

一种结合双特征的运动捕获数据行为分割方法

彭淑娟 柳欣

(华侨大学计算机科学与技术学院 厦门 361021)

摘要 运动捕获数据行为分割的目的是将长序列数据划分为单个运动类型短片段的序列集合,使集合中每个片段具有特定的运动语义。针对相邻运动片段的过渡区间存在部分运动帧序列的语义归属歧义,提出了一种结合双特征的运动捕获数据行为分割方法。该方法首先从原始数据中提取角度和距离两组不同类型的运动特征集,并分别基于PPCA方法构建规格化的综合特征函数;然后利用子区间标准差阈值限定方法分别对综合特征函数进行粗分割,从而将运动捕获数据划分为若干具有独立语义特性的可信区域与待定区域;最后采用高斯混合模型方法判别待定区域的具体归属,从而得到最终的分割结果。实验结果表明,该算法能对模糊歧义区域进行细分割,具有较好的分割效果。

关键词 运动行为分割,双特征,综合特征函数,可信区域,待定区域

中图法分类号 TP391 **文献标识码** A

Double-feature Combination Based Approach to Motion Capture Data Behavior Segmentation

PENG Shu-juan LIU Xin

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

Abstract The objective of motion capture data behavior segmentation is to divide the original long motion sequence into several motion fragments, and each motion fragment incorporates a particular semantic behavior. In general, the transition parts of some neighboring motion fragments are always encountered with the semantic ambiguity. To this end, this paper presented a double-feature combination based approach to tackle this problem. The proposed approach first extracts two different types of motion features, i. e., angle, distance, and then utilizes the PPCA algorithm to construct two different comprehensive characteristic functions individually. Subsequently, a subinterval standard deviation approach associated with threshold limiting strategy is employed to segment the comprehensive characteristic functions into several confidence regions and pending regions roughly. Finally, by utilizing the Gaussian mixture model to further determine the pending regions, the robust segmentation result can be obtained. The experimental results show that the proposed approach performs favorably compared to the state-of-the-art methods.

Keywords Motion behavior segmentation, Double feature, Comprehensive characteristic functions, Confidence region, Pending region

1 引言

运动捕获(Mocap)是一种获取真实运动数据的有效方法,指通过传感设备记录运动物体在三维空间中的运动轨迹,并将其转化为抽象的运动数据,该方法可广泛用于驱动虚拟物体或者人进行运动仿真。由于运动捕获设备购买和使用价格昂贵以及数据捕获过程费时费力等因素,使用者常常需要将长序列的高维数据进行有效划分,以便于运动捕获数据的重用、编辑、修改以及合成等应用,从而达到生成满足不同运动场景需求的动作序列。作为运动捕获数据分析的基础,运动分割技术将运动序列中具有不同语义类型的片段进行分段、分类,并将其划分为若干单个运动类型的短片段序列集合,这对后续运动行为分析有着非常重要的作用。近年来,如何快速准确地对各种行为运动捕获数据进行有效的语义分

割,已成为研究人员广泛关注的焦点之一。

在运动捕获数据中,人体运动数据的应用最为广泛。有关其行为语义分割方法已出现在公开的文献和技术资料中,国内外比较有代表性的工作包括:王天树等人^[1]针对原始数据,使用最小描述长度准则约束下的HMM模型对连续运动序列进行建模和聚类,实现对运动序列的提取、分割和标记。该算法只能提取运动序列中的基本行为单元,而不能较好区分不同行为片段之间的过渡区域。P. Beaudoin等人^[2]针对整体数据序列,提出一种基于字符串运动主题的标记算法用于运动压缩与分割。该算法设定由用户控制运动主题的长度和单个主题中包含的运动数目。然而,此种设定不一定能保证提取出的运动片段精准匹配特定的运动语义。

针对原始捕获数据维度较大容易引起“维数灾难”的问题,J. Barbic等人^[3]分别采用主成分分析(PCA)和概率主成

到稿日期:2012-10-02 返修日期:2012-12-27 本文受国家自然科学基金项目(61202298,61202297,61102163)资助。

彭淑娟(1982—),女,博士,讲师,CCF会员,主要研究方向为人体动画、图形处理和统计模式识别,E-mail: pshujuan@hqu.edu.cn;柳欣(1982—),男,博士,讲师,主要研究方向为计算机视觉、模式识别和数据挖掘。

分分析(PPCA)算法以及混合高斯模型(GMM)对长运动序列进行行为分析,以达到其运动行为分割的目的。PCA 降维法和 PPCA 方法旨在寻找高维运动数据在低维空间的相似性,并通过对低维空间数据集的直接研究来实现原始数据的行为分析。GMM 的应用可对不同行为片段进行有效建模和描述,并能够检测出各个行为片段之间的跳变点。然而,此方法需预先人工指定合适的分类个数,并对每个行为的具体信息判别有所欠缺。对于一组较为复杂的运动捕获数据序列,具有特定行为的运动片段不一定仅仅属于同一分类,有可能呈现出在不同类别之间的规律跳变。为了检测这种跳变,杨跃东等人^[4]将原始运动序列表示为字符运动串形式,并对其聚类,进而对运动串进行序列分析和语义标注,从而提取出所有属于同一分类以及在各个分类之间规律跳变的运动片段来完成长运动序列中不同行为片段的检测。

鉴于实际应用中使用的动作捕获数据具有片段小、离散以及相邻片段间存在语义关联等特性,D. Bouchard 等人^[5]提出了一种基于模板运动分析(Laban Movement Analysis, LMA)的运动数据行为分割方法。该方法利用神经网络算法对 LMA 序列进行深层语义特性挖掘,使对相邻片段边界的分割具有鲁棒性。肖俊等人^[6]提出了一种三维人体运动特征可视化与交互式运动分割方法。该方法通过定义一组人体主要骨骼之间的几何特征来表示原始三维人体运动序列,并将其转换为一维特征信号进行不同行为片段的划分。K. Kaho 等人^[7]首先使用一个动态层次数据结构表示人体运动数据,进而采用贝叶斯分类器进行运动分割。这两种方法都采用交互手段实现运动分割,但算法不一定能保证行为分割的高查准确率。此外,G. N. Pradhan 等人^[8]提出一种双阈值多维分割算法用于周期性行为数据的划分,其将复杂人体行为序列自动分解为一个简单动态线性模型序列。近期,D. Endres 等人^[9]提出一种基于贝叶斯绑定(Bayesian Binning, BB)的无监督分割算法。该方法利用分段多项式模型结合贝叶斯绑定策略定义一个基准数据划分原则,其能够自动测定分割点的个数和位置。此算法建立在将人体行为视为可清晰分离成离散原子序列的假定基础上,可扩展性受到一定限制。

总体上说,以上方法大部分采取对整体捕获数据序列建模分析,将高维数据映射到低维度空间或者采取单类特征描述其运动特征。然而,相邻运动片段间的过渡区域往往存在模糊歧义,单一方法或者单类特征往往不能较好地精准分析此类区域中部分运动帧序列的具体语义归属。为了提高模糊区域的分割精度,本文提出一种双特征相结合的运动行为分割算法。该算法首先提取两组不同类特征对原始运动序列进行简化表示并结合规格化的综合特征函数进行运动粗分割,然后利用高斯混合模型方法判别模糊区域的具体归属,进行细分割,得到了较好的效果。

2 双特征选择

2.1 人体模型

本文采用图 1 所示的人体骨架模型,该模型是 CMU 人体运动数据库的标准骨架,总共包括 31 个关节点,各关节点采用树形结构组织,其中臀为 ROOT 根节点,|J|表示关节点的集合,用数学可描述为 $|J|=31$ 。

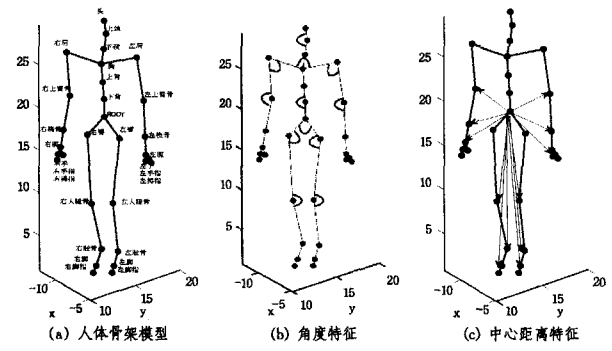


图 1

2.2 角度特征与中心距离特征

传统的三维数据特征无法解决运动的逻辑相似和数据相似之间的语义鸿沟。例如,对于两段不同速率的走路运动片段,后一个运动要比前一个运动快 5 倍,两个运动逻辑相似,但其运动特征并不一定相似。人体运动时常常呈现一种规律性的形态,特别是关键部位(如手、脚、关节等)的运动空间变换具有特定规律。一般情况下,不同的运动类型呈现不同的空间变换规律,仅单类的运动特征有时无法全面地反映运动细节信息。例如,人在跑动的过程中,手臂的前后摆动和弯曲程度呈现规律性的变化,两腿前后摆动和弯曲程度也呈现交替性的规律变化,此时关节间的夹角的变化可以反映运动姿态的行为变化;针对绕动手腕的运动序列,姿态变化时关节夹角值保持不变,而主要关节点到 ROOT 中心的距离则随时间发生变化。

为了使提取到的运动特征更全面地表现细节信息,本文提取两组不同的特征来表征人体运动。类似文献^[6]的方法,第一组特征为骨段角度特征,定义 B_i^j 为第 i 帧第 $\lfloor j/2 \rfloor$ 角度特征的骨段。如图 1(b)所示,本文提取 12 个重要的角度特征。对于一段运动序列的给定帧 i ,其角度特征的计算公式如下:

$$a_i^k = \frac{\vec{B}_i^{2k} \cdot \vec{B}_i^{2k+1}}{\|\vec{B}_i^{2k}\| \|\vec{B}_i^{2k+1}\|}, k=1,2,\dots,12 \quad (1)$$

特征向量集合 $a_i=(a_i^1, a_i^2, \dots, a_i^{12})$ 表示运动姿态序列中第 i 帧的角度特征。这里约定角度 a_i^k 取值范围在 $[0 \ \pi]$ 之间。

第二组特征为中心距离特征,如图 1(c)所示,分别提取 ROOT 根节点到上肢和下肢 12 个关键点的距离。给定帧 i ,中心距离特征向量 $d_i=(d_i^1, d_i^2, \dots, d_i^{12})$ 的表示方式如表 1 所列。

表 1 中心距离特征的计算方式

上肢	下肢
d_i^1 : ROOT→左上臂骨	d_i^1 : ROOT→左大腿骨
d_i^2 : ROOT→左桡骨	d_i^2 : ROOT→左胫骨
d_i^3 : ROOT→左腕	d_i^3 : ROOT→左脚
d_i^4 : ROOT→右上臂骨	d_i^4 : ROOT→右大腿骨
d_i^5 : ROOT→右桡骨	d_i^5 : ROOT→右胫骨
d_i^6 : ROOT→右腕	d_i^6 : ROOT→右脚

3 基于 PPCA 的综合特征函数构造

M. E. Tipping 等人^[10]提出的概率主成分分析(Probabilistic PCA, PPCA)是 PCA 和隐变量模型的结合。PPCA 将

PCA 中丢弃的维度信息作为高斯噪声进行参数估计和重新利用,以得到最佳概率模型。本节基于 PPCA 来构造综合特征函数。

3.1 运动帧的表示

对于 m 帧的高维人体运动序列,角度与中心距离特征均可表示成维度为 12 的样本特征序列 $T=(t_1, t_2, \dots, t_m)$, T 为运动帧的全体。将动作每一帧的特征向量 t_i 看作随机向量,根据文献[10]中 PPCA 的假定,随机样本 t 可以由向量 z 按照下式产生:

$$t=Wz+\mu+\epsilon \quad (2)$$

式中, z 为隐含向量; W 为 $12 \times q$ 权重矩阵; μ 为总体样本 M 的均值; ϵ 为噪声,且服从各项同性的高斯分布 $\epsilon \sim N(0, \epsilon I_d)$ 。这里,隐含向量 z 为表现系数,且假定其服从多元高斯分布 $z \sim N(0, I_q)$, I 为单位矩阵。则:

$$P(t) \sim N(t, \mu, WW^T + \epsilon I) \quad (3)$$

PPCA 是概率化的 LPCA,因此权重矩阵 W 对应的 q 个分量即对应 PPCA 中的 q 个主成分。根据文献[10]中的结论,PPCA 对 $P(t)$ 中参数 ϵ 和 W 的最大似然估计值为:

$$\epsilon = \frac{1}{12-q} \sum_{i=1}^{12} \lambda_i \quad (4)$$

$$W = \sum_q (V_q - \epsilon I_q)^{1/2} R \quad (5)$$

式中, $\sum_q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_q)$ 是基于样本 $t_1, t_2, \dots, t_m$ 估计得到的协方差矩阵的前 q 个特征值; $V_q = (v_1, v_2, \dots, v_q)$ 为 q 个特征向量组成的矩阵; R 为任意 $q \times q$ 正交的矩阵,为简化运算,在本文中取 $R=1$ 。

3.2 综合特征函数构造

对于给定的 12 维样本 T ,选取 $q(q < 12)$ 个基向量构成子空间的基信息,将主成分外的 $12-q$ 维信息作为高斯噪声进行参数估计。随机向量 $t_i (i=1, \dots, m)$ 的中心点为:

$$\bar{t} = \frac{1}{m} \sum_{i=1}^m t_i \quad (6)$$

并基于 PPCA 进行降维,构造综合特征函数 Z 。为了减低噪声对后续分割的干扰,对 Z 进行高斯滤波去噪,生成新的规格化的综合特征函数 Z' 。具体算法步骤如下:

1) 定义均值零化的特征向量:

$$\bar{G}_i = t_i - \bar{t} \quad (i=1, 2, \dots, m) \quad (7)$$

2) 基于特征向量 $\bar{G}_i$ 生成样本相关矩阵:

$$C_{(12 \times 12)} = A^T \cdot A \quad (8)$$

其中, $A^T = [G_1^T, \dots, G_m^T]$, $A = U \Sigma V^T$, 矩阵 U 和 V 单位正交。

3) 计算相关矩阵 C 的特征值和特征向量,并按特征值大小排序:

$$\Sigma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{12}) \quad (9)$$

且 $\lambda_1 > \lambda_2 > \dots > \lambda_{12} > 0$, 其对应的特征向量分别为 $V_{(12 \times 12)} = (v_1, v_2, \dots, v_{12})$ 。

4) 投影原特征向量 G 到特征子空间,生成新的主成分特征向量 H :

$$H_{(m \times 12)} = G_{(m \times 12)} \cdot V_{(12 \times 12)} \quad (10)$$

其中,对于 $H_{(m \times 12)}$ 中的特征向量 $H = (h_1, h_2, \dots, h_{12})$, H 中各个分量按照其对应特征值从大到小排序。

5) 构造综合特征函数:

$$Z = \sum_{i=1}^q \omega_i h_i + \frac{\epsilon}{12-q} \sum_{i=1}^{12} h_i \omega_i \quad (11)$$

约定前 q 个特征向量对应的权重如下:

$$\omega_i = \frac{\lambda_i}{\sum_{i=1}^{12} \lambda_i} \quad (i=1, 2, \dots, q) \quad (12)$$

其中,高斯参数 ϵ 由最大似然估计值估算得到。

6) 高斯低通滤波器对 Z 进行去噪、平滑处理:

$$Gauss = \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{z^2}{2\sigma_2^2}} \quad (13)$$

其中, σ_2 为高斯噪声中的标准偏差。

7) 规格化得到新的综合特征函数 Z' ,使其值域在 $[0, 180]$ 之间。由于角度特征和距离特征的值域区间不同,为了方便后续的融合分割,将综合特征函数 Z 规格化为标量,使之在 $[0, 180]$ 区间。定义规格化函数如下:

$$Z'(t) = 180 \frac{Z(t) - \text{Min}(Z(t))}{\text{Max}(Z(t)) - \text{Min}(Z(t))}, t \in [1; m] \quad (14)$$

例如,一段 1600 帧的运动捕获数据序列 Seq1^[11] (来源于 CMU 人体运动数据库,编号 14_06 的前 1600 帧)表示的动作行为语义为“跳动摆臂、慢跑、蹲坐”。通过角度特征和特征距离计算,得到如图 2 所示的综合特征函数波形(横轴为帧数,纵轴为规格化标量)。综合波形图 2 中角度和距离特征的直观分析可知,运动中发生了 2 次剧烈的变化,分别发生在 500 帧和 1000 帧附近,可粗分割为 3 段不同语义的运动序列,大致发生于区间 $[0, 500]$ 、 $[500, 1000]$ 和 $[1000, 1600]$ 之间。在 $[0, 500]$ 帧之间,可以明显看出有 3 个波峰和 2 个波谷,曲线变化非常有规律,并且波峰和波谷的数值都非常接近,变化的速度也相近,推测前面这段形成 3 个波峰和波谷的运动序列是属于同一类型运动;在 $[500, 1000]$ 帧之间,2 组特征的曲线呈现平稳的态势;而对于 $[1000, 1600]$ 之间的运动序列,单一的距离特征不能很好地分割,而观察发现其角度特征的曲线变化却非常有规律。通过人工观察测试运动序列发现,测试运动序列与我们分析波形图的情况基本吻合:前一段运动“跳动摆臂”、“跳动摆臂”运动序列具有强烈的震荡性,体现在手臂的反复上举和下放,属于剧烈运动;中间一段运动“慢跑”序列,其主要体现在两腿的弯曲程度上,两腿的弯曲程度呈现周期性的小幅震荡,对应于综合特征曲线图 $[500, 1000]$ 区间曲线的平稳小幅度的波动曲线;最后一段运动“蹲坐”序列,其主要体现在两腿间夹角有规律的周期变化,而距离特征变化不明显,对应角度综合特征曲线图中 $[1000, 1600]$ 区间曲线的有规律周期波动。这表明:综合特征函数计算所得的两组特征能够较好地反映动作序列语义关系,且能够满足对不同运动分割的互补,建立起了运动数据与运动语义之间的联系桥梁。

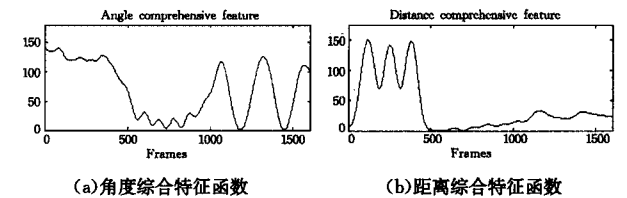


图 2 规格化的综合特征函数

4 运动行为分割

本文提出的运动行为分割方法主要包括以下两个步骤:

1) 利用子区间标准差阈值限定方法分别对角度与距离综合特征进行粗分割,并以此得到分割出的可信区域与待定区域;

2)针对待定区域的语义归属歧义,利用 GMM 方法进行细分割,从而得到最终的分割结果。

4.1 运动粗分割

定义 设 A 和 B 是初分割区域集合, f_i 是运动序列中的第 i 帧, M 为运动序列全体, $A \subset M, B \subset M$, 则它们的并 $A \cup B$ 、交 $A \cap B$ 、差 $A - B$ 和 A 的补 $\bar{A}$ 分别定义为如下帧区间:

$$\begin{aligned} A \cup B &= \{f_i | f_i \in A \vee f_i \in B\} \\ A \cap B &= \{f_i | f_i \in A \wedge f_i \in B\} \\ A - B &= \{f_i | f_i \in A \wedge f_i \notin B\} \\ \bar{A} &= M - A = \{f_i | f_i \notin A\} \end{aligned} \quad (15)$$

运动粗分割是将输入的运动序列全体 M 分割为可信区域与待定区域两大类。

- A)可信区域指根据两种特征得到的分割子区域的交集;
- B)待定区域指可信区域的补集。

对于规格化的角度与距离综合特征函数 Z' , 本文提出一种基于子区间标准差阈值限定方法对两种特征序列进行粗分割, 得到两组不同的子区间划分结果。子区间的交集即为可信区域, 可信区域的补集为待定区域, 其中可信区域与待定区域交替出现。步骤如下:

1)计算 Z' 的局部极大值与极小值序列集合, 分别记为 $zMax, zMin$, 其中极大值与极小值集合中元素个数分别记为 K_{max}, K_{min} 。

2)处理极大值集合 $zMax$, 其中需要给定阈值 $tval$ 参数。

①首先初始化一个空数组 $zMaxSeg = \text{NULL}$, 用于存储极大值分割区间, 设置初值 $i=1, index+i=1$, 将集合 $zMax$ 中的第一个元素和最后一个元素分别加入到 $zMaxSeg$ 。

②遍历 $zMax$ 。如果 $index+i \leq K_{max}$, 计算 $zMax$ 中第 $index$ 到第 $index+i$ 个元素的标准差 $zstd$ 和均值 $zmean$ 。如果标准差 $zstd$ 和均值 $zmean$ 相减的绝对值大于阈值 $tval$, 将 $zMax$ 中第 $index+i-1$ 和第 $index+i$ 分别加入到 $zMaxSeg$ 中, 置 $index=index+i$ 且 $i=1$; 否则 $i=i+1$ 。

③ $index+i \leq K_{max}$ 时, 重复步骤②, 否则继续。

④得到全部局部极大值分割区间 $zMaxSeg$ 。

3)按照与步骤 2)同样的方式(阈值 $tval$ 参数设置与 2)相同), 得到全部局部极小值分割区间 $zMinSeg$ 。

4)基于 $zMaxSeg$ 和 $zMinSeg$ 重新挑选分割区间, 得到综合特征函数 Z' 的初始分割区间 $zSeg$ 。

①如果 $zMaxSeg$ 与 $zMinSeg$ 分出来的区间相等, 则合并两个区间重新排序: $zTmpSeg = \text{sort}(zMaxSeg, zMinSeg)$, 将集合 $zTmpSeg$ 中的其它元素按等间隔 N 帧采样 (N 取小于 5), 得到分割区间 $zSeg$; $zTmpSeg$ 第一个元素加入到 $zSeg$ 中, 遍历集合。

②否则, 以 $zMaxSeg$ 和 $zMinSeg$ 区间长度的最大值作为分割区间 $zSeg$ 。

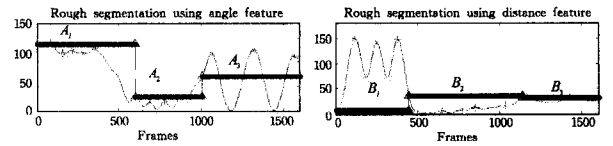
5)随机分割 $zSeg$ 中小区间段, 与最近邻区间段合并, 得到一组新的分割子区域 $zSeg$ 。

图 3 为运动序列 Seq1 的 2 组特征的初分割结果。图 3 (a)为角度综合特征函数初分割的 3 个子区域集合:

$$A_1 = [1 \ 601], A_2 = [602 \ 1009], A_3 = [1010 \ 1600]$$

图 3(b)为距离综合特征函数初分割的 3 个子区域集合:

$$B_1 = [1 \ 447], B_2 = [448 \ 1140], B_3 = [1141 \ 1600]$$

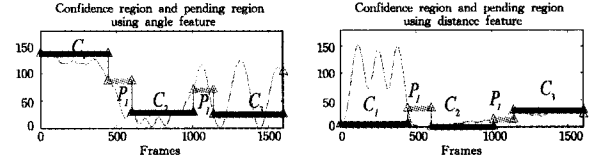


(a)角度综合特征分割子区域 (b)距离综合特征分割子区域

图 3 分割子区域

根据前面的定义, 得到图 4 所示的 3 个可信区域 C_1, C_2, C_3 (黑色段) 和两个待定区域 P_1, P_2 (绿色段):

$$\begin{aligned} C_1 &= A_1 \cap B_1 = [1 \ 447] \\ C_2 &= A_2 \cap B_2 = [602 \ 1009] \\ C_3 &= A_3 \cap B_3 = [1141 \ 1600] \\ P_1 &= [448 \ 601] \quad P_2 = [1010 \ 1140] \end{aligned} \quad (16)$$



(a)角度综合特征 (b)距离综合特征

图 4 可信区域与待定区域

4.2 待定区域细分割

混合高斯模型(GMM)是聚类算法和概率密度拟合中常用的模型, 其基本假设是认为待研究数据可以由数个高斯分布生成。由于 Gaussian 分布在计算上有很好的性质以及 GMM 本身可以变得任意复杂, 通过增加混合模型的个数可以任意地逼近任何连续的概率分布, 因此 GMM 在实际中应用广泛。每个 GMM 由 M 个 Gaussian 分布组成, 这些高斯分布线性组合在一起形成 GMM 的概率密度函数:

$$P(X|\theta) = \sum_{i=1}^M w_i g(X|\mu_i, \Sigma_i) \quad (17)$$

式中, $P(X|\theta)$ 是 d 维观测值 X 在 GMM 模型 θ 下的概率密度函数值, w_i 是高斯分布混合时的权重值, $g_i(X) (i=1, 2, \dots, M)$ 是第 i 个单个高斯分布概率密度函数, 其形式为:

$$g(X|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)\right] \quad (18)$$

式中, μ_i 是均值矢量, Σ_i 是协方差矩阵。另外, 混合权重值满足 $\sum_{i=1}^M w_i = 1$ 。高斯混合模型参数 θ 由各个单高斯模型的均值矢量、协方差矩阵和混合权重值共同组成:

$$\theta = \{w_i, \mu_i, \Sigma_i\}, i=1, 2, \dots, M \quad (19)$$

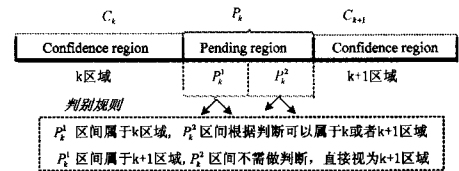


图 5 待定区域细分割方法

本文将相邻的运动可信区域用一个高斯混合模型来参数化, 由中心极限定理知, 这样的参数化是合理的。即分别求出其高斯模型参数 θ_k 和 θ_{k+1} , 其表示第 k 和 $k+1$ 个可信区域。判别准则如图 5 所示, 对于给定的一个待定区域运动序列, 我们将其均分为两段, 对于任意一段 $X = \{x_1, x_2, \dots, x_L\}$, 假设

其观测序列之间相互独立,可算出此序列属于两个相邻可信区域的概率,即后验概率:

$$P(X|\theta_k)=\prod_{i=1}^I P(x_i|\theta_k) \tag{20}$$

首先采用 k-means 方法初始化模型参数值,然后结合期望极大化方法(EM)迭代估计最大似然模型参数 θ 。约定角度与距离特征的参数分别记为 θ^A 、 θ^D ,对于任意一段待定区域 X 属于可信区域 k 还是 $k+1$,我们采取表 2 中的 4 种不同情形进行判定。

表 2 待定区域 X 划分的 4 种不同情形(k 和 $k+1$ 为两个相邻的可信区域)

情形	角度	距离	归属区域
1	$P(X \theta_k^A) \geq P(X \theta_{k+1}^A)$	$P(X \theta_k^D) \geq P(X \theta_{k+1}^D)$	k
2	$P(X \theta_k^A) < P(X \theta_{k+1}^A)$	$P(X \theta_k^D) < P(X \theta_{k+1}^D)$	$k+1$
3	$P(X \theta_k^A) \geq P(X \theta_{k+1}^A)$	$P(X \theta_k^D) < P(X \theta_{k+1}^D)$	if $\frac{P(X \theta_k^A)}{P(X \theta_{k+1}^A)} \geq \frac{P(X \theta_k^D)}{P(X \theta_{k+1}^D)}$; 判断为 k 区域; else 归属为 $k+1$ 区域
4	$P(X \theta_k^A) < P(X \theta_{k+1}^A)$	$P(X \theta_k^D) \geq P(X \theta_{k+1}^D)$	if $\frac{P(X \theta_{k+1}^A)}{P(X \theta_k^A)} \geq \frac{P(X \theta_k^D)}{P(X \theta_{k+1}^D)}$; 判断为 $k+1$ 区域; else 归属为 k 区域

4.3 算法描述

- 综合上述第 2—4 节的分析,本文算法可以描述如下:
- 1)根据式(1)和表 1 分别计算角度特征与中心距离特征;
 - 2)选取 q 个基向量,并基于 PPCA 方法分别构造基于角度和距离特征的综合特征函数;
 - 3)利用提出的区间标准差阈值限定方法,针对得到的综合特征函数,分割出可信区域与待定区域;
 - 4)对可信区域 GMM 建模并对待定区域重新划分,利用图 5 和表 2 的判别准则进行待定区间细分割,得到最终结果。

5 实验分析

模拟仿真实验采用卡麦基-梅隆大学(CMU)^[11]提供的运动数据,该数据库中人体运动捕获序列的采样频率是 120 帧/秒。实验选取两组不同运动语义且不同复杂程度的人体运动序列,并将本文算法与现有方法(PCA、PPCA)以及人工方法进行对比分析。

Exp1 对前述 Seq1 中的 1600 帧语义为“跳动摆臂、慢跑、蹲坐”的运动捕获数据进行行为分割,该序列中不同运动类别之间存在突兀跳变,如图 6 所示。

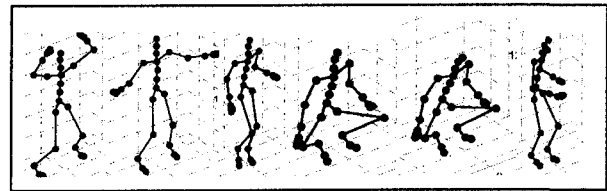


图 6 待分割的人体运动数据 Seq1

参数设置为: $q=3, \mu_2=7, trval=12, M=3$ 。其中, q 为综合特征函数构造时选取的基向量的个数, μ_2 为式(13)中高斯噪音中的标准偏差, $trval$ 为运动粗分割算法中阈值参数的设定值, M 为待定区域细分割 GMM 中高斯混合模型的状态个

数。图 7 为不同方法对 Seq1 的分割结果,4 种方法均将原始数据分为 3 个子区间段,图中白色三角框的上顶点表示运动分割点(下同)。

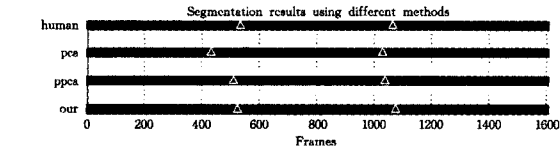


图 7 不同方法对 Seq1 的分割结果

Exp2 运用本文算法对不同运动复杂程度、不同运动姿态及不同运动类型的运动捕获序列进行分割测试。选取 2465 帧 Seq2 运动捕获数据进行分割,其语义为“跳爵士、侧扭扭、下蹲、慢跑、站立”(来源于 CMU 人体运动数据库,其编号为 13_30),如图 8 所示。

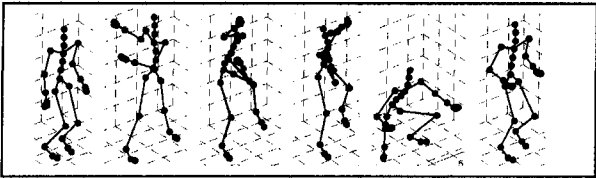


图 8 待分割的人体运动数据 Seq2

参数设置为: $q=3, \mu_2=7, trval=15, M=3$ 。图 9 为不同方法对 Seq2 的分割结果。PCA 方法将原始数据分割为 6 个子区间段,其余 3 种方法分割为 5 个子区间段。

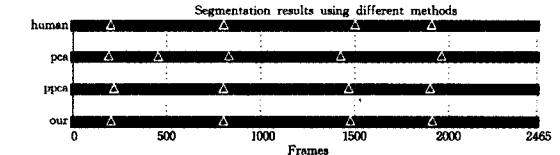


图 9 不同方法对 Seq2 的分割结果

本文以人工方法得到的分割结果作为参考模板,指标误分率 ER 用来定量衡量算法的有效性,定义如下:

$$ER=\frac{N_{miss}+N_{mddn}}{N} * 100\% \tag{20}$$

式中, N_{miss} 表示子区间段中误分多余的帧数, N_{mddn} 表示子片段中误分缺少的帧数, N 为手工分割方法得到的子区间段的总帧数。

Seq1 序列分割得到的 3 段子区间分别记为 Seq1_1、Seq1_2、Seq1_3。由于 PCA 方法将 Seq2 分割为 6 个子区间段,为了方便对比分析,这里将中间的子区间段 2 和段 3 组合为一个子区间段,组合后 Seq2 对应的 5 段子区间分别记为 Seq2_1、Seq2_2、Seq2_3、Seq2_4、Seq2_5。不同方法分割结果的误分率如表 3 和表 4 所列。

表 3 不同方法对 Seq1 分割的误分率

算法	Seq1_1	Seq1_2	Seq1_3
pca	15.5%	17.4%	6.7%
ppca	4.5%	5.4%	4.8%
our	2.6%	4.7%	3.9%

表 4 不同方法对 Seq2 分割的误分率

算法	Seq2_1	Seq2_2	Seq2_3	Seq2_4	Seq2_5
pca	3.7%	6.7%	11.4%	15.6%	6.0%
ppca	4.0%	4.6%	5.4%	5.7%	3.1%
our	3.2%	4.1%	4.6%	5.4%	3.3%

如图 7 和图 9 所示,本文算法稍优于 PCA 和 PPCA 单一特征分割方法。主要原因在于 PCA 分割方法直接对原始人体运动序列进行降维后提取其主要成分,没有考虑其运动姿态帧间的约束关系。PPCA 单一特征分割方法需结合 PCA 子空间分析和高斯建模,实现起来较 PCA 方法复杂;此方法分割的结果虽比 PCA 方法好,但对模糊歧义区间分析不够准确,导致精准度有所欠缺。相比之下,本文算法提取的双特征集能够较全面地反映运动特性,且特征之间对运动语义表达具有互补性,从而得到了较好的分割效果。

表 3 和表 4 的两组对比数据显示,本文算法对不同子区间误分率均在 6% 以内,明显小于 PCA 分割。对于中等帧长的突兀跳变数据,本文算法误分率要稍好于 PPCA;对于较复杂的运动数据,本文算法能够获得较多的精准分割区间。综合两组数据来看,本文算法分割性能整体略优于 PCA 和 PPCA 单一特征分割方法。本文算法利用双特征仔细分析运动数据,具有较好的直观性和概括性,且后续通过 GMM 方法对特定区域进行细分,提高了整体运动分割精度。

从以上对比实验可以看出,针对相邻运动片段的过渡区间,提取单类特征后分割的结果往往不够准确。相比之下,基于双特征相结合的运动分割方法可以进一步对过渡区间进行判别归属,从而达到较为满意的分割结果。本文提出的方法,可以明显地检测到不同运动序列转折的跳变分割点。对于分割精度要求不高的序列,可以直接采用本文第 4.1 节所提出的基于子区间标准差阈值限定的粗分割。对于稍微复杂的序列,通过 GMM 判别准则处理特定区间,进而分配到相邻区间,可以得到较为精准的分割点。本文提出的算法在运动分割中具有较好的适用性,但也存在如下限制:1)对于剧烈跳变的短周期运动序列可能达不到理想的分割效果;2)对于不同空间位置的同类型运动,提取得到的 2 组规格化特征近似,算法会存在一定程度的误分割。

结束语 针对相邻运动片段的过渡区间存在部分运动序列的归属歧义,本文提出了一种结合双特征与高斯混合模型的运动分割算法。该算法通过选择角度与距离双特征,利用概率主成分分析(PPCA)降维,建立规格化的综合特征函数,采用基于子区间标准差阈值限定方法将原始运动序列粗分割为可信区域与待定区域;对于待定区域,利用高斯混合模型(GMM)方法对其进行进一步判别归属,从而最终达到细分的目的,仿真实验验证了算法的有效性。同时,本文中所提出的粗分割的方法可以直接应用于对分割精度要求不是很高的场合。本文算法对于 3000 帧左右长度的运动序列可以得到

较理想的效果,但是对于过长的运动序列,分割效果相比其他方法,优势不明显。主要原因在于后续分割的数据来源于提取的两组不同类特征,这两组特征能够较好地反映人体运动的逻辑性以及相关性,处理起来简单、直观,但是由于特征信息量还有所欠缺,存在对运动细节表现能力不足的缺点,对过复杂的运动序列的分割效果不够理想。在后续研究工作中,我们拟对人体运动分割进行进一步的深入研究,主要包括:

1)结合全局与局部运动姿态序列分析建模,从空间邻域的角度考虑运动语义分割;

2)研究核空间中运动姿态的距离度量理论,用于估算运动捕获数据中不同运动姿态的相似度。

参考文献

- [1] 王天树,郑南宁,徐迎庆. 人体运动非监督聚类分析[J]. 软件学报,2003,14(2):209-214
- [2] Beaudoin P, Coros S, Panne M V D, et al. Motion-motif graphs [C] // Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation. 2008;117-126
- [3] Barbic J, Safonova A, Pan J Y, et al. Segmenting motion capture data into distinct behaviors [C] // Proc. Graphics Interface. 2004,62;185-194
- [4] 杨跃东,王莉莉,郝爱民. 运动串:一种用于行为分割的运动捕获数据表示方法[J]. 计算机研究与发展,2008,45(3):527-534
- [5] Bouchard D, Badler N. Semantic segmentation of motion capture using laban movement analysis [C] // Proc. Intelligent Virtual Agents. 2007;37-44
- [6] 肖俊,庄越挺,吴飞. 三维人体运动特征可视化与交互式运动分割[J]. 软件学报,2008,19(8):1995-2003
- [7] Kahol K, Tripathi P, Panchanathan S. Gesture segmentation in complex motion sequences [C] // Proc. IEEE International Conference on Image Processing. 2003;105-108
- [8] Pradhan G N, Li C J, Prabhakaran B. Hierarchical indexing structure for 3D human motions [C] // Proc. on Multimedia Modeling. 2007,4351;386-396
- [9] Endres D, Christensen A, Omlor L, et al. Emulating human observers with bayesian binning; segmentation of action streams [J]. ACM Trans. Appl. Percept., 2011,8(3):1544-3558
- [10] Tipping M E, Bishop C M. Probabilistic Principal Component Analysis[J]. Journal of the Royal Statistical Society; Series B(Statistical Methodology), 1999,61(3):611-622
- [11] CMU Graphics Lab Motion Capture Database[OL]. <http://mocap.cs.cmu.edu/>

(上接第 302 页)

- [4] Mazloom S, Eftekhari-Moghadam A M. Color image encryption based on coupled nonlinear chaotic map [J]. Chaos, Solitons & Fractals, 2009,42(3):1745-1754
- [5] 卢辉斌,刘海莹. 基于耦合混沌系统的彩色图像加密算法[J]. 计算机应用,2010,30(7):1812-1814,1817
- [6] 何松林. 基于混沌序列的数字彩色图像加密算法[J]. 计算机工程,2011,37(10):114-116

- [7] 卢辉斌,郑恒娜,韩秀峰. 基于 Lorenz 三维混沌序列的彩色图像加密算法[J]. 电子测量技术,2008,31(11):34-36
- [8] 张燕,黄贤武,刘家胜. 基于三维混沌系统的彩色图像加密新算法[J]. 计算机工程与应用,2008,44(20):202-205
- [9] 陈国亮. 基于混沌理论的彩色图像加密算法研究[D]. 兰州:兰州大学,2012
- [10] 王兴元,王明军. 二维 Logistic 映射的混沌控制[J]. 物理学报,2008,57(2):731-736