

模板化的人体运动合成

夏贵羽¹ 孙怀江¹

摘要 为解决现有运动合成方法中控制方式过于复杂的问题, 提出一种模板化的运动合成模型, 旨在降低运动合成技术的应用门槛. 利用稀疏主成分分析 (Sparse principal component analysis, SPCA)、Group lasso 和 Exclusive group lasso 对人体运动进行建模, 使其对应的每一个低维参数只依赖于少数几个人体关节, 构成人体运动的一个内在自由度 (Degree of freedom, DOF), 并具有直观语义; 同时, 每个关节被尽量少的低维参数所控制, 以减少低维参数对彼此所控制的自由度的交叉影响. 实验表明, 通过直观地修改低维参数, 就能够实时地控制每个参数对应的摆臂幅度、踢腿高度、跳跃距离等运动属性. 这种“模板学习、模板定制”的两步方法, 有效地降低了运动合成控制的复杂度, 即便非专业人员也可以用其进行艺术创作.

关键词 运动合成, 模板化, 运动参数, 语义特征

引用格式 夏贵羽, 孙怀江. 模板化的人体运动合成. 自动化学报, 2015, 41(4): 758–771

DOI 10.16383/j.aas.2015.c140457

Templated Human Motion Synthesis

XIA Gui-Yu¹ SUN Huai-Jiang¹

Abstract Since the existing approaches to control human motion synthesis are too complicated, we propose a templated motion synthesis model to reduce the difficulty of using motion synthetic technology. We use sparse principal component analysis (SPCA), group lasso and exclusive group lasso to model human motions so that each low-dimensional parameter depends on a few human joints which form an intrinsic degree of freedom (DOF) with intuitive meanings. Meanwhile, our approach makes each joint controlled by as few low-dimensional parameters as possible to reduce the interferences between different DOFs. Our experiments demonstrate that users can control the motion features like amplitude of swing arm, kick height and jump distance by modifying the low-dimensional parameters intuitively in real time. This two-step approach of “template learning and template customization” can effectively reduce the complexity of synthesis control, and allows inexperienced users to create a realistic human animation quickly and easily.

Key words Motion synthesis, template, motion parameter, semantic feature

Citation Xia Gui-Yu, Sun Huai-Jiang. Templated human motion synthesis. *Acta Automatica Sinica*, 2015, 41(4): 758–771

近年来, 随着虚拟现实技术的快速发展, 人体运动捕获数据在影视制作、电子游戏以及仿人机器人^[1] 领域得到了广泛的应用. 但由于运动捕获设备的价格过于昂贵, 设备的使用对专业性要求过高, 以及演员能够表演的动作内容有限等原因, 只有很少的专业部门才有能力获取运动捕获数据. 而运动合成技术能够利用已有运动数据高效地生成逼真的新运动, 大大地提高了运动数据的重用程度. 但运动合成是一项非常艰巨的任务, 因为人体运动是一种复杂的现象, 即使是简化的人体骨架模型也有着很高的自由度, 而人眼又很善于发现运动中不协调的地方, 这就要求合成的运动有很高的逼真度. 另外, 运动合成的过程需要有便于实时控制的交互接口, 这

也对运动合成算法交互性能和实时性能提出了更高的要求.

目前, 数据驱动的运动合成方法已成为主流, 出现了一些具有代表性的运动合成方法, 如运动图^[2]、运动过渡^[3]、可变形运动模型^[4] 以及各种参数化的运动模型^[5–11] 等, 但这些合成方法的控制方式都较为复杂, 不适合一些面向非专业用户的应用, 如以角色动画为基础的虚拟现实游戏 *Second Life*, 其玩家不可能都具备足够的专业知识, 所以需要一些“低门槛”的运动合成方法, 以便于非专业用户在实际的应用中合成出生动的角色动画. 因此本文提出一种模板化的人体运动合成的概念, 旨在降低运动合成技术使用的门槛, 简化运动合成的控制过程, 提高控制精度和合成效果. 现有的合成方法, 尽管能够根据用户需求完成运动合成工作, 但对用户专业能力的要求高. 例如需要用户提供运动的关键帧^[2] 或是特定关节的运动轨迹^[12], 对于没有经验的用户来说很难将这些约束条件与实际合成运动的相关特性直观

收稿日期 2014-06-25 录用日期 2014-09-11
Manuscript received June 25, 2014; accepted September 11, 2014

本文责任编辑 贾云得

Recommended by Associate Editor JIA Yun-De

1. 南京理工大学计算机科学与工程学院 南京 210094

1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094

地对应起来,也就很难提供精确的约束来指导运动合成. Min 等^[4]利用可变形运动模型开发了一种适合非专业用户使用的运动合成方法,但当需要对运动合成进行细粒度的控制时,该方法将无法实时完成合成任务,并且,如果用户提供的约束不合理,将会导致最终无法合成逼真且满足约束的运动. 本文提出的模板化的运动合成方法,将稀疏主成分分析(Sparse principal component analysis, SPCA)^[13]、Group lasso^[14]以及 Exclusive group lasso^[15]结合起来,找到每一种运动内在的具有语义特征的自由度(例如踢腿高度和摆臂的幅度),并将其包装成模板化的运动参数留给用户,用户便可以根据该参数的语义描述,直观地进行实时的模板化运动合成. 该方法提供了一种新的运动合成概念和方式,即“模板化”,如上文描述的现有的运动合成方法通常从运动本身出发,其控制方式就会与运动内容相关,如需要用户提供关键帧、运动轨迹、特定关节的位置等,势必要求用户有较高的专业能力,这就限制了运动合成技术的使用范围. 而模板化运动合成的概念则是从应用的需求出发,对于一些面向非专业用户的应用,需要一些“低门槛”的运动合成方法,而本文方法提供的运动模板,只要简单地操纵每种模板对应的运动参数,就能够合成出个性化的人体运动. 当然,专业人员也可以利用运动模板制作动画,制作的过程中可以将更多的精力放在艺术创作上,而不是对运动合成的控制上. 用户使用运动模板进行运动合成的过程,就如同使用傻瓜相机拍照的过程,操作过程简单,合成内容丰富.

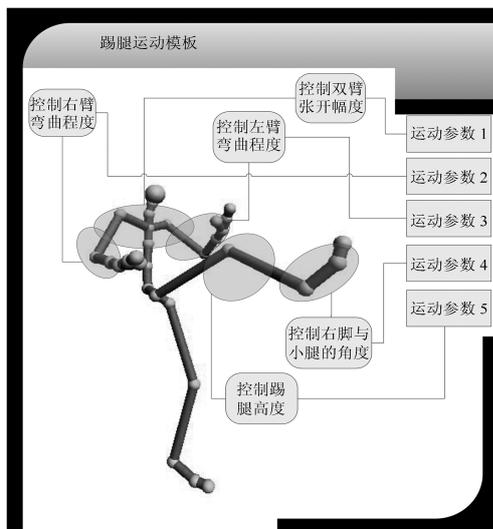


图1 踢腿运动模板示意图

Fig. 1 Diagram of a kick template

人体运动虽然复杂,但却是有规律可循的,每一种类型的运动数据在高维度的表象下,只对应着少数几个内在的自由度,而文中模型的任务就是通过

学习自动发现每一种运动的固有维度(内在自由度),并将其作为该运动模板的运动参数,以获得每一种运动的最终控制方式. 这种最佳的控制方式最终通过运动模板来实现,用户并不需要分析该运动的相关特性,因为其已经被包装成运动模板的对外接口,即运动参数;也不需要担心对模板的操作会破坏运动的逼真程度,因为模型通过对参数的调整范围的约束保证了在模板限定的框架内的运动的逼真性.

1 相关工作

人体运动作为计算机动画的一个重要分支,其在影视制作、电子游戏以及仿人机器人领域的应用越来越广泛. 随着人体运动捕获技术日趋成熟,数据驱动的人体运动合成方法已经成为计算机动画领域的主流技术,有着很好的应用前景. 然而运动合成却是一项艰巨的任务,因为人眼对于运动本身的协调性十分敏感,这就要求运动合成方法有很高的逼真性,以避免合成一些“畸形”的人体运动.

目前,基于运动图的运动合成^[2,16]是一种很有效的运动合成方法,运动图^[2]是一种表示运动数据之间关系的数据结构,节点表示运动帧,边表示不同帧之间的运动过渡. 根据用户设定的不同阶段的运动状态,在运动图中搜索最优的路径,并合成出符合要求的运动. 其优势在于对建图的运动数据的要求比较宽泛,并且能产生长度不同、过渡自然的运动,但合成运动的表达能力依赖于原运动数据,且容易造成根节点的位置计算不准确,导致脚步的滑动和朝向抖动.

参数化的运动模型^[6-7,17]在运动图的基础上,利用运动的物理特性有效地解决以上问题, Kwon 等^[17]将运动的类型、落脚点、速度、转角以及加速度等以参数的形式引入运动合成模型,通过控制这些参数指导运动的合成. Heck 等^[6]结合了运动图和运动混合的优点,将运动图的节点构造为连续参数空间,为原本非常有限的拼接组合方式带来了细粒度的控制. 参数化模型的引入,很大程度上提高了运动合成过程的可控性,但参数的内容需要预先人为指定,无法自动适应运动类型的变化.

因此,一些研究者跳出了运动图的框架,提出了自适应提取参数的运动模型,他们根据运动数据维度之间高度相关的特性,通过降维技术获得了其内在的低维结构,利用优化算法求得满足用户需求的低维变量取值,再重构回原始空间以合成符合要求的新运动^[10-12,18-20]. 例如王宇杰等^[11]先对原始运动进行密集重采样后,投影到低维流形上,然后学习低维样本点与原始运动的对应关系,形成低维参数空间到原始空间的逆映射,从而达到了通过修改低维变量完成运动合成的目的. 蓝荣祯等^[21]则利用

主测地线分析 (Principal geodesic analysis, PGA) 对运动进行建模, 根据其获取的非线性低维空间, 将风格经过编辑的运动末端曲线重构为原始运动形式, 以合成具有不同风格的人体运动. 为了提高合成精度, 也有研究者将合成过程分段进行, 选取数据库中和最近合成的运动帧相似的运动作为先验知识, 在主成分分析 (Principal component analysis, PCA) 降维后的低维参数空间, 根据约束条件, 求取一个最大后验解作为运动合成的结果^[12]. 为了能够在低维空间表示降维后运动样本的位置, 这类方法大多选择两到三维的参数空间, 为了尽可能多地保留原始数据的信息, 在维数过低前提下, 普遍以重构误差最小或保持拓扑关系为准则, 诸如 PCA 和多元尺度法 (Multi dimensional scaling, MDS) 之类的降维方法被广泛使用. 因此, 无论是降维方法直接得到的载荷向量还是通过学习得到的低维空间到原始空间的逆映射, 都和所有的人体关节有关. 所以这类方法虽然能够自动获取运动参数, 且形式统一, 但是由于每一个低维参数都依赖于原始空间中的所有变量, 控制人体的所有关节, 导致了参数缺乏语义含义.

而稀疏语义参数化模型 (Sparse semantic parametric model, SSPM)^[22] 既能够自适应地提取参数, 又使得每一个运动参数只依赖于少数几个人体关节, 从而有效地为运动参数赋予了语义含义, 但在该模型中, 同一个关节出现在过多的运动参数中, 这就导致了低维参数对彼此所控制的自由度造成交叉影响, 即一个运动特性需要多个运动参数控制, 这就增加了控制的难度, 降低了运动参数以及合成过程的可理解性, 而本文改进了稀疏语义参数化模型, 提出了模板化的运动合成方法, 该方法不仅能够根据运动类型自适应地提取有效的运动参数, 同时由于 Exclusive group lasso 的引入, 降低了同一个关节被不同运动参数所控制的可能性, 从而尽可能地做到运动参数与运动自由度一一对应.

2 模板化运动合成模型

三维人体动画是一种非常生动的多媒体艺术, 但由于运动捕获数据的获取和合成都很难, 所以导致动画制作需要专业人员才能完成, 不能像拍摄照片和录制视频一样得到普及. 模板化的运动合成模型, 可以将每一种运动制成一个运动模板, 把该运动的具有语义特性的内在自由度 (如踢腿运动的踢腿高度) 作为接口留给用户, 用户就可以简单地控制这些接口, 在已有的运动模板的基础之上, 合成符合自己要求的人体运动. 模板化的运动合成方法, 降低了运动合成技术的使用难度, 系统开发者可以先将各种类型的运动制成运动模板, 用户即可选择自己

需要的运动模板, 并在此模板上进行个性化的调节, 操作过程就如同拍照和录制视频一样简单, 以达到普及动画制作的目的.

SPCA 与 PCA 的主要区别在于降维后得到的载荷向量是稀疏的, 因而使得每一个主成分对应的变量 (在本文中也就是运动参数) 具有了语义含义. 本文中的方法在利用 SPCA 降维的前提下, 结合使用 Group lasso 和 Exclusive group lasso 进行变量选择, 可以得到 K 个“成组”稀疏的载荷向量和重构矩阵, 每个载荷向量中只有少数的几个关节对应的权重不为 0, 该载荷向量对应的变量即为一个运动参数, 控制该运动的一个内在的自由度. 稀疏载荷矩阵如图 2 所示, 一个方格表示同一关节在不同帧中的所有旋转变量, 白色表示值为 0, 具体分组方式在第 2.2 节中有详细介绍.

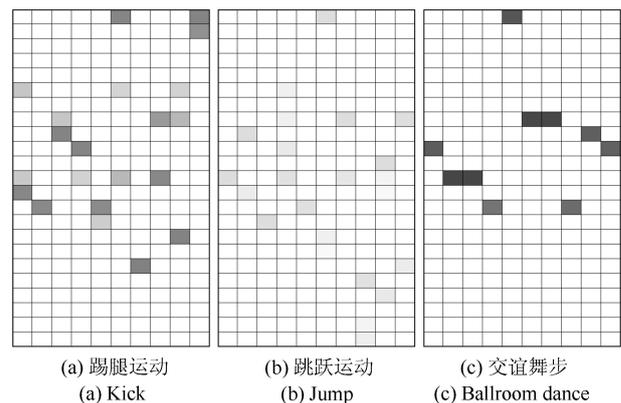


图 2 本文方法得到的三种运动的稀疏载荷向量组示意图
Fig. 2 Diagram of sparse loading matrices of the three motions using our method

在合成时, 只需要调节任意一个或多个运动参数的值, 然后将所有运动参数构成的向量作为低维参数空间的一个样本, 利用求得的重构矩阵将其重构回原始空间, 就能得到符合要求的人体运动. 由于合成的过程不需要做任何优化求解, 都是解析化的处理, 所以速度很快, 用户在调节任意运动参数时, 都能看到实时的合成结果, 使得整个合成的过程更加的简便、精确和直观. 图 3 描述了模板化运动合成的主要步骤.

2.1 SPCA、Group lasso 和 Exclusive group lasso

本文用到的降维方法是具有稀疏特性的稀疏主成分分析 (SPCA) 加上两个特征选择工具 Group lasso 和 Exclusive group lasso. 下面将对以上三种模型做简单的介绍, 为了对它们以及本文提出的模型进行统一的表示, 表 1 对几种模型共用的符号作以下定义 (文中未特别指明的向量均为列向量):

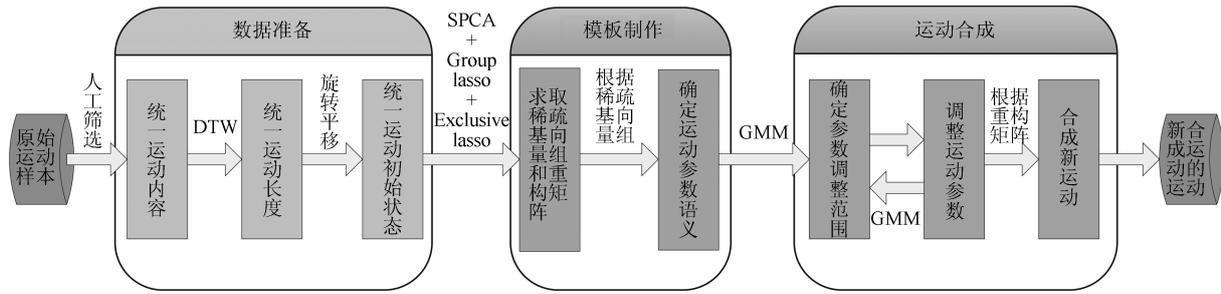


图3 模板化运动合成流程图

Fig. 3 General steps of templated motion synthesis

表1 共用符号定义

Table 1 Definitions of sharing notations

符号	含义	备注
N	样本数	
d	样本维数	
K	主成分的数目	
λ	正则化因子	
\mathcal{G}	变量的分组集合	
g	\mathcal{G} 中一个分组	$g \in \mathcal{G}$
$\ *\ _1$	向量的 L_1 范数	
$\ *\ _2$	向量的 L_2 范数	
$\ \beta_g\ _{p,q}$	β 的 L_p/L_q 复合范数	$\ \beta_g\ _{p,q} = \sum_{g \in \mathcal{G}} \ \beta_g\ _q^p$
$\ \beta_g\ _q^p$	β_g 的 L_q 范数的 p 次方	$\ \beta_g\ _q^p = (\sum_{k \in g} \beta^k ^q)^{p/q}$
β^k	β 的第 k 个元素	

由于 PCA 得到的载荷向量不稀疏, 每一个主成分都依赖于所有的原始变量, 所以导致主成分缺乏语义特征. 为了解决这一问题, Zou 等^[13] 把 PCA 等价地看作一个岭回归问题^[23], 然后把岭回归替换为弹性网回归 (Elastic net)^[24] 以引入 Lasso 惩罚, 从而得到稀疏的载荷向量. 以下为求取前 K 个主成分 SPCA 的目标函数:

$$(\hat{A}, \hat{B}) = \arg \min_{A, B} \sum_{n=1}^N \|\mathbf{x}_n - \mathbf{x}_n B A^T\|^2 + \lambda \sum_{j=1}^K \|\beta_j\|^2 + \sum_{j=1}^K \lambda_{1,j} \|\beta_j\|_1 \quad (1)$$

其中, $\mathbf{x}_n \in \mathbf{R}^{1 \times d}$ 为第 n 个观察样本, 样本总个数为 N ; $A \in \mathbf{R}^{d \times K}$ 为正交矩阵, 即 $A^T A = \mathbf{I}_{K \times K}$; $B = [\beta_1, \dots, \beta_K] \in \mathbf{R}^{d \times K}$ 为回归系数矩阵. 式中的 $\sum_{n=1}^N \|\mathbf{x}_n - \mathbf{x}_n B A^T\|^2$ 的作用是为了保证样本 \mathbf{x}_n 经投影后的重构误差尽可能小, $\lambda \sum_{j=1}^K \|\beta_j\|^2 + \sum_{j=1}^K \lambda_{1,j} \|\beta_j\|_1$ 则是弹性网正则化项, 作用是保证了载荷向量的稀疏性, λ 是 K 个载荷向量共享的正则化因子, 而 $\lambda_{1,j}$ 则是为 β_j 单独设置的正则化因子, 对 β_j 进行 L_1 范数惩罚能够带

来稀疏的载荷矩阵.

人体运动数据是一种时序信号, 由很多帧组成, 所以在对运动的某一关节进行调整时, 需要对该关节位于所有帧中的状态同时进行调整, 也就是说运动参数需要同时控制一个运动序列中的一组变量, 而这一需求与 Group lasso 的思想完美契合. Group lasso 是 Yuan 等^[14] 在 Lasso 回归^[24] 的基础上提出的一种允许变量分组的回归模型 (本文将运动序列对应的原始变量按时序进行分组, 即不同帧的同一关节的旋转变量分在一组), 它将 Lasso 中的 L_1 范数正则项替换为 L_1/L_2 复合范数正则项, 具体形式如下:

$$\hat{\beta} = \arg \min_{\beta} \frac{1}{2} \|\mathbf{y} - X\beta\|^2 + \lambda \|\beta_g\|_{1,2} \quad (2)$$

其中, $\mathbf{y} \in \mathbf{R}^{N \times 1}$ 是响应向量, \mathcal{G} 是变量的分组集合, $\beta_g \in \mathbf{R}^{p_g \times 1}$ 是分组 g 中的变量所对应的系数构成的向量, 变量个数为 p_g . L_1/L_2 复合范数惩罚和 L_1 范数惩罚一样可以产生稀疏性, 并且是“成组稀疏性”, 即在 β 的多个分组中, 只有少数组被激活. 如果一个分组被激活, 则组中各分量基本都不为 0, 而未被激活的分组中各分量都为 0, 这就达到了成组选择变量的目的.

为了解决文献 [22] 中同一个关节被过多的运动参数所控制的问题, 本文引入了 Exclusive group lasso, 其思想来源于 Exclusive lasso^[25]. 在多任务回归的框架之下进行特征选择时, 研究者的关注点大多数在于变量之间的关系, 而忽视了任务间的关系, Exclusive lasso 关注的则是任务间的负相关性. 具体的说, 一个变量如果对一个任务或者一些任务敏感, 即权值很大, 那么它就不太可能对其他任务也敏感, 所以对应于其他任务的权值就应该很小甚至是为 0. 基于这样一个假设, Zhou 等^[25] 提出了 Exclusive lasso, 而 Exclusive group lasso 与之相比的区别则在于用一组变量代替 Exclusive lasso 模型中单个变量, 用一组变量组成向量的 L_2 范数代替单个变量的 L_1 范数, Exclusive group lasso 的目标函

数如下:

$$\hat{B} = \arg \min_B \|Y - XB\|^2 + \frac{1}{2} \sum_{i=1}^d \|\beta_{i\mathcal{G}}\|_{1,2}^2 \quad (3)$$

其中, $Y \in \mathbf{R}^{N \times K}$ 为响应矩阵, $X \in \mathbf{R}^{N \times d}$ 为数据矩阵, $B = [\beta_1, \dots, \beta_d]^T \in \mathbf{R}^{d \times K}$ 为待求的系数矩阵, β_i^T 为第 i 个变量对应于所有任务的系数组成的行向量, \mathcal{G} 是原始变量的分组集合, $\beta_{i\mathcal{G}} \in \mathbf{R}^{p_g \times 1}$ 是分组 g 中的变量所对应的系数构成的向量, 变量个数为 p_g .

在本文提出的模板化运动合成的模型中, 载荷向量 (B 的列向量) 的每一个分量对应着人体运动一个关节的旋转变量, 在 SPCA 的框架之下, 使用 Group lasso 对载荷向量进行变量选择, 以获得载荷向量的“成组稀疏”性, 使得载荷向量只依赖于少数的人体关节, 从而使其对应的运动参数具有了语义特征, 同时使用 Exclusive group lasso 寻找每组变量 (每个关节) 在主成分 (控制参数) 间的负相关性, 让每个关节出现在尽量少的控制参数上, 降低了控制的复杂度.

2.2 运动模板制作

与文献 [4] 类似, 制作运动模板之前需要准备大量的运动数据, 并使用动态时间弯曲算法^[26] 进行预处理, 以得到一组类型一致、长度统一、节奏相似的人体运动 $\{\mathbf{s}_n(t) | n = 1, \dots, N; t = 1, \dots, T\}$ 和每个运动相对于参考运动的时间弯曲函数 \mathbf{w}_n , 其中 N 表示该类型的运动的个数, T 是经归一化后的统一的运动帧数, $\mathbf{s}_n(t)$ 表示第 n 个运动, 第 t 帧的人体姿势, 包括根节点位置和所有关节的朝向, \mathbf{w}_n 记录经配准后运动与原运动在时间上的对应关系. 为了建模方便, 将每个运动数据统一拉成一维向量 \mathbf{x}_n (维度为一个运动帧维度与 T 的乘积, 记为 d), 所有的 \mathbf{x}_n 经中心化 (去均值) 处理之后构成了数据 $X \in \mathbf{R}^{N \times d}$.

本方法的目标函数 (如式 (4)) 以载荷矩阵 B 为优化变量, 式 (4) 中的后两个正则项的作用是分别对 B 中的不同元素针对不同的分组方式 (分别记为 \mathcal{G}^1 和 \mathcal{G}^2) 进行 L_1/L_2 范数惩罚. 为了更好地描述这两种分组方式, 用 J_{ijk} 来表示载荷矩阵 B 中的元素, i 表示该元素对应的关节编号, j 表示对应的帧号, k 表示对应的主成分 (参数), 图 4 以 3 个关节、2 帧数据和 5 个运动参数为例, 直观地展示了对载荷矩阵 B 的两种分组方式. \mathcal{G}^1 和 \mathcal{G}^2 中每个元素都是某一个主成分中所有帧的同一关节的旋转变量, 即 $\{J_{i1k}, J_{i2k}, J_{i3k}, \dots, J_{iT k}\}$, 其中 T 表示运动序列的帧数. 不同的是 \mathcal{G}^1 中的所有元素构成一个运动序列的所有变量 (如图 4(c)), 而 \mathcal{G}^2 中的所有元素构成

某一个关节对应于所有运动参数的所有变量 (如图 4(e)).

如前所述, 本文在 SPCA 的框架之下, 综合使用 Group lasso 和 Exclusive group lasso 对载荷矩阵 B 进行变量选择, 以获得稀疏的载荷矩阵, 以下为目标函数:

$$\begin{aligned} (\hat{A}, \hat{B}) = \arg \min_{A, B} & \sum_{n=1}^N \|\mathbf{x}_n - \mathbf{x}_n B A^T\|^2 + \\ & \lambda \sum_{j=1}^K \|\beta_{1,j}\|^2 + \sum_{j=1}^K \lambda_{1,j} \|\beta_{1,j\mathcal{G}^1}\|_{1,2} + \\ & \sum_{i=1}^S \lambda_{2,i} \|\beta_{2,i\mathcal{G}^2}\|_{1,2} \end{aligned} \quad (4)$$

其中, $A = [\alpha_1, \dots, \alpha_K] \in \mathbf{R}^{d \times K}$, 且满足 $A^T A = I_{K \times K}$, $B = [\beta_{1,1}, \dots, \beta_{1,K}] \in \mathbf{R}^{d \times K}$. $\beta_{1,j}$ 是 B 的列向量, 也是待求的载荷向量, 对其进行 L_1/L_2 惩罚以产生“成组稀疏”的效果, 使其只依赖于很少的关节, 从而获取该运动的一个自由度, 它所对应的低维参数则作为该运动模板的一个接口; 如图 4(e), $\beta_{2,i}^T$ 是元素重排后的载荷矩阵 (记为 $B' = [\beta_{2,1}, \dots, \beta_{2,S}]^T \in \mathbf{R}^{S \times TK}$) 的行向量, 其中, S 表示关节数, $TK = T \times K$ 表示帧数与参数个数的乘积, 对其进行 Exclusive group lasso 惩罚, 本质上是让所有运动参数“竞争”获取每一个关节的“控制权”. SPCA 目标函数的求解是一个迭代的过程, 轮流固定 A 和 B 中的一个求另一个, 本文也采取类似的求解方式. 根据文献 [13] 中的证明, 当固定 B 时, 对 $X^T X B$ 进行奇异值分解得 $X^T X B = U D V^T$, 则 $A = U V^T$, 其中 X 是中心化后的数据矩阵, 详细的证明过程可以参考文献 [13] 的附录; 当固定 A 时, 式 (4) 就转化成了一个多任务学习的问题:

$$\begin{aligned} \hat{B} = \arg \min_B & \sum_{j=1}^K \|X \alpha_j - X \beta_{1,j}\|^2 + \\ & \lambda \sum_{j=1}^K \|\beta_{1,j}\|^2 + \sum_{j=1}^K \lambda_{1,j} \|\beta_{1,j\mathcal{G}^1}\|_{1,2} + \\ & \sum_{i=1}^S \lambda_{2,i} \|\beta_{2,i\mathcal{G}^2}\|_{1,2} \end{aligned} \quad (5)$$

式 (5) 是一个非光滑的凸函数, 所以需要为其寻找一个光滑的近似, 以便使用基于梯度的算法进行优化.

本文借鉴于 Chen 等^[15] 处理非光滑凸函数的方法求解式 (5), 向量的 L_1/L_2 范数导致目标函数不光滑, 所以只需要找到 L_1/L_2 范数的光滑近似,

就等于找到了目标函数的光滑近似. Nesterov^[27]提出了一些用光滑函数近似非光滑函数的方法, 根据这些方法可知, 对于任意的向量 $\beta \in \mathbf{R}^d$, 都有 $\|\beta\| = \max_{\|v\| \leq 1} \langle \beta, v \rangle$. 基于这样一个简单的性质, Chen 等^[15]用 $f_{\mathcal{G}, \delta}(\beta) = \max_{\|v_{\mathcal{G}}\|_{\infty, 2} \leq 1} \langle \beta, v \rangle - \delta \|v\|^2 / 2 = \|\beta_{\mathcal{G}}\|_{1, 2} - \delta \|v\|^2 / 2$ 作为 $\|\beta_{\mathcal{G}}\|_{1, 2}$ 的光滑近似 (δ 是一个控制近似精度的参数), 并给出了 $f_{\mathcal{G}, \delta}(\beta)$ 梯度的计算方法: $\nabla f_{\mathcal{G}, \delta}(\beta) = V_{\mathcal{G}, \delta}(\beta) = (v_{g_1, \delta}(\beta)^T, v_{g_2, \delta}(\beta)^T, \dots, v_{g_L, \delta}(\beta)^T)^T$ ($g_l \in \mathcal{G}$), 其中对于任意的 $g \in \mathcal{G}$, 有 $v_{g, \delta}(\beta) = \beta_g / \max\{\delta, \|\beta_g\|\}$.

本文中对矩阵 B 的分组方式以及式 (5) 两个正则项与文献 [15] 有所不同, 但可以利用其对 L_1/L_2 范数光滑近似的方法, 得到式 (5) 后两个正则项的光滑近似:

$$\Omega_{\delta}(B) = \sum_{j=1}^K \lambda_{1,j} f_{\mathcal{G}^1, \delta}(\beta_{1,j}) + \sum_{i=1}^S \lambda_{2,i} f_{\mathcal{G}^2, \delta}(\beta_{2,i}) \quad (6)$$

可求得 $\Omega_{\delta}(B)$ 的梯度为

$$\nabla \Omega_{\delta}(B) = \sum_{j=1}^K \lambda_{1,j} \nabla f_{\mathcal{G}^1, \delta}(\beta_{1,j}) + 2 \sum_{i=1}^S \lambda_{2,i} f_{\mathcal{G}^2, \delta}(\beta_{2,i}) \nabla f_{\mathcal{G}^2, \delta}(\beta_{2,i}) \quad (7)$$

至此, 式 (5) 的梯度就可以近似地表示为

$$2X^T(XB - XA) + 2\lambda B + \nabla \Omega_{\delta}(B) \quad (8)$$

根据以上梯度函数, 选用 APG (Accelerated proximal gradient method)^[28] 进行求解, 得到了载荷矩阵 B , 再将其加入求解 SPCA 下一轮的迭代中, 直到 A 和 B 都趋于稳定, B 即是决定低维运动参数语义特征的稀疏载荷矩阵, A 即为将低维运动参数重构回原始空间的正交矩阵. 令 $\mathbf{x}_{ave} = \frac{\sum_{n=1}^N \mathbf{x}_n}{N}$ 为所有样本的平均, $\gamma \in \mathbf{R}^{1 \times K}$ 为某运动样本 \mathbf{x} 经载荷矩阵 B 投影到低维参数空间所得的低维运动参数, 则 \mathbf{x} 可近似地重新表示为 $\mathbf{x} = \mathbf{x}_{ave} + \gamma A^T$. 分析目标函数 (4) 可知, $\sum_{j=1}^K \lambda_{1,j} \|\beta_{1,j}\|_{1, 2}$ 保证了 B 的稀疏性, 每一个载荷向量只依赖于很少的关节,

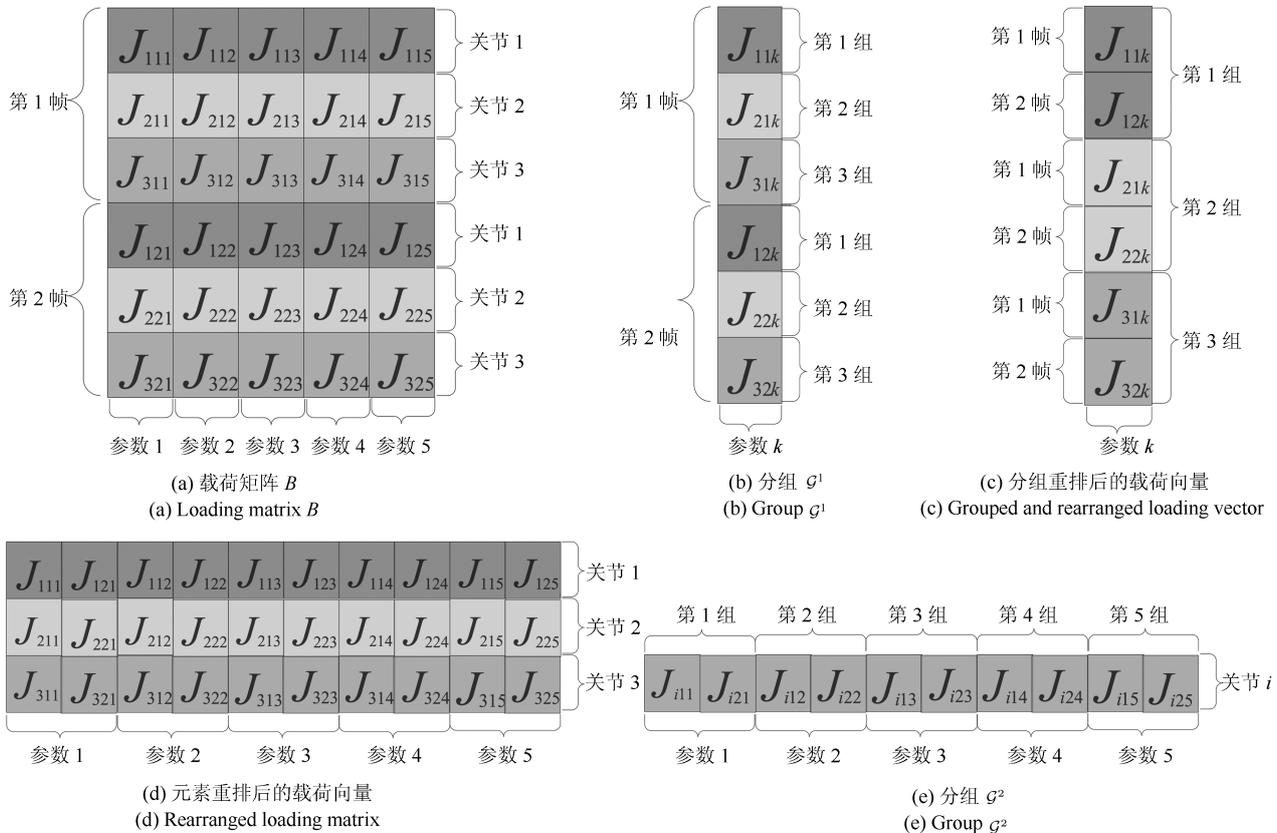


图 4 载荷矩阵 B 的两种分组方式

Fig. 4 Two grouping ways of loading matrix B

其意义易于理解,从而使得原始运动经 B 降维后获得的低维参数具有明显的语义,所以 B 决定低维参数(运动模板对外接口)的语义特征,控制着运动的内在自由度.在进行运动合成时,只需要对具有语义的低维运动参数进行调整,再由 A^T 重构回原始空间,合成的运动只会在局部关节上有较为明显的改变,从而达到了对运动进行局部控制的目的. $\sum_{i=1}^S \lambda_{2,i} \|\beta_{2,i} g_2\|_{1,2}^2$ 使得低维运动参数“竞争”获取每一个运动参数的“控制权”,降低了一个关节被过多的运动参数控制的可能性,尽可能地做到运动参数与运动自由度一一对应.

以上内容主要是针对人体运动包含的空间信息进行建模,但时序信息也是人体运动重要的组成部分,它包含了运动的快慢、节奏等重要特征,所以需要通过建模获取控制运动时序特征的低维参数(以下称作时序参数,以区别于上文中控制人体姿态的运动参数)并对其进行有效的调整.本文的策略是对运动的时序信息进行 PCA 以获取这些时序参数,在对其进行调整之后再利用 PCA 得到的基向量将其重构回原始空间.数据预处理时得到的时间弯曲函数 $w_n \in \mathbf{R}^{1 \times T}$ 是一个有约束的函数(非负且严格单调),对其直接进行 PCA 会产生无效的时间序列,因此需要将 w_n 转换成无约束的函数 z_n ,然后对 z_n 进行 PCA, w_n 和 z_n 之间的转换公式为

$$z_n(t) = \ln(w_n(t) - w_n(t-1)) \quad (9)$$

$$w_n(t) = \sum_{i=1}^t \exp(z_n(i)) \quad (10)$$

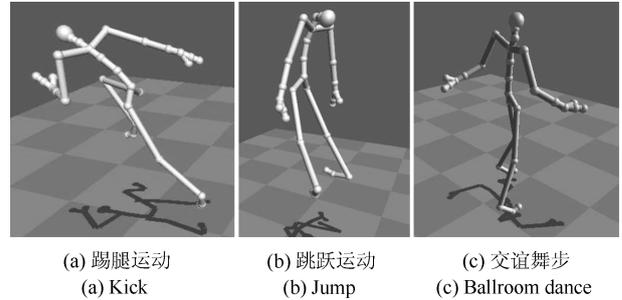
其中, $t = 1, 2, 3, \dots, T, w_n(0) = 0$. 经 PCA 后得到一组基向量 $Q \in \mathbf{R}^{T \times H}$ 和降维后的主成分 $(\eta_1^T, \eta_2^T, \eta_3^T, \dots, \eta_N^T)^T, \eta_i \in \mathbf{R}^{1 \times H}$ 表示第 i 个运动的时序参数, H 表示时序参数的个数, N 表示运动样本的个数. 令 $z_{ave} = \frac{\sum_{n=1}^N z_n}{N}$ 为所有时序特征的平均, η 为某时序特征 z 经 Q 投影后的时序参数,则 z 可以近似地重新表示为 $z = z_{ave} + \eta Q^T$. 在对 η 进行适当的调整之后,利用 z_{ave} 和 Q 就可以将其重构回原始空间得到调整后的时序特征.

因此,用来进行合成的运动模板由空间信息和时序信息两个部分组成,空间信息包括稀疏载荷向量组 B 、平均运动 x_{ave} 、重构矩阵 A 、运动参数向量 γ (初始置为 0); 时序信息包括基向量组 Q 、平均时序特征 z_{ave} 、时序参数向量 η (初始置为 0), 记为 $(B, x_{ave}, A, \gamma | Q, z_{ave}, \eta)$, 其中 B 用于确定运动参数语义含义,其他用于实时的运动合成. 运动模板的制作过程虽然相对复杂,但运动模板的制作和使用是可以分离的,这也是“模板化”概念的一个重要内容,即将复杂的运动合成工作以运动模板的形式呈

现给用户,用户面对的只是简单便捷的操作方式,而并不需要了解复杂的制作过程和相关的专业知识.

2.3 确定运动参数调整范围

运动参数大小的不合理会导致重构后产生“畸形”的人体运动,图 5 即为第 3 节实验中的三种运动在运动参数值不合理的情况下合成的“畸形”姿态. 为了避免这种情况的发生,需要确定运动参数有效的调整范围,但不能简单地以原始运动中该参数的最大值和最小值作为其调整范围,因为 SPCA 与 PCA 相比,由于对载荷向量稀疏性的“严格”要求破坏了主成分间的不相关性. 因此,虽然大多数的运动参数只控制着一个运动自由度,但一个参数的调整会影响到其他参数的可调范围. 具体到文中模型,将参数 γ 分成 (γ_a, γ_b) 两部分, γ_a 为正在被修改的参数, γ_b 为已调整完毕的参数, γ_b 的改变会影响到 γ_a 的调整范围,但在实际情况下无法获取大量的不同 γ_b 条件下的 γ_a 的统计数据,即此时 γ_a 是难以观测的随机变量,而后验分布可以作为难以观测变量的概率分布的有效估计,所以本文的策略是根据待调整参数的后验分布来确定参数 γ_a 的调整范围,具体实施方法如下.



(a) 踢腿运动 (a) Kick (b) 跳跃运动 (b) Jump (c) 交谊舞步 (c) Ballroom dance

图 5 三种运动的“畸形姿态”

Fig. 5 Deformities of the three motions

利用多元高斯模型,对原始样本投影后的低维运动参数 γ 进行统计建模,于是运动参数 γ 有以下概率分布:

$$p(\gamma) = \mathcal{N}(\gamma | \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{K}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\gamma - \mu)^T \Sigma^{-1}(\gamma - \mu)\right) \quad (11)$$

其中, $\mu \in \mathbf{R}^{1 \times K}$ 是 γ 的均值, $\Sigma \in \mathbf{R}^{K \times K}$ 是协方差矩阵, $|\Sigma|$ 是 Σ 的行列式. 在本文模型中待调整参数 γ_a 的后验分布,即为以 γ_b 为条件的条件高斯分布 $p(\gamma_a | \gamma_b)$, 且多元高斯分布的条件概率分布仍然是高斯分布,所以单参数的条件概率分布即为一元高斯分布,其均值即为 γ_a 的最大后验概率取值. $p(\gamma_a | \gamma_b)$ 可以根据 $p(\gamma)$ 进行求解: 将

μ 分成 (μ_a, μ_b) 两部分, μ_a 为正在被修改的参数的均值, μ_b 为其余的参数的均值向量, 将 Σ 分成 $\begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$ 4 部分, 则 γ_a 的条件概率分布表示如下:

$$p(\gamma_a | \gamma_b) = \mathcal{N}(\gamma_a | \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2\sigma^2}(\gamma_a - \mu)^2\right) \quad (12)$$

其中, $\mu = \mu_a + (\gamma_b - \mu_b) \Sigma_{bb}^{-1} \Sigma_{ba}$, $\sigma = (\Sigma_{aa} - \Sigma_{ab} \Sigma_{bb}^{-1} \Sigma_{ba})^{1/2}$. 因此, 根据高斯分布的特点取 $(\mu \pm 3\sigma)$ 作为该参数的调整范围.

2.4 时序参数的自适应调整

在本文方法中, 时序参数与运动参数有所不同, 前者对应的载荷向量不稀疏, 所以它没有明确的语义含义, 因此对其直接修改缺乏明确的语义指导. 而文献 [4] 中则是根据用户提供的约束对两种类型的低维参数同时进行优化, 这就对用户提供的约束的合理性提出了更高的要求, 无形中增加了运动合成控制的复杂度. 因此, 为了简化运动合成的控制方式, 模板化的运动合成方法不再需要用户提供约束, 而是使用与第 2.3 节类似的策略, 根据运动参数的值自适应地调整时序参数. 具体地, 将运动参数与时序参数组合在一起, 构成一个新的参数 $\alpha = (\gamma, \eta)$, 并对其进行类似式 (11) 的高斯建模. 同样, 将其均值分成 (μ_γ, μ_η) 两部分, 将其协方差矩阵分成 4 部分 $\begin{pmatrix} \Sigma_{\gamma\gamma} & \Sigma_{\gamma\eta} \\ \Sigma_{\eta\gamma} & \Sigma_{\eta\eta} \end{pmatrix}$, 在运动参数 γ 调整完毕之后, 以 γ 为条件计算 η 的条件概率分布得:

$$p(\eta | \gamma) = \mathcal{N}(\eta | \mu_{\eta|\gamma}, \Sigma_{\eta|\gamma}) = \frac{1}{(2\pi)^{\frac{H}{2}}} \frac{1}{|\Sigma_{\eta|\gamma}|^{\frac{1}{2}}} \times \exp\left(-\frac{1}{2}(\eta - \mu_{\eta|\gamma})^T \Sigma_{\eta|\gamma}^{-1} (\eta - \mu_{\eta|\gamma})\right) \quad (13)$$

其中, $\mu_{\eta|\gamma} = \mu_\eta + (\gamma - \mu_\gamma) \Sigma_{\gamma\gamma}^{-1} \Sigma_{\gamma\eta}$, $\Sigma_{\eta|\gamma} = (\Sigma_{\eta\eta} - \Sigma_{\eta\gamma} \Sigma_{\gamma\gamma}^{-1} \Sigma_{\gamma\eta})^{1/2}$. 均值 $\mu_{\eta|\gamma}$ 是使得时序参数条件概率最大的取值, 所以本文以均值 $\mu_{\eta|\gamma}$ 作为时序参数自适应调整的结果, 既能减少不必要的参数控制, 又能保证合成运动在时序上的合理性.

本文提出的运动合成模型, 制作模板的过程虽然相对比较耗时, 但可以在线下完成, 对在线合成速度不会造成任何影响. 整个重构过程, 只涉及到向量和矩阵的相乘, 剩下的只有三维动画的显示工作, 所以能够做到实时的运动合成. 更重要的是, 用户可以

在设定的范围内任意地调整运动参数, 并且实时地显示参数调整的合成结果. 这样, 就可以根据实时合成的结果判断当前的参数调整是否符合要求, 能够有效地提高合成的精度, 降低控制的复杂度.

3 实验结果

三组实验数据 (踢腿运动、跳跃运动和交谊舞步) 来源于 HDM05 和卡耐基梅隆大学的运动捕获数据库, 数据格式为 ASF/AMC. 实验中使用四元数来表示关节的旋转量, 同时去掉了一些对运动影响不大的关节 (锁骨和手指), 最终保留了 22 个关节. 对踢腿运动的 67 个样本, 跳跃运动的 49 个样本和交谊舞步的 78 个样本进行了建模, 开发机器的硬件参数为 2.93 GHz 的 Pentium Dual-Core 处理器, 内存为 2 GB, 软件环境为 Matlab 2012 (模板制作) 和 Visual Studio 2008 (运动合成), 表 2 为以上三个运动在实验中的相关数据.

从表 2 中可以看出, 稀疏载荷的计算相对比较耗时, 但可以在线下完成, 并不影响运动合成的相关应用, 而运动合成所需要的时间却很短, 与运动的长度成正比, 每一帧的平均合成时间大约在 0.2 毫秒左右, 这样的合成速度在面对实际应用时 (例如游戏或者动画制作) 可以给用户带来良好的用户体验. 并且对存储空间的需求也很低, 每一个运动模板需要的存储空间是一个固定值, 并不会随着建模样本数量的增减而改变, 但却能有效地保存所有样本的大部分信息.

模板化运动合成的特点在于便捷、直观的操作方式, 图 6 是运动合成软件的交互界面, 界面右侧是一些对运动姿态影响较大的运动参数, 用户可以通过鼠标直接操纵滑块对参数 (初始值为 0) 进行直观的调整, 左边的 3D 视图会实时地显示参数调整后的合成结果, 并且在参数的调整过程中, 系统会根据当前参数的值实时地更新参数的可调范围, 以确保合成出逼真的人体运动.



图 6 模板化运动合成软件界面

Fig. 6 Software interface of templated motion synthesis

表 2 3 个运动的相关实验数据
Table 2 Experimental data of the three motions

运动类型	样本数	参考运动帧数	原始样本所占空间 (MB)	运动模板所占空间 (kB)	稀疏载荷计算时间 (h)	合成时间 (ms)
踢腿运动	67	126	4.91	820	2.8	29
跳跃运动	49	73	1.15	468	1.1	16
交谊舞步	78	196	8.89	1 240	4.5	41

实验结果不仅展示了低维运动参数明显的语义特征,而且同一关节被多个运动参数控制的情况与文献 [22] 相比有了明显的改善. 图 2 是实验中对踢腿、跳跃运动和交谊舞步建模所得的稀疏载荷矩阵 B , 一个方格代表 \mathcal{G}^1 和 \mathcal{G}^2 中的一个分组. 纵向对应骨架的每一个关节 (包括根节点的位置和朝向), 从上到下对应的关节依次是根节点、下半背部、上半背部、胸部、下半颈部、上半颈部、头部、右大臂、右肘、右腕、右手、左大臂、左肘、左腕、左手、右大腿、右膝、右脚、右脚趾、左大腿、左膝、左脚、左脚趾, 横向对应主成分 (运动参数), 从左到右依次是 γ^1 到 γ^{10} . 方格的颜色深度表示该分组中所有分量绝对值的大小, 颜色越深表示越大, 白色表示该组中所有分量都为 0, 从图 2 中可以发现, 每个运动参数只依赖于少数几个关节, 运动参数的意义从而可以被理解, 例如踢腿运动的第 8 组载荷向量只有对应左大臂和右大臂的两个关节的分量不为 0, 这就说明运动参数 γ_{Kick}^8 只控制着两个关节, 经重构后发现, γ_{Kick}^8 控制的是两臂的张开幅度. 因此, γ_{Kick}^8 的语义含义即可以理解为控制踢腿运动两臂张开幅度的运动参数. 而有些参数控制的关节似乎语义特征并不够明显, 例如 γ_{Kick}^9 控制的是上半颈部、右大臂和右大腿, 但仔细观察发现, 右大腿对应的分组方格的颜色很深, 说明该组分量数值的绝对值很大, 而另两个分组数值的绝对值则相对较小, 表明对于 γ_{Kick}^9 , 右大腿相对于另两个关节更重要. 经合成后发现, γ_{Kick}^9 控制的其实是踢腿的高度, 这也从数据上说明了人体运动关节之间的相关性: 控制踢腿运动时, 如果只单纯地提高踢腿的高度, 而不对其他关节做任何改变, 势必会造成运动的不自然, 正常情况下, 人体提高抬腿高度时, 头部势必轻微前倾, 手臂势必轻微上抬, 而 γ_{Kick}^9 对应的关节情况正好验证了这一点, 既保证了对运动的局部改变, 又避免了只对个别关节进行调整造成运动失真.

3.1 运动合成结果

图 7~18 给出了其中一些在均值运动 (运动参数都为 0) 的基础上调整单个运动参数的合成结果, 表明在一个统一的运动模板下仍然能够合成大量的具有表现力的新运动. 图 7 给出了踢腿运动中帧

距的 4 帧运动姿态, γ_{Kick}^1 控制着左臂弯曲的程度以及身体的向左的倾斜度, 而身体向左的倾斜度并不是一个独立的运动自由度, 它与手臂弯曲程度是相关的. 图 8 中则是控制右臂弯曲程度的合成效果图, 与图 7 有类似的效果.

γ_{Kick}^9 控制着踢腿的高度, 图 9 中线条表示脚尖的的运动轨迹, 能够看出图 9 (a) 中的踢腿高度比图 9 (b) 中的踢腿高度明显高出许多, 且人体姿态并没有随着参数的调整而显得不自然, 图 10 则很好地验证了上文提到的 γ_{Kick}^8 控制着两臂张开幅度.

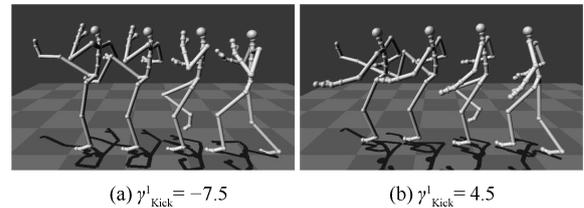


图 7 γ_{Kick}^1 控制左手弯曲程度效果图

Fig. 7 Results of γ_{Kick}^1 controlling bending degree of left arm

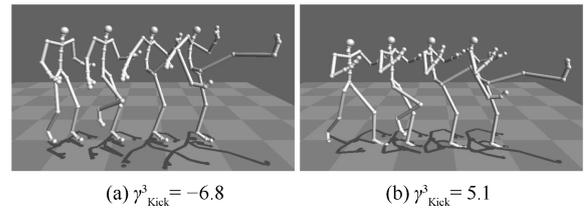


图 8 γ_{Kick}^3 控制右手弯曲程度效果图

Fig. 8 Results of γ_{Kick}^3 controlling bending degree of right arm

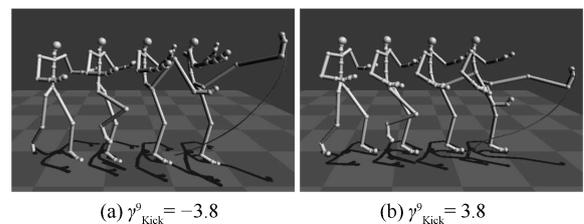


图 9 γ_{Kick}^9 控制踢腿高度效果图

Fig. 9 Results of γ_{Kick}^9 controlling kick height

除了人体关节, 运动参数还能控制运动的轨迹和方向, 图 2 (a) 中的第 6 个和第 10 个载荷向量中,

根节点对应的颜色很深, 说明 γ_{Kick}^6 和 γ_{Kick}^{10} 可以对根节点进行操作, 经实验发现, γ_{Kick}^6 控制根节点的轨迹, γ_{Kick}^{10} 控制踢腿的方向, 以 γ_{Kick}^{10} 为例, 图 11 为 γ_{Kick}^{10} 控制根节点的效果图, 线条为指尖的轨迹。

在跳跃运动上也有类似的实验效果. 图 12 为 γ_{Jump}^2 取不同值时, 手臂呈现不同的弯曲程度的结果; 如图 13, γ_{Jump}^6 控制着根节点, 所以 γ_{Jump}^6 取不同值时, 合成的运动呈现不同的跳跃距离; 如图 14, γ_{Jump}^4 控制着两臂张开的程度; 而 γ_{Jump}^7 则控制了前后摆臂的幅度, 如图 15 所示。

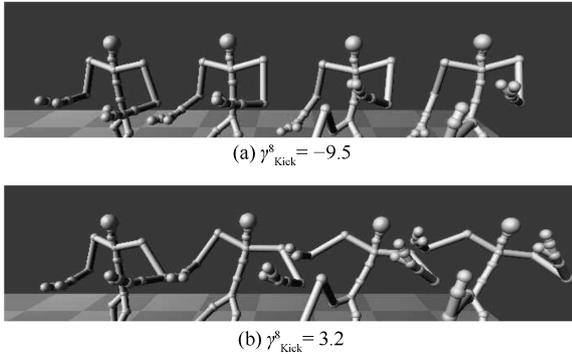


图 10 γ_{Kick}^8 控制两臂张开程度效果图
Fig. 10 Results of γ_{Kick}^8 controlling open degree of two arms

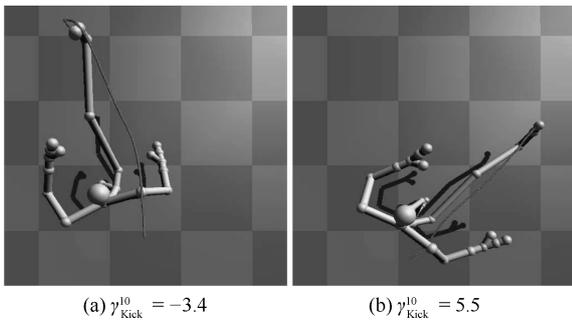


图 11 γ_{Kick}^{10} 控制踢腿方向效果图
Fig. 11 Results of γ_{Kick}^{10} controlling kick direction

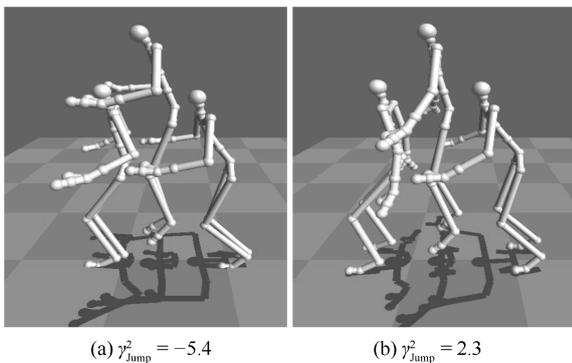


图 12 γ_{Jump}^2 控制手臂弯曲程度效果图
Fig. 12 Results of γ_{Jump}^2 controlling bending degree of arms

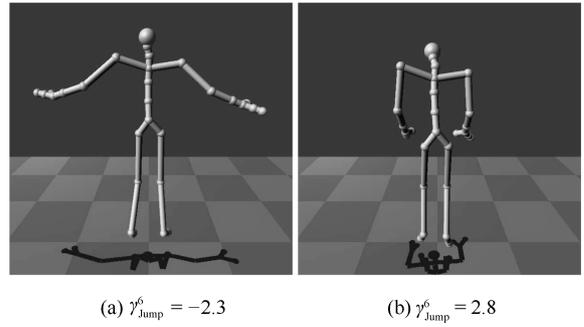


图 13 γ_{Jump}^6 控制跳跃距离效果图
Fig. 13 Results of γ_{Jump}^6 controlling jump distance

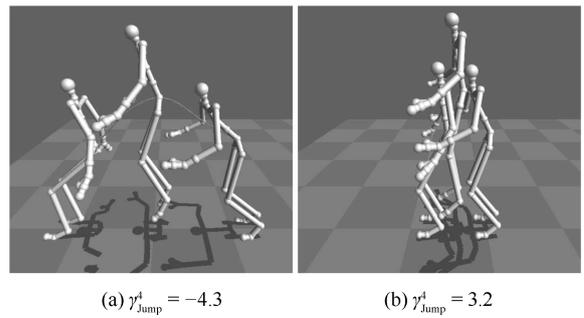


图 14 γ_{Jump}^4 控制两臂张开程度效果图
Fig. 14 Results of γ_{Jump}^4 controlling open degree of two arms

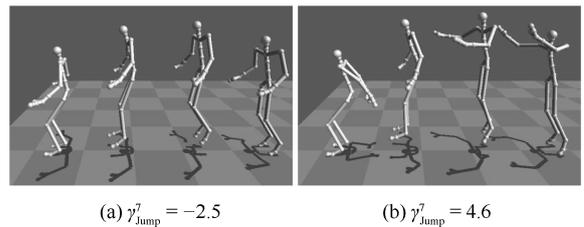
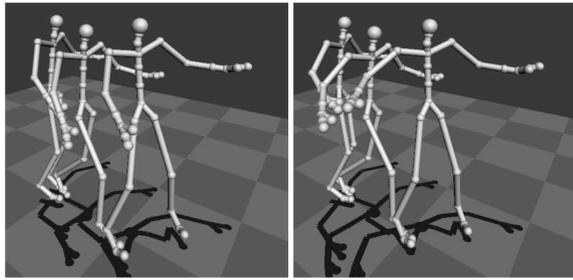


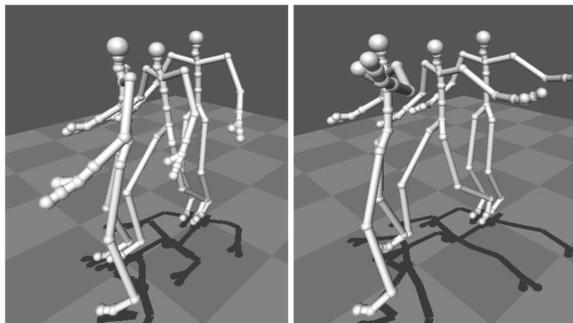
图 15 γ_{Jump}^7 控制前后摆臂幅度效果图
Fig. 15 Results of γ_{Jump}^7 controlling amplitudes of arms

第 3 组实验选择的是一个经典的交谊舞步, 因为该运动的建模样本中下半身动作差异较小, 主要区别在手部的动作和运动的路线上, 在图 2(c) 中有较为明显的体现, 出现在运动参数上的都是双臂上的关节和根节点, 所以针对交谊舞步模板的运动参数主要控制着手臂的姿态和舞步的轨迹. 图 16 和 17 中显示的是该运动的三个关键姿态, 即初始态、跨右脚和跨左脚. 从图 16 中可以看出, 运动参数 γ_{Dance}^6 控制着跳舞时右臂的抬起高度; 从与之对应的图 17 中可以看出, γ_{Dance}^2 控制着跳舞时左臂的抬起高度. 除了手臂动作之外, 跳舞时的运动路线以及身体转向也是可以控制的, 如图 18 所示, 地板上线条为根节点的运动轨迹, 箭头表示人体的正面朝向, 所以 γ_{Dance}^5 控制着该运动的路线和舞步转向。



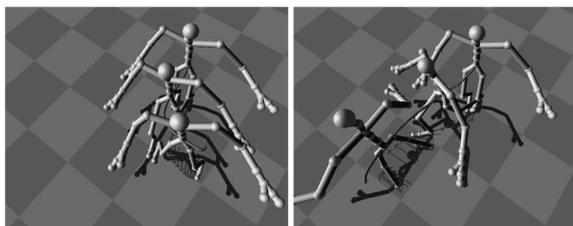
(a) $\gamma_{Dance}^6 = 2.5$ (b) $\gamma_{Dance}^6 = -4$
图 16 γ_{Dance}^6 控制右臂抬起高度效果图

Fig. 16 Results of γ_{Dance}^6 controlling height of right arm



(a) $\gamma_{Dance}^2 = 6$ (b) $\gamma_{Dance}^2 = -6$
图 17 γ_{Dance}^2 控制左臂抬起高度效果图

Fig. 17 Results of γ_{Dance}^2 controlling height of left arm



(a) $\gamma_{Dance}^5 = 6$ (b) $\gamma_{Dance}^5 = -6$
图 18 γ_{Dance}^5 控制运动轨迹和转体方向效果图

Fig. 18 Results of γ_{Dance}^5 controlling the trajectory and direction

3.2 运动合成的精确控制

本文提出的模板化的运动合成模型, 不仅能够赋予控制参数语义含义, 而且能够为用户提供精确的运动合成控制. 当一个运动被制作成模板之后, 其相关的运动特性 (如上文所述的踢腿高度、双臂张开幅度、跳跃距离、身体转向等) 就由某一个运动参数控制, 如果将这些特性量化成一个具体的变量, 那么它就会与对应的运动参数之间存在一个函数关系. 以踢腿高度为例, 用 h 表示踢腿高度, 那么就存在一个函数 f 满足 $h = f(\gamma_{Kick}^9)$, 理论上用户就可以根据这样的函数关系精确地控制踢腿高度, 但实际上并不需要明确知道函数 f , 因为本文提出的模型能够做到实时的运动合成, 而 f 在定义域内 (运动参数的可调范围) 又都是单调的, 所以用户完全可以根

据画面中实时的合成效果来确定当前参数的合理性, 这就达到了对合成过程的精确控制, 有效地提高了合成效果. 图 19 是精确控制合成过程和结果的示意图. 图 19 (a) 中, 由远及近的 4 个姿态对应的 γ_{Kick}^9 分别为 -3.7 、 -1.3 、 1.2 、 3.7 , 对应的踢腿高度呈现近似线性下降的趋势; 图 19 (b) 中, 从左到右的四对姿态对应的 γ_{Jump}^6 分别为 2.8 、 1.1 、 -0.6 、 -2.3 , 对应的跳跃距离呈现近似线性上升的趋势. 如图 20 所示, 图 20 (a) 和 (b) 分别是 γ_{Kick}^9 与踢腿高度和 γ_{Jump}^6 与跳跃距离之间的关系图, 这种精确且简单的对应关系就保证了运动参数对合成过程及结果的精确控制.

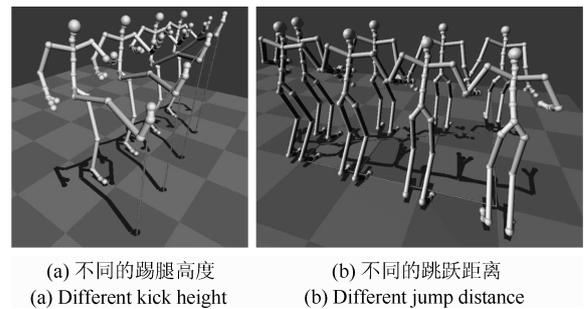


图 19 运动合成的精确控制示意图

Fig. 19 Accurate control of motion synthesis

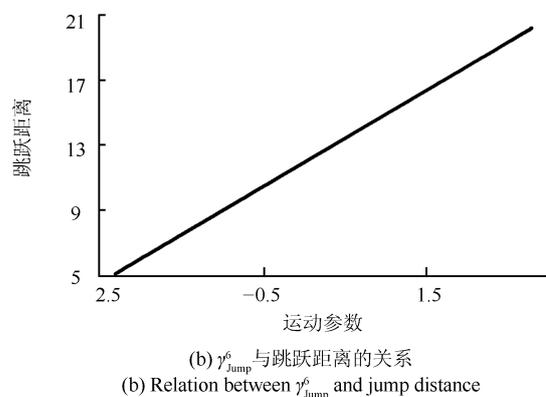
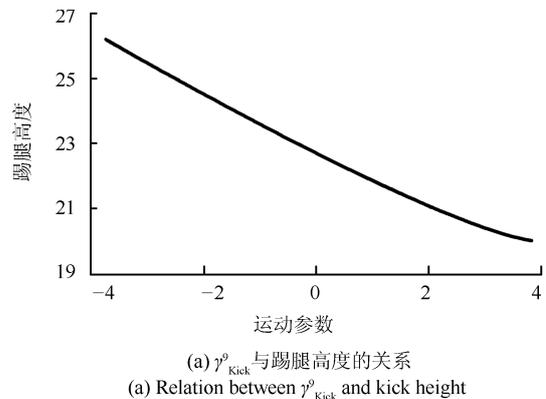


图 20 运动参数与运动特性之间的关系图

Fig. 20 Relation between motion parameters and motion attributes

3.3 Exclusive group lasso 特性体现

对于一个关节, 如果被过多的运动参数所控制, 就会导致运动参数对彼此所控制的自由度造成交叉影响, 即一个运动特性需要多个运动参数控制, 这就增加了控制的难度, 降低了运动参数以及合成过程的可理解性, 所以希望同一个关节被尽量少的运动参数控制. 本文在模型中引入了 Exclusive group lasso, 在保证运动参数语义特征的同时, 还降低了同一个关节被不同运动参数所控制的可能性, 从而尽可能地做到运动参数与运动自由度一一对应, 在很大程度上能够改善文献 [22] 所存在的上述问题. 图 21 是文献 [22] 中方法求得的三个运动的稀疏载荷向量组的示意图, 与图 2 对应, 从图 2 中可以看出, 一些关节 (如踢腿运动的根节点、右大腿、右大臂) 相较于图 21 出现在更少的控制参数上, 说明本文的方法有效地减少了控制同一关节的运动参数的个数, 降低了运动合成控制的复杂度.

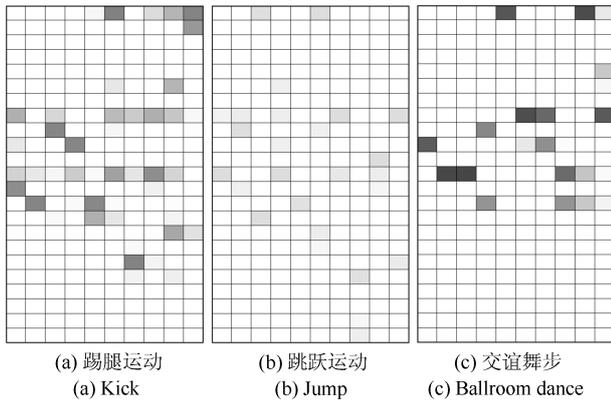


图 21 SSPM 模型得到三种运动的稀疏载荷向量组示意图
Fig. 21 Diagram of sparse loading matrices of the three motions using SSPM

在运动合成的过程中, 以上改进主要体现在两个方面: 1) 减少运动参数对彼此所控制的自由度的交叉影响, 如图 22(a) 和 (b) 所示, 这是文献 [22] 中方法得到的参数 γ_{Kick}^0 分别等于 3.8 和 -3.8 的合成

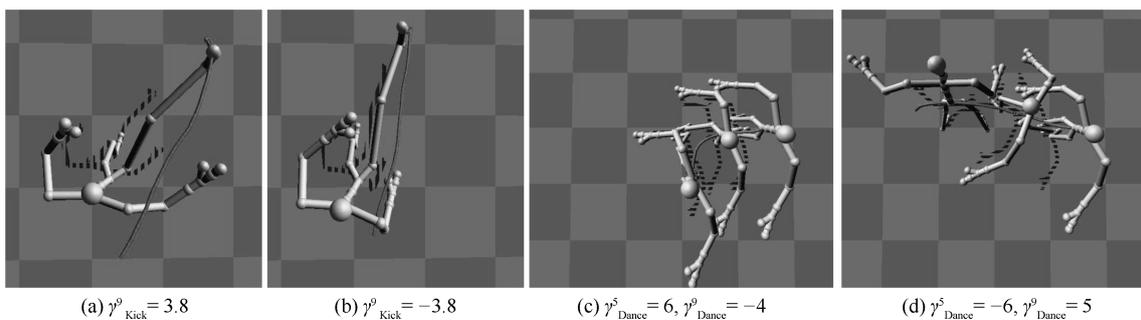


图 22 文献 [22] 中方法的合成结果

Fig. 22 Synthesis results of the method in [22]

结果, γ_{Kick}^0 控制的是踢腿高度, 但由于根节点也出现在了 γ_{Kick}^0 中, 如图 21(a) 所示, 这就导致其对踢腿方向也造成了一定的影响, 如图 22(a) 和 (b) 中的线条所示, 而踢腿方向这一运动特性本该由 $\gamma_{\text{Kick}}^{10}$ 来控制. 2) 降低需要多个参数控制一个运动特性的可能性, 如图 22(c) 和 (d) 所示, 这是文献 [22] 中方法得到的参数 γ_{Dance}^5 和 γ_{Dance}^0 分别是 6、-4 和 -6、5 的合成结果, 对比图 18 发现, 与本文方法仅由 γ_{Dance}^5 控制的合成效果是类似的, 而多参数控制一个运动特性的情况势必增加了控制的复杂度, 降低了运动参数的可理解性.

4 讨论

模板化运动合成模型的一个重要任务是利用 SPCA 获取稀疏的载荷向量, 使得运动参数只依赖于少数的人体关节, 从而具有了语义含义. 可是对载荷向量稀疏性的要求破坏了主成分分析方差最大化的特性, 即对于前 K 个主成分, SPCA 保留的原始运动的信息量比 PCA 少. 但现有的很多利用低维参数进行运动合成的方法^[10-11, 19-20] 为了显示运动样本投影到低维空间后的位置 (以便于对其进行编辑), 大多选择两到三维的参数空间, 而本文模型选择 10 维的低维参数空间, 虽然其对应的载荷向量是稀疏的, 但通过对目标函数 (4) 中 λ 和 $\lambda_{1,j}$ 的控制, 很容易使得 SPCA 前 10 个主成分保留的信息多于 PCA 前 3 个主成分保留的信息, 所以本文中的参数模型并没有因稀疏性而过多牺牲重构的准确性.

模板化运动合成模型与现有的一些根据用户约束进行优化的运动合成方法^[2, 4, 12] 有所不同, 后者本质上是通过优化合成出一种与现有运动最“相似”又满足用户约束的新运动, 当用户提供约束的合理性无法保证时, 就可能造成合成结果破坏了运动约束, 如接触约束、物理约束等. 但在本文中, 用户通过在一定范围内调节运动参数进行运动合成, 并不需要用户提供约束, 第 2.3 节中参数范围的确定过程可以看成对合成结果进行运动约束的过程, 因为

参数范围是根据真实运动样本统计分布确定的,真实的运动样本都满足运动约束,所以根据其统计分布特征确定的范围内的样本也应满足大部分的运动约束.当然,由于骨架尺寸的不同会导致手脚末端的接触问题,但这一点在有场景的应用中可以通过对关节长度或是根节点坐标的适当调整来解决,例如在网购衣服时,网站可以提供动画人物试衣的服务,用户可以根据需要的尺寸调整动画人物的尺寸,并可以在特定的场景中利用本文的方法控制其做出不同的动作以观察衣服的效果.综上所述,运动模板不仅为用户提供了控制运动合成的接口,还限定了合成运动的范围,保证了合成结果的逼真性.

模板化运动合成模型的其中一个重要特点是自动地获取运动本身的自由度,对运动的类型并没有要求,但运动的自由度不宜过多,所以其主要适用于单元运动(如一个走路周期、一个舞步或是一个拳击动作)合成,不适合长运动合成,因为长运动势必会有更多的自由度,而这些自由度之间很容易造成交叉影响,增加运动合成控制的复杂度和不确定性.并且,运动模板除了为用户提供控制合成的接口之外,也框定了运动的内容,比如将一长段芭蕾舞表演制作成模板,那么用其只能合成模板所限定的内容,这就破坏了舞蹈内容的多样性,在需要合成长运动时可以利用运动拼接技术将各个运动单元进行拼接,这样既降低了合成控制的复杂度,又保证了合成的长运动内容的丰富程度.

该方法也存在一些问题,主要包括三点:1) 对前期的数据准备的要求严格,因为参与建模数据的质量将直接决定合成运动的质量;2) 不同关节对运动的重要程度难以区分,因为关节对于不同运动的重要性是不同的,不好统一设置;3) 该方法只能合成运动单元,不能单独完成长运动合成,若要合成内容丰富的长运动还需辅以运动拼接技术.

5 总结

本文提出了一种模板化运动合成的方法,该方法结合 SPCA、Group lasso 和 Exclusive group lasso 对人体运动进行建模,求得一组稀疏的载荷向量和重构矩阵,稀疏载荷向量中只有少数关节对应的分量不为 0,这使得低维的运动参数只控制着少数的关节因而具有的语义特征,此外,Exclusive group lasso 的引入使得运动参数之间需要“竞争”获取每一个关节的控制权,以免同一关节出现在过多的运动参数上,减少了运动合成控制的不确定性.该方法有效地降低了运动合成控制的复杂度,有利于专业人员将更多的精力放在艺术创作上,而普通大众也可以用该方法合成自己需要的人体动画.

References

- Zhang Li-Ge, Bi Shu-Sheng, Gao Jin-Lei. Human motion data acquiring and analyzing method for humanoid robot motion designing. *Acta Automatica Sinica*, 2010, **36**(1): 107–112
(张利格, 毕树生, 高金磊. 仿人机器人复杂动作设计中人体运动数据提取及分析方法. *自动化学报*, 2010, **36**(1): 107–112)
- Kovar L, Gleicher M, Pighin F. Motion graphs. *ACM Transactions on Graphics (TOG)*, 2002, **21**(3): 473–482
- Kovar L, Gleicher M. Flexible automatic motion blending with registration curves. In: *Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Switzerland: Eurographics Association, 2003. 214–224
- Min J Y, Chen Y L, Chai J X. Interactive generation of human animation with deformable motion models. *ACM Transactions on Graphics (TOG)*, 2009, **29**(1): Article No. 9
- Kwon T, Shin S Y. Motion modeling for on-line locomotion synthesis. In: *Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Los Angeles, CA, USA: ACM, 2005. 29–38
- Heck R, Gleicher M. Parametric motion graphs. In: *Proceedings of the 2007 Symposium on Interactive 3D Graphics and Games*. Seattle, Washington, USA: ACM, 2007. 129–136
- Kovar L, Gleicher M. Automated extraction and parameterization of motions in large data sets. *ACM Transactions on Graphics*, 2004, **23**(3): 559–568
- Park S I, Shin H J, Kim T H, Shin S Y. On-line motion blending for real-time locomotion generation. *Computer Animation and Virtual Worlds*, 2004, **15**(3–4): 125–138
- Li Jin-Dan, Mao Tian-Lu, Wang Zhao-Qi, Liu Jin-Gang. Motion graph construction based on parametric motion synthesis. *Computer Simulation*, 2009, **26**(3): 208–212
(李锦丹, 毛天露, 王兆其, 刘金刚. 基于参数化运动合成的运动图构建及其应用. *计算机仿真*, 2009, **26**(3): 208–212)
- Shin H J, Lee J. Motion synthesis and editing in low-dimensional spaces. *Computer Animation and Virtual Worlds*, 2006, **17**(3–4): 219–227
- Wang Yu-Jie, Xiao Jun, Wei Bao-Gang. 3D human motion synthesis based on nonlinear manifold learning. *Journal of Image and Graphics*, 2010, **15**(6): 936–943
(王宇杰, 肖俊, 魏宝刚. 基于非线性流形学习的 3 维人体运动合成. *中国图象图形学报*, 2010, **15**(6): 936–943)
- Liu H, He F, Cai X T, Chen X, Chen Z. Human motion synthesis using window-based local principal component analysis. In: *Proceedings of the 12th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics)*. Washington D. C., USA: IEEE, 2011. 282–287
- Zou H, Hastie T, Tibshirani R. Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 2006, **15**(2): 265–286
- Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2006, **68**(1): 49–67
- Chen X Y, Yuan X T, Yan S C, Tang J H, Rui Y, Chua T S. Towards multi-semantic image annotation with graph regularized exclusive group lasso. In: *Proceedings of the 19th ACM International Conference on Multimedia*. New York, NY, USA: ACM, 2011. 263–272

- 16 Gleicher M, Shin H J, Kovar L, Jepsen A. Snap-together motion: assembling run-time animations. In: Proceedings of the ACM SIGGRAPH 2008 Classes. New York: ACM, 2008: Article No. 52
- 17 Kwon T, Shin S Y. Motion modeling for on-line locomotion synthesis. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. New York, NY, USA: ACM, 2005. 29–38
- 18 Safonova A, Hodgins J K, Pollard N S. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. *ACM Transactions on Graphics (TOG)*, 2004, **23**(3): 514–521
- 19 Li Chun-Peng, Wang Zhao-Qi, Xia Shi-Hong. Motion synthesis for virtual human using functional data analysis. *Journal of Software*, 2009, **20**(6): 1664–1672
(李淳芑, 王兆其, 夏时洪. 人体运动的函数数据分析与合成. 软件学报, 2009, **20**(6): 1664–1672)
- 20 Liu Geng-Dai, Xu Ming-Liang, Zhang Ming-Min. Human motion synthesis based on independent spatio-temporal feature space. *Chinese Journal of Computers*, 2011, **34**(3): 464–472
(刘更代, 徐明亮, 张明敏. 基于独立时空特征空间的人体运动合成. 计算机学报, 2011, **34**(3): 464–472)
- 21 Lan Rong-Yi, Sun Huai-Jiang. Style analysis and human locomotion synthesis based on inverse kinematics and reconstructive ICA. *Acta Automatica Sinica*, 2014, **40**(6): 1135–1147
(蓝荣祎, 孙怀江. 基于逆运动学和重构式 ICA 的人体运动风格分析与合成. 自动化学报, 2014, **40**(6): 1135–1147)
- 22 Lan Rong-Yi, Sun Huai-Jiang. A sparse semantic parametric model for interactive motion synthesis. *Journal of Computer-Aided Design & Computer Graphics*, 2013, **25**(3): 341–349
(蓝荣祎, 孙怀江. 人体运动的稀疏语义参数化模型与交互式合成. 计算机辅助设计与图形学学报, 2013, **25**(3): 341–349)
- 23 Hoerl A E, Kennard R W. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 1970, **12**(1): 55–67
- 24 Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2005, **67**(2): 301–320
- 25 Zhou Y, Jin R, Hoi S C H. Exclusive lasso for multi-task feature selection. In: Proceedings of the 2010 International Conference on Artificial Intelligence and Statistics. 2010. 988–995
- 26 Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1978, **26**(1): 43–49
- 27 Nesterov Y. Smooth minimization of non-smooth functions. *Mathematical Programming*, 2005, **103**(1): 127–152
- 28 Tseng P. On accelerated proximal gradient methods for convex-concave optimization. *SIAM Journal on Optimization*, 2008.



夏贵羽 南京理工大学博士研究生. 2012 年获得南京理工大学学士学位. 主要研究方向为模式识别, 人体运动捕获数据重用.

E-mail: xiaguiyu1989@sina.com

(XIA Gui-Yu Ph.D. candidate at Nanjing University of Science and Technology. He received his bachelor degree from Nanjing University of Science and Technology in 2012. His research interest covers pattern recognition and human motion capture data reusing.)



孙怀江 南京理工大学计算机科学与工程学院教授. 1995 年获得西北工业大学博士学位. 主要研究方向为神经网络与机器学习, 人体运动分析与合成. 本文通信作者.

E-mail: sunhuaijiang@njust.edu.cn

(SUN Huai-Jiang Professor at Nanjing University of Science and Technology. He received his Ph.D. degree at Northwestern Polytechnical University in 1995. His research interest covers neural networks and machine learning, human motion analysis and synthesis. Corresponding author of this paper.)