策树之间相关度  $\rho$  的平均值;  $s$  为决策树  $h(\mathbf{X}, \theta_k)$  的分类效能. 为了提高 RF 的分类准确率, 在减少树与树之间相关度的同时, 增大单棵树的分类效能.

1.5 文中算法

结合金字塔模型和随机森林对运动捕获序列进行语义标注. 输入: 训练集运动片段  $\{\mathbf{tr}_1, \mathbf{tr}_2, \dots, \mathbf{tr}_n\}$ , 测试运动序列  $\mathbf{test}$ . 输出: 加标签的语义片段  $\{\mathbf{te}_1, \mathbf{te}_2, \dots, \mathbf{te}_l\}$ . 具体有以下 6 个步骤.

**步骤 1** 预处理. 将训练集中的每个运动片段的欧拉角数据转化为关节点空间位置坐标,  $\mathbf{tr}_i = \{\mathbf{R}^{93 \times \text{frame}_{tr_i}}\}$ , 重新计算各个关节点位置,  $\mathbf{tr}_i^* = \{\mathbf{R}^{93 \times \text{frame}_{tr_i}}\}$ . 提取角度特征  $\mathbf{g}_{\text{angle}} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_8]^T$  和中心距离特征  $\mathbf{g}_{\text{dist}} = [\mathbf{g}_9, \mathbf{g}_{10}, \dots, \mathbf{g}_{20}]^T$ . 分别归一化  $\mathbf{g}_{\text{angle}}^* = \frac{\mathbf{g}_{\text{angle}} - \min(\mathbf{g}_{\text{angle}})}{\max(\mathbf{g}_{\text{angle}}) - \min(\mathbf{g}_{\text{angle}})}$ ,  $\mathbf{g}_{\text{dist}}^* = \frac{\mathbf{g}_{\text{dist}} - \min(\mathbf{g}_{\text{dist}})}{\max(\mathbf{g}_{\text{dist}}) - \min(\mathbf{g}_{\text{dist}})}$ , 并组合为特征矩阵  $\mathbf{g}^* = \mathbf{g}_{\text{angle}}^* \cup \mathbf{g}_{\text{dist}}^* = [\mathbf{g}_1^*, \mathbf{g}_2^*, \dots, \mathbf{g}_{20}^*]^T$ .

**步骤 2** 构建傅里叶时间金字塔模型. 递归地将特征矩阵  $\mathbf{g}^*$  按时间顺序划分为金字塔, 对每层金字塔时间片段中的每个特征函数  $g_i^*[t]$  使用短时傅里叶变换. 将得到低频系数作为特征并组合, 特征矩阵转化为特定长度的时空特征向量  $\mathbf{x}_{tr_i} \in \mathbf{R}^N$ .

**步骤 3** 加标签. 经过步骤 1~2, 训练集数据可以表示为矩阵  $\{\mathbf{x}_{tr_1}, \mathbf{x}_{tr_2}, \dots, \mathbf{x}_{tr_n}\}$ , 加入类别标签  $y_{tr_i}$  为  $\{\mathbf{x}_{tr_i}, y_{tr_i}\}$ .

**步骤 4** 训练模型. 假设决策树的数量为  $k$ . 对  $\{\mathbf{x}_{tr_i}, y_{tr_i}\}$  进行训练, 可得到随机森林模型  $\{h_i\}, i=1, 2, \dots, k$ .

**步骤 5** 运动序列语义划分. 使用 PPCA 方法将测试序列  $\mathbf{test}$  划分为具有不同语义的子运动片段组合  $\{\mathbf{te}_1, \mathbf{te}_2, \dots, \mathbf{te}_l\}$ .

**步骤 6** 语义片段标注. 使用步骤 1, 2 对语义片段进行处理, 得到各片段时空特征向量, 则语义片段可以表示为  $\{\mathbf{x}_{te_1}, \mathbf{x}_{te_2}, \dots, \mathbf{x}_{te_l}\}$ . 使用步骤 4 训练好的随机森林模型对  $\{\mathbf{x}_{te_1}, \mathbf{x}_{te_2}, \dots, \mathbf{x}_{te_l}\}$  进行分类, 可得到各个语义片段  $\{\mathbf{te}_1, \mathbf{te}_2, \dots, \mathbf{te}_l\}$  的类别标签.

2 结果及分析

2.1 实验数据选取及描述

为了验证文中方法的有效性, 从 HDM05 数据库中选取 408 段不同长度的运动片段组成运动捕获数据语料库, 并按照不同的运动风格分为 12 个基本类别, 如表 1 所示. 表 1 中: 每类运动包含 15~57 段不同的运动片段, 每个片段包括动作从开始到结束的完整过程, 但可以多次重复, 如“移动”、“击拳”、“单脚跳跃”等类型的运动片段包含了 2~3 次相同的动作.

表 1 运动捕获数据语料库  
Tab. 1 Corpus of motion capture data

编号	名称	描述	编号	名称	描述
C1	锻炼	肘部向膝盖弯曲, 下蹲	C7	击拳	用左手或者右手向前或者向旁边击拳
C2	右手放置	用右手将物体放置在高处或者低处	C8	旋转手臂	用单手或者双手向前或者向后旋转
C3	单脚跳跃	用左脚或者右脚跳跃	C9	坐躺	坐在椅子或者地板上, 躺在地板上
C4	跳跃	双脚跳跃	C10	站立	从地面、椅子、桌子站起
C5	踢腿	用左脚或者右脚向前或者向旁边踢	C11	右手扔	用右手将物体向高处或者低处扔
C6	移动	从左脚或者右脚开始行走或者慢跑	C12	转身	向左或者向右转身

2.2 语义片段识别

为了评估文中语义片段识别方法的有效性, 从语料库中随机抽取 2/3 的运动片段用于训练, 其余用于测试. 在相同实验条件下, 将文中方法与 DTW<sup>[1]</sup>, HMM<sup>[4]</sup>, SVM<sup>[5]</sup> 等 3 种经典的运动识别方法进行比较. 基于 HMM 的方法, 为每类运动建立一个隐马尔科夫模型, 每个模型的隐藏状态数设置为 3, 每个隐藏状态所包含的混合高斯分布数也设置为 3, 通过 Baum-Welch 算法对训练集进行学习得到模型的各个参数. 在分段基础上计算识别率, 如图 4 所示. 测试集中每类运动中分类正确的片段数与每类运动

总的片段数之比为每个类的识别率,并用条形图表示.整个测试集中分类正确的片段数与测试集总片段数之比为总识别率( $\eta$ ),并用线段表示.

由图 4 可知:对于各类语义片段,当随机森林分类过程中决策树的数量设置为 120 时,文中方法均能取得较好的识别正确率,而基于 SVM 的方法在惩罚项  $c$  设置为 100,核函数半径  $g$  设置为 0.001 时,在某些类别中也能取得较好的分类结果.由于基于 DTW 的相似度匹配方法,只注重运动序列的局部缩放,在全局缩放及统一缩放尺度下效果不佳,且计算比较耗时.基于 HMM 的方法由于需要提前指定隐藏状态,而对不同的运动指定相同数量的状态对其识别结果造成了一定影响.此外,文中方法总的识别率分别为 0.54,0.82,0.91,0.96,可以看出文中方法明显优于其他 3 种方法,具有更好的识别效果.

2.3 运动序列语义标注

为了验证具有多个运动语义的运动序列的标注效果,选取 9 个复杂运动序列用于测试分段识别效果,如表 2 所示.这 9 个动作序列来自 3 个表演者 BD,BK,DG.每个运动序列均包含多个动作,最短的序列长度为 2 469 帧,包含 16 个不同的动作,最长的序列长度为 8 527 帧,包含 18 个不同的动作.9 个复杂运动序列分割和识别的查准率( $P(D)$ )和查全率( $R(D)$ ),如表 3 所示.

为了便于对比分析,手动分割和标注相应复杂运动序列,将查准率和查全率分别定义为

$$P(D) = \frac{|A(D) \cap M(D)|}{|A(D)|}, \quad R(D) = \frac{|M(D) \cap A(D)|}{|M(D)|}.$$

上式中: $M(D)$ 为手动标注的结果; $A(D)$ 为自动标注的结果.

在一段连续的运动序列中,由于动作的起始和结束位置一般不能精确确定,因此,以分段层次的查准率和查全率定义识别结果,只考虑自动标注与手动标注的结果重叠部分,当自动标注与手动标注有重叠且为相同分类时,则认为该片段正确标注.首先,使用相同的分割算法 PPCA 对运动序列进行分段;之后,采用在运动片段识别中识别率较高的 HMM,SVM 与文中方法进行比较.

由表 3 可知:文中方法能够得到较理想的查准率和查全率,且更接近真实标注结果.同时,用文中方法对最短和最长的运动序列进行标注,耗时分别为 19,56 s,与手动分割和标注方法相比,可节省大量时间,具有一定的时间效率和可用性.

表 3 查准率和查全率比较

Tab. 3 Comparison of precision and recall

编号	$P(D)$			$R(D)$		
	HMM	SVM	文中方法	HMM	SVM	文中方法
01_BD	0.81	0.87	0.93	0.75	0.92	0.94
02_BK	0.94	0.95	0.95	1.00	1.00	0.97
03_DG	0.95	0.89	0.89	0.95	0.74	0.82
04_BD	0.90	0.60	0.71	1.00	0.81	0.85
05_BK	0.75	0.90	0.90	0.83	0.72	0.82
06_DG	0.89	1.00	1.00	0.70	0.86	0.92
07_BD	0.78	0.83	0.83	0.55	0.84	0.88
08_BK	0.67	0.91	1.00	0.56	0.95	1.00
09_DG	0.89	0.82	0.88	0.75	0.81	0.88
平均值	0.84	0.86	0.90	0.79	0.85	0.90

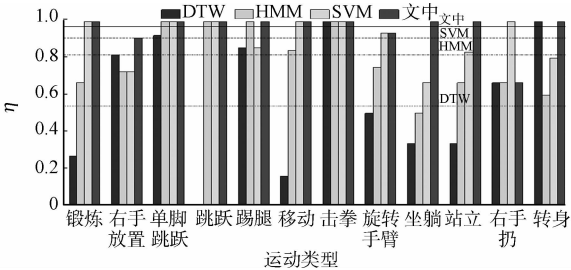


图 4 4 种方法识别率比较

Fig. 4 Comparison of recognition rate in four methods

表 2 复杂运动捕获数据序列

Tab. 2 Complex motion capture data sequences

编号	帧数	片段数	运动类别
01_BD	8 523	18	移动、转身
02_BK	7 601	19	移动、转身
03_DG	6 987	17	移动、转身
04_BD	3 429	16	击拳、踢腿
05_BK	3 590	17	击拳、踢腿
06_DG	3 234	16	击拳、踢腿
07_BD	2 921	16	移动、转身、右手放置
08_BK	3 112	15	移动、转身、右手放置
09_DG	2 469	15	移动、转身、右手放置

### 3 结束语

针对复杂运动捕获序列标注,基于分段识别的思想,首先,使用概率主成分分析方法确定动作边界临近的过渡帧,将运动序列自动划分为具有特定语义的运动片段.然后,提取2种典型的几何特征,在进行数据降维的同时,更好地保留了单帧运动姿态信息.最后,结合傅里叶时间金字塔模型和随机森林进行自动标注,对运动片段进行对齐,并加入时间信息,进一步提高了方法的识别率.该方法存在一些不足:1)由于一段运动序列中相邻运动片段间的过渡区域往往存在歧义,概率主成分分析方法不能精准地分析此类区域中运动帧的具体语义归属,从而给后序标注造成影响;2)进一步提高方法的时间效率,同时,将实验扩展到连续在线运动的实时分割及标注.

### 参考文献:

- [1] ADISTAMBHA K, RITZ C H, BURNETT I S. Motion classification using dynamic time warping[C]// Proceedings of the Multimedia Signal Workshops. Cairns: IEEE Press, 2008: 622-627.
- [2] LYU Fengjun, NEVATIA R. Recognition and segmentation of 3D human action using HMM and multi-class ada-boost[C]// Proceedings of the European Conference on Computer Vision. Graz: Springer Press, 2006: 359-372.
- [3] ZHU Hongli, DU Pengying, XIANG Jian. 3D Motion Recognition based on ensemble learning[C]// Proceedings of the International Conference on Image Analysis for Multimedia Interactive Services Workshops. Santorini: IEEE Press, 2007: 1-4.
- [4] XIA Lu, CHEN C C, AGGARWAL J K. View invariant human action recognition using histograms of 3D joints [C]// Proceedings of the International Conference on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2012: 20-27. DOI: 10.1109/CVPRW.2012.6239233.
- [5] BENGALUR M D. Human activity recognition using body pose features and support vector machine[C]// Proceedings of the International Conference on Advances in Computing, Communications and Informatics. Mysore: IEEE Press, 2013: 1970-1975. DOI: 10.1109/ICACCI.2013.6637484.
- [6] HAN Lei, WEI Liang, WU Xinxiao, *et al.* Human action recognition using discriminative models in the learned hierarchical manifold space[C]// Proceedings of the International Conference on Automatic Face and Gesture Recognition. Amsterdam: IEEE Press, 2008: 1-6.
- [7] LI Chuanjun, ZHENG Soqing, PRABHAKARAN B. Segmentation and recognition of motion streams by similarity search[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2007, 3(3): 79-82. DOI: 10.1145/1236471.1236475 • Source: DBLP.
- [8] MULLER M, ROER T, CLAUSEN M, *et al.* Documentation: Mocap da-tabase HDM05[EB/OL]. [2016-03-02]. <http://www.mpi-inf.mpg.de/resources/HDM05>.
- [9] BARBI V C J, SAFONOVA A, PAN J, *et al.* Segmenting motion capture data into distinct behaviors[C]// Proceedings of the International Conference on Graphics Interface. London: Canadian Human-Computer Communications Society, 2004: 185-194.
- [10] SHUM H, HO E S. Real-time physical modelling of character movements with Microsoft kinect[C]// Proceedings of the Symposium on Virtual Reality Software and Technology. Toronto: ACM Press, 2012: 17-24.
- [11] 杨跃东, 王莉莉, 郝爱民, 等. 基于几何特征的人体运动捕获数据分割方法[J]. 系统仿真学报, 2007, 19(10): 2229-2234. DOI: 10.3969/j.issn.1004-731X.2007.10.022.
- [12] 彭淑娟. 基于中心距离特征的人体运动序列关键帧提取[J]. 系统仿真学报, 2012, 24(3): 565-569.
- [13] WANG Jiang, LIU Zicheng, WU Ying, *et al.* Learning actionlet ensemble for 3D human action recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(5): 914-927.
- [14] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5-32.

(责任编辑: 钱筠 英文审校: 吴逢铁)