

基于子空间集成学习的3维人体运动识别

向坚¹⁾ 叶绿¹⁾ 朱红丽²⁾

¹⁾(浙江科技学院信息与电子工程学院, 杭州 310023) ²⁾(浙江大学城市学院信息与电气工程分院, 杭州 310015)

摘要 如何对3维运动数据进行自动识别,是提高数据利用效率和进行计算机动画创作的前提。为了有效地识别运动数据,需要把运动数据投影到非线性流型低维子空间中,先识别出运动的内在结构,然后对运动的各个关节点分别进行学习,最后基于集成学习的方法产生强的隐马尔可夫学习器,以便能够对一些常见的运动类型进行自动识别。实验结果表明,这种基于子空间集成的人体运动识别方法较传统方法的检索精度、检索速度均有较大提高。

关键词 3维运动 集成学习 子空间

中图法分类号:TP391.41 文献标识码:A 文章编号:1006-8961(2008)10-2003-04

3D Human Motion Recognition Method Based on Ensemble Learning in Subspace

XIANG Jian¹⁾, YE Lv¹⁾, ZHU Hong-li²⁾

¹⁾(School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023)

²⁾(College of Information and Electronic Engineering, Zhejiang University City College, Hangzhou 310015)

Abstract In this paper, a motion retrieval and recognition system is investigated from an ensemble learning model. In order to recognize and retrieve 3D motion data, first motion features are extracted from motion data. Due to the high dimensionality of motion's features, a generalized isomap nonlinear dimension reduction based on the estimation of underlying eigenfunction is used for training data of ensemble HMM learning. Then each action class is learned with one HMM. Since ensemble learning can effectively enhance supervised learning, ensembles of weak HMM learners are built. Experimental results show that our approaches are effective for information retrieval from large scale motion database.

Keywords 3D motion, ensemble learning, subspace

1 引言

随着运动捕获技术的发展,大规模的3维人体运动数据库日益普及,因此如何高效地利用这些数据库进行动画制作、电影特技等实际应用成为研究的热点。而要想高效地利用开发数据库,首先就需要有快速准确的检索识别技术。

本文针对基于3维捕获数据及人体运动分析的

一些关键技术进行研究,首先在成功提取了运动的特征数据^[1,2]之后,因为运动特征的高维复杂性,导致无法有效地直接对运动特征进行处理,所以第2步就需要对运动特征进行降维,即先通过有效合理的降维算法,把高维原始数据投影到一个低维的子空间里,然后通过这个低维空间揭示人体运动的内在结构。

本文采用 ISOMAP^[3]非线性流形学习的降维方法来发现嵌入在高维空间的低维坐标,从而得

基金项目:浙江省教育厅科研计划项目(20070777)

收稿日期:2008-06-20;改回日期:2008-07-05

第一作者简介:向坚(1976~),男,2007年获浙江大学博士学位,讲师。主要研究方向为统计学习、多媒体分析和检索、计算机动画等。E-mail: freexiang@gmail.com

到运动特征的低维子空间。在成功实现降维之后,即可对运动数据降维之后的低维子空间特征进行学习,也就是先构建隐马尔科夫学习器(HMM)^[4],然后利用集成学习^[5]的方法形成针对一些常用运动类型的强集成 HMM 学习器,以便能够自动快速地从人体运动数据库中识别检索出一些常见的运动。

2 ISOMAP 子空间生成

本文是采用 ISOMAP 非线性流形学习的降维方法来生成低维数据子空间,即首先使用最近邻图中最短路径得到近似的测地线距离,用以代替不能表示内在流形结构的欧氏(Euclidean)距离,然后输入到高维尺度分析(MDS)中进行处理,进而发现嵌入在高维空间的低维坐标,即可得到运动特征的低维子空间。

图 1 是两个踢腿运动用 ISOMAP 降维方法映射到 3 维子空间的运动投影,图中虚曲线表示的是侧踢腿运动,实曲线表示的是前踢腿运动。黑色矩形框表示站立姿势。

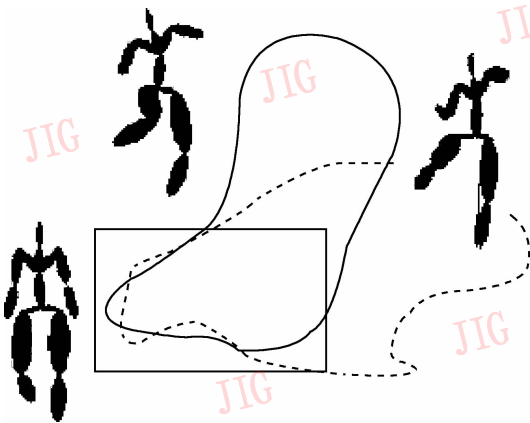


图 1 ISOMAP 方法在 3 维子空间的运动投影

Fig. 1 ISOMAP projection on subspace

3 基于集成隐马尔科夫模型运动识别

3.1 基于 3 维运动特征的隐马尔科夫模型

这里定义常见运动类型的数量为 M ,由于每个关节的几何特征都是独立的,因此可以把每个关节作为一个特征属性,将关节的数量定为 N 。本文用隐马尔科夫模型来捕捉人体运动的一些动态信息。如果对于第 k 个关节($k = 1, 2, \dots, 16$),可为

每一种运动类型学习一个隐马尔科夫模型 $I_i (i = 1, 2, \dots, M)$,那么基于第 j 个关节的运动识别就可以通过寻找有着最大 $P(\mathbf{O}|\mathbf{I})$ 值的第 i 个运动类型来实现,如下所示:

$$K(\mathbf{O}) = \operatorname{argmax}_{i:i=1,\dots,M} (P(\mathbf{O}|\mathbf{I}_i))$$

通过以上表达式, M 个运动类型和 $N (N = 16)$ 个关节就组成了一个 $M \times N$ 大小的隐马尔科夫模型矩阵,若把第 i 种运动类型的第 k 个关节的隐马尔科夫模型表示为 $HMM_{i,j}$,其对应的参数为 $l_{i,j}$,则与列 j 对应的隐马尔科夫模型集合就对应第 k 个关节的分类器。

在本系统里,每个隐马尔科夫模型都有 3 个隐藏状态,每个隐藏状态含有的混合高斯分布总数也为 3。应用隐马尔科夫模型识别时,首先通过最大期望算法(也称 Baum-Welch 算法)训练得到每个隐马尔科夫模型的以下参数:状态之间的转移概率、起始状态的概率、每个状态 s 的参数($s = 1, 2, 3$)、期望 $\mu_{s,m}$ 、协方差矩阵 $\Sigma_{s,m}$ 、每个高斯混合模型的权值 $w_{s,m} (m = 1, 2, 3)$ 。

对于前向后向算法和最大期望算法来说,都需要计算给定状态 s 在 t 时刻观测事件 O_t 的概率 $P(O_t|s_t = s)$ 。和离散隐马尔科夫模型不同,连续隐马尔科夫模型是用概率密度函数来计算 $P(O_t|s_t = s)$,因为对于连续分布来说,没有点有概率。其表达式如下:

$$\sum_{m=1}^3 \left(w_{s,m} \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_{s,m}|^{\frac{1}{2}}} e^{-\frac{1}{2}(O_t - \mu_{s,m})^T \Sigma_{s,m}^{-1} (O_t - \mu_{s,m})} \right)$$

3.2 集成学习的隐马尔科夫模型

由于人的完整运动是由所有的关节的运动信息和谐构成的,如果仅从某一个关节的运动信息去识别人体运动,则显然是不全面的,因此本文引入集成学习的方法,即通过对一个运动类型的所有关节的隐马尔科夫模型进行集成学习来得到一个能反映人体所有关节运动信息的加强型集成分类器,以便实现正确高效的运动识别。

例如给定 n 个训练样本 $(x_1, y_1), \dots, (x_n, y_n)$ 的训练集,其中 $y_i = |0, 1|$, ($i = 1, 2, \dots, n$) 对应样本识别的错误和正确,样本中有 u 个错误样本, v 个正确样本;待分类运动的每个关节的时空特征表示为 $f_j(g)$,其中 $1 \leq j \leq 16$;对于第 i 个训练样本 x_i ,它的特征为 $f_j(x_i)$ 。第 j 个关节的时空特征的弱分类器 $h_j(x)$ 由一个特征 f_j ,一个阈值 q_j 和一

个指示不等式方向的偏置值 p_j (只有 ± 1 两种情况) 构成:

训练目标:通过对正反例进行分析来选择错误率最低的 T 个弱分类器,使其最优化组合成一个强分类器。

训练过程如下:

(1) 设 $w_{t,i}$ 为第 t 次循环中第 i 个样本的误差权重。对训练样本中的误差权重按如下公式进行初始化:对于 $y_i = 0$ 的样本, $w_{1,i} = \frac{1}{2u}$; 对于 $y_i = 1$ 的样本, $w_{1,i} = \frac{1}{2v}$;

(2) for $t = 1$ to T

通过权重归一化,使得 $w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$;

(3) 对于每个关节点,通过训练出其弱 HMM 分类器来确定阈值 q_j 和偏置值 p_j , 以使得目标函数

$$e_j = \sum_{i=1}^n w_{t,j} |h_j(\mathbf{x}_i) - y_i| \text{ 达到最小;}$$

(4) 从上面的弱分类器中,找到一个具有最小错误 e_i 的分类器 h_i 对所有样本的权重进行更新:

$$w_{t+1,i} = w_{t,i} b_i^{1-e_i};$$

(5) 确定 e_i , 如果刚才 \mathbf{x}_i 已被 h_i 正确识别,则 $e_i = 0$; 反之 $e_i = 1$; 而且 $b = \frac{e_i}{1 - e_i}$ 能保证后面训练提取的弱分类器会更加强化这些识别错误样本的训练;

(6) 最后得到以下所有关节点的弱分类器集成的强分类器:

$$h(\mathbf{x}) = \begin{cases} 1 & \sum_{t=1}^T a_t h_t(\mathbf{x}) \geq 0.5 \sum_{t=1}^T a_t \\ 0 & \text{其他} \end{cases}$$

其中, $a_t = \log \frac{1}{b_t}$ 。

4 实验结果

本文是采用 Matlab 实现了基于时空特征的运动检索,并在实验室的 Mocap 设备中捕获得到的一系列运动数据的基础上展开实验的。其中用作测试的运动数据库包含有 1 500 多个运动序列,上百种运动类型。

利用单个的弱 HMM 学习器和集成学习之后

的集成 HMM 强学习器来对常见的运动进行分析 and 识别,发现集成学习器有着好得多的学习精度和识别性能。识别结果如表 1 所示。为了比较集成学习 HMM 分类器对常见运动的识别效率,本文采用常见的 25 种运动类型对其进行分析,并用传统的运动数据识别方法中的具有代表性的基于聚类的弱学习的识别方法 (CW) 来作为参考。据检索结果的分析,很显然基于子空间的集成学习方法 (SE) 对常见运动的识别有着更快的速度和更高的精度 (表 2)。图 2 显示了两种方法的识别时间比较。

表 1 集成学习的隐马尔科夫模型和弱隐马尔科夫模型的识别性能比较

Tab.1 Comparison of ensemble HMM and weak HMM recognition rate

学习模型	不同人体运动类型的识别率 (%)			
	走路	体操	舞蹈	游泳
弱隐马尔科夫学习	62.1	54.2	45.2	25.5
集成隐马尔科夫学习	98.1	91.2	85.3	79.1

表 2 查准率和查全率

Tab.2 Recall and precision

运动类型	查全率 (%)		查准率 (%)	
	CW 方法	SE 方法	CW 方法	SE 方法
走路	81	98	90	97
体操	61	89	73	91
舞蹈	55	87	65	89
游泳	51	89	49	91

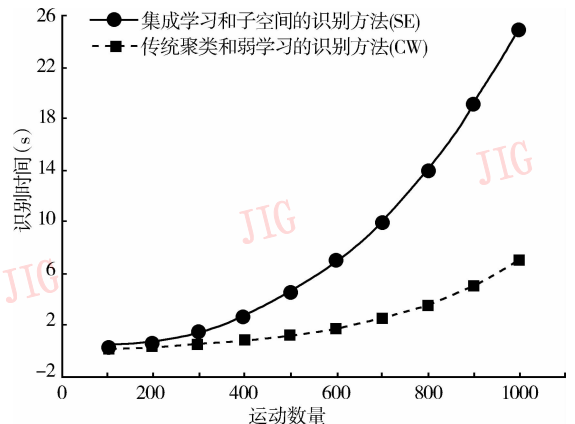


图 2 运动识别时间比较

Fig.2 Comparison of motion recognition time

由表 2 可见,SE 检索方法的检索精度高于

CW 方法,而且 CW 检索方法的检索速度和效率极大地取决于数据库的规模,随着数据库规模的扩大,其检索时间急剧增加;但是集成学习的检索效率就基本和数据库规模无关。图 3 展示了直接对原始数据进行处理和利用低维子空间的方法先对运动数据进行降维之后,再进行训练学习的时间比较。由图 3 可见,基于子空间的分析方法提取了运动的内在本征特征,由于避免了直接对原始数据进行计算,因此随运动数量的提高训练时间增加很少。

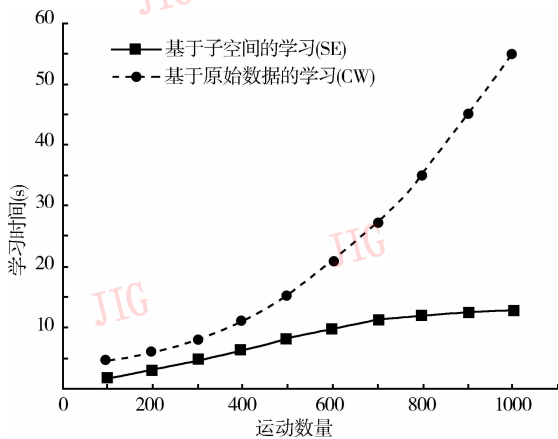


图 3 训练时间比较

Fig. 3 Comparison of training time

5 结 论

本文通过子空间降维、集成学习决策树学习,将隐马尔科夫模型引入到 3 维运动数据的检索和识别中,不仅使得能够快速完成 3 维运动数据的检索以及一些最常见的运动类型的识别,而且有利于运动数据分析的提速和后续智能化处理。

参考文献 (References)

- 1 Liu Feng, Zhuang Yue-ting, Wu Fei, *et al.* 3D motion retrieval with motion index tree [J]. *Computer Vision and Image Understanding*, 2003, **92**(3):265 ~ 284.
- 2 Lee Jehee, Chai Jin-xiang, Reitsma Paul S A, *et al.* Interactive control of avatars animated with human motion data [A]. In: *Proceedings of SIGGRAPH [C]*, San Antonio, Texas, USA, 2002: 491 ~ 500.
- 3 Tenenbaum Joshua B, Silva Vin De, Langford John C. A global geometric framework for nonlinear dimensionality reduction [J]. *Science*, 2000, **290**(5500):2319 ~ 2323.
- 4 Rabiner L. A Tutorial on hidden Markov models and selected applications in speech recognition [J]. *Proceedings of the IEEE*, 1989, **77**(2): 257 ~ 286.
- 5 Dietterich Thomas G. Machine learning research: four current directions [J]. *Artificial Intelligence Magazine*, 1997, **18**(4): 97 ~ 136.