

基于学习的群体动画生成技术研究

魏迎梅 瞿 师 吴玲达

(国防科技大学信息系统与管理学院, 长沙 410073)

摘 要 为了降低群体动画中生成大量自然而又相似的人体运动的难度和复杂性,研究了一种基于学习的群体动画生成技术。该技术首先通过建立基于高斯过程隐变量模型和隐空间动态模型的运动姿势学习模型,将高维运动姿势映射到低维隐空间中,并在低维隐空间对相邻姿势的动态演化进行建模;然后通过对已有运动数据的学习来获得组成该运动的姿势的概率分布,再通过隐空间中的动态预测和 Hybrid Monte Carlo 采样来得到符合给定概率分布的隐轨迹;最后通过姿势重构来得到与原运动非常相似但又不同的一系列自然的运动,以产生群体动画,从而避开了传统的基于几何和物理约束的逆运动方法固有的困难和复杂性。

关键词 机器学习 群体动画 高斯过程隐变量模型 高斯动态模型

中图法分类号: TP391.9 TP391.41 文献标志码: A 文章编号: 1006-8961(2010)06-971-07

Group Animations Based on Machine Learning

WEI Yingmei, QU Shi, WU Lingda

(School of Information System and Management, National University of Defense Technology, Changsha 410073)

Abstract A group animation generation method based on machine learning was proposed in order to reduce the complexity of generating mass of similar but different natural human motions in group animations. There are two models. Poses learning model was built based on Gaussian process latent variable model to characterize a specific motion and dynamic model was built in latent space to characterize the dynamic evolving process of neighboring poses in latent space. These models can be represented as probability distribution over all poses composing the motion by learning from existing motion data. Dynamic prediction can be made in latent space for giving initial state, then hundreds of latent trajectories by Hybrid Monte Carlo sampling according to given probability distribution can be obtained. Group animations can be implemented by generating a series of similar but different natural motions reconstructed from these latent trajectories, thereby avoid the difficulty and complexity of calculating geometric relationship and physical constrains in inverse kinematics.

Keywords machine learning, group animation, Gaussian process latent variable model, Gaussian process dynamic model

0 引言

群体动画一直是计算机动画中的难点问题,特别是由人形角色组成的群体的运动生成,它不仅需要生成自然合理的人体运动,而且需要生成若干个相似但又不相同的运动。

对于角色动画生成而言,比较有效的方法是逆运动(Inverse kinematic)方法。其本质是从一组约

束中计算合理的人体姿势,尽管其计算的工作是由计算机完成的,但它存在着固定的不确定性,例如,用户可指定一个角色的手脚位置,但对于复杂的人体运动链而言,有许多可能的姿势都满足这些约束,而且有些姿势看起来是不太协调不太像的姿势。但是,如果想使逆运动的解唯一,并且能确定出最像最协调的姿势,那么就需要指定更多的约束,这无论对用户还是对系统来说都是非常困难的,其根源在于用简单的生物力学难以解释复杂的人体姿势^[1],而

基金项目:国家高科技研究发展计划(863)项目(2009AA01Z335)

收稿日期:2009-06-01;改回日期:2009-08-02

第一作者简介:魏迎梅(1972—),女,副教授。2000年于国防科技大学获计算机科学与技术专业博士学位。研究方向为多媒体信息系统与虚拟现实。E-mail: weiyingmei@nudt.edu.cn

对于群体动画而言,即使已生成了一段合理的人体运动,而当需要生成若干个与之相似但又不相同的运动时,问题则变得更加复杂。如果在原始运动数据上直接编辑,就存在着局部修改与整体协调性之间的关系问题,如果从改变约束入手,则又回到了逆运动固有的问题复杂性上。

近年来,统计学习模型的发展,使其能够较好地描述复杂高维数据的空间分布,统计学习在诸多领域的应用为人们提供了一条新的思路^[2-4],人们不再考虑寻找描述人体运动的参数或约束,而是基于运动捕获数据或已得到认可的动画数据库,通过计算机自动学习某类运动的姿势的似然来将新运动生成问题转化为根据用户给定的新约束、在姿势空间中寻找最像的姿势的问题。

1 3 维角色运动关节模型

由于 3 维角色模型是一个复杂的三角网格模型,而对运动起决定性作用的,则是关节的位置及自由度,因此需要为 3 维角色建立关节模型,以刻画角色的运动姿势。在关节模型中,人体可以看作是通过关节顺次连接而成的一系列开式运动链,并形成树的结构。每一层关节的位置是由其自身的自由度(DOF)和它的上一层的关节位置决定的。每个关节都有描述其运动能力的自由度,一般为 1~3 个,称为关节变量。在具有 d 个自由度的运动链中, d 个关节变量决定了人体的一个运动状态(称为姿势)。人体动画则是由连续变化的 d 个关节变量所决定的。本文采用的运动关节模型是在美国 Carnegie Mellon 大学 mocap 人体运动捕获数据库中人体关节模型基础上,去掉了与整体运动影响不大的手部和脚部几个关节而得到的(如图 1 所示),共有 50 个自由度。其中根结点有 6 个自由度,表示全局位置(3DOF)和方向(3DOF),通过相邻姿势的根结点全局位置相减,即可得到 3 个全局平移速度。这样,就可以用一个由根节点的全局方向角和移动速度以及各关节角组成的向量 \mathbf{y} 来表示 3 维姿势,为了后续处理的方便,应该将每个姿势向量减去均值,而一个运动则可由 m 个姿势向量组成的 $m \times d$ 维矩阵表示, m 为采样的运动帧数, d 为姿势向量的维度,即 $d = 50$ 。

在实际应用中,还可以根据对运动细节的要求程度,去掉一些不太敏感的关节及自由度,以减少姿势向量的维数。

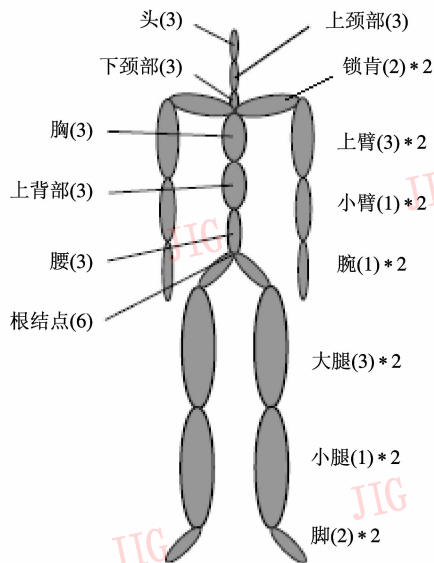


图 1 人体运动关节模型

Fig. 1 The DOFs for the human model

2 基于学习的群体动画生成

本文的目的是从已有运动数据中来学习获得对组成该运动的姿势空间的描述,即对所有可能的正确的姿势的似然建模,以便得到对一个姿势的理想程度的度量。

2.1 问题描述

由于姿势向量是高维的,因此直接在高维空间中获得所有有效姿势的概率分布是非常困难的,而对于组成一个运动的每个姿势来说,各自由度之间却存在着很大的相关性,即组成某个运动的所有姿势常常聚集在某个维数较低的子空间中,因此可以考虑对高维的姿势观察空间进行降维,用低维隐空间来更直观地表示组成该运动的所有姿势的似然。

最经典的降维方法是主成分分析(PCA),其基本思想是借助于一个正交变换,将其分量相关的原随机向量转化成其分量不相关的新随机向量,但由于主成分分析方法只是通过简单地丢弃主子空间以外的信息来达到降维的目的,故缺乏概率模型描述,而且 PCA 方法从观察空间到主子空间完全是线性映射,不足以描述复杂的人体运动空间。

考虑到人体运动的复杂性,从高维观测空间到低维隐空间的映射必定是非线性的,而且由于组成的姿势的数目是不定的,因此本文用非线性的高斯过程隐变量模型(GPLVM)^[3]来构建运动姿势学习

模型;由于人体运动的时序性,还需要在隐空间中对人体运动进行动态性建模,考虑到大多数运动可由速度和加速度来刻画,因此可在隐状态上放置两阶马尔可夫链,即用另一个高斯过程来对隐空间中相邻隐状态的进化过程进行建模,以构建高斯动态模型(GPDM)^[5]。这样就可以用得到的低维隐空间(2维或3维)中的轨迹来描述高维空间中的运动,隐空间中相邻的点在观测空间中也是相似的,其得到的隐轨迹越光滑,则模型的学习能力就越强。

假设在 d 维的观测空间(具有 d 个自由度的姿势空间)中,有一 m 帧的角色动画序列,记作 $m \times d$ 大小的矩阵 \mathbf{Y} ,矩阵 \mathbf{Y} 即为训练数据集,其每一行(\mathbf{y}_t)表示一个姿势向量(用于刻画动画的每一帧),其中 t 为时间索引,每一列表示一个给定的 DOF 的轨迹。

假设存在一个 q 维的隐空间($q < d$),则观察空间中的每一个姿势状态 \mathbf{y}_t 都在隐空间中存在一个对应的隐状态 $\mathbf{x}_t, \mathbf{x}_t$ 由前两个时间阶的隐状态 \mathbf{x}_{t-1} 和 \mathbf{x}_{t-2} 进化而来,这样就存在以下两个映射:

$$h: \mathbf{x}_t = \sum_j \mathbf{b}_j \psi_j(\mathbf{x}_{t-1}, \mathbf{x}_{t-2}) + \mathbf{n}_{x,t} \quad (1)$$

$$g: \mathbf{y}_t = \sum_i \mathbf{a}_i \phi_i(\mathbf{x}_t) + \mathbf{n}_{y,t} \quad (2)$$

其中, h 为隐空间中从前两个隐状态 $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}$ 到下一隐状态 \mathbf{x}_t 的映射,其参数为 $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots]^\top$, g 为从隐空间到观察空间的映射,其参数为 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots]^\top$; ψ_i 和 ϕ_j 为非线性基函数, $\mathbf{n}_{x,t}$ 和 $\mathbf{n}_{y,t}$ 分别是具有零均值等方性的高斯噪声。

本文的目的是获得姿势空间及相应隐空间的概率分布函数,即获得观测姿势及其隐状态的似然。

2.2 学习模型

本文用高斯过程动态模型来对人体运动进行建模,它包括从隐空间到观测空间的非线性映射 g (式(2)),和反映隐空间中隐状态演化的动态映射 h (式(1))。

高斯过程(GP)描述的就是函数上的概率分布,从概率的角度来看,人们更关心的是参数 \mathbf{A} 和 \mathbf{B} 的不确定性,而不需要像传统拟合那样去估计参数 \mathbf{A} 和 \mathbf{B} 的具体值。正如高斯随机变量由均值和方差定义一样,高斯过程由均值函数 $\mu: \mathbf{R}^q \rightarrow \mathbf{R}^d$ 和核函数(方差函数) $k: \mathbf{R}^q \rightarrow \mathbf{R}^d$ 定义,它描述的是所有函数 $g: \mathbf{R}^q \rightarrow \mathbf{R}^d$ 的分布。由于已做过减均值处理,因此均值函数可设为零均值。高斯过程的回归可描述为:通过调整协方差函数的参数,使给定观测数据的似然最大。根据高斯过程的定义,可以脱离基函数直接定义核函数,从本质上看,核函数提供了对输入

空间中两个点 \mathbf{x}_i 和 \mathbf{x}_j 的相似度度量。

对于从隐空间到观测空间的非线性映射,可采用以下径向基函数(RBF)核:

$$k_g(\mathbf{x}_i, \mathbf{x}_j) = \beta_1 \exp\left(-\frac{\beta_2}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) + \beta_3^{-1} \delta_{i,j} \quad (3)$$

β_1 为输出比例, β_2 为 RBF 核函数的宽度的倒数, β_3 是高斯噪声 $\mathbf{n}_{y,t}$ 的方差的倒数。 $\delta_{i,j}$ 为 Kronecher delta 函数(即当 $i=j$ 时,该值为 1,否则为 0)。对于 m 个输入向量 $\{\mathbf{x}_i\}$,定义 $m \times m$ 大小的核矩阵 \mathbf{K}_g ,其中 $(\mathbf{K}_g)_{i,j} = k_g(\mathbf{x}_i, \mathbf{x}_j)$ 。

给映射参数 \mathbf{B} 指定高斯先验分布 $\mathbf{b}_i \sim \mathcal{N}(\mathbf{b}_i | 0, I)$,通过边缘化 \mathbf{B} ,即可得到观测空间中的姿势向量的概率密度。

$$p(\mathbf{Y} | \mathbf{X}, \boldsymbol{\beta}) = \int \prod_{i=1}^m p(\mathbf{y}_i | \mathbf{x}_i, \mathbf{B}, \boldsymbol{\beta}) p(\mathbf{B}) d\mathbf{B} = \frac{1}{(2\pi)^{\frac{dm}{2}} |\mathbf{K}_g|^{-\frac{d}{2}}} \exp\left(-\frac{1}{2} \text{tr}(\mathbf{K}_Y^{-1} \mathbf{Y} \mathbf{Y}^\top)\right) \quad (4)$$

其中, $\boldsymbol{\beta}$ 为超参数 $\{\beta_1, \beta_2, \beta_3\}$ 。

映射 h 以两阶马尔可夫动态链为基础,其先验分布为

$$p(\mathbf{X}) = p(\mathbf{x}_1) p(\mathbf{x}_2 | \mathbf{x}_1) \prod_{i=3}^m p(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_{i-2}) \quad (5)$$

考虑到相邻运动姿势的线性变化特性(例如,局部位移变化接近线性)以及复杂运动的非线性约束(与地面接触等约束为非线性),映射 h 上高斯过程的核函数采用线性 + RBF 核,即

$$k_h(\mathbf{x}_i, \mathbf{x}_j) = \alpha_1 \exp\left(-\frac{\alpha_2}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) + \alpha_3 \mathbf{x}_i^\top \mathbf{x}_j + \alpha_4^{-1} \delta_{i,j} \quad (6)$$

其中, α_1 为输出比例, α_2 为 RBF 核函数的宽度的倒数, α_3 为线性过程的方差, α_4 是高斯噪声 $\mathbf{n}_{x,t}$ 的方差的倒数。对于 $m-2$ 个输入向量 $\mathbf{X}_{2,m-1,1,m-2}$ (其第 i 项为 $(\mathbf{x}_{i+1}, \mathbf{x}_i)$),定义 $(m-2) \times (m-2)$ 大小的核矩阵 \mathbf{K}_h ,其中 $(\mathbf{K}_h)_{i,j} = k_h(\mathbf{x}_i, \mathbf{x}_j)$ 。

给映射参数 \mathbf{A} 指定高斯先验分布 $\mathbf{a}_i \sim \mathcal{N}(\mathbf{a}_i | 0, I)$,即可通过边缘化 \mathbf{A} 来得到以下隐空间中的隐状态的联合概率密度:

$$p(\mathbf{X} | \boldsymbol{\alpha}) = \int p(\mathbf{X} | \mathbf{A}, \boldsymbol{\alpha}) p(\mathbf{A} | \boldsymbol{\alpha}) d\mathbf{A} = p(\mathbf{x}_1) p(\mathbf{x}_2 | \mathbf{x}_1) \int \prod_{i=3}^m p(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_{i-2}, \mathbf{A}, \boldsymbol{\alpha}) p(\mathbf{A} | \boldsymbol{\alpha}) d\mathbf{A} = \frac{p(\mathbf{x}_1) p(\mathbf{x}_2 | \mathbf{x}_1)}{(2\pi)^{\frac{q(m-1)}{2}} |\mathbf{K}_h|^{-\frac{q}{2}}} \exp\left(-\frac{1}{2} \text{tr}(\mathbf{K}_h^{-1} \mathbf{X}_{3,m}, \mathbf{X}_{3,m}^\top)\right) \quad (7)$$

其中, α 为超参数 $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$

对超参数赋以先验概率 $p(\alpha) \propto \prod \alpha_i, p(\beta) \propto \prod \beta_j$, 即得到所有数据维的联合似然

$$p(\mathbf{X}, \mathbf{Y}, \alpha, \beta) = p(\mathbf{Y} | \mathbf{X}, \beta) p(\mathbf{X} | \alpha) p(\alpha \beta) \quad (8)$$

2.3 学习过程

建立学习模型的目的是为 d 维姿势数据构建 q 维隐空间, 使得在姿势空间中相似的动作, 在隐空间中的位置也是相邻的。学习的过程实际上就是通过数值优化的方法来求解高斯动态模型中的未知参数的过程, 即通过调整核函数的参数 (α 和 β), 使得观测数据的似然最大。这在理论上要求知道作为训练集的观测数据 \mathbf{Y} 和对应的隐坐标 \mathbf{X} 。由于 \mathbf{X} 是未知的, 因此可以先使用 PCA 方法得到其初值, 然后迭代地执行 GP 回归来校正对 \mathbf{X} 的估计, 不断调整超参数 α 和 β , 直到收敛或到达最大的迭代次数。

在实际实现时, 可通过最小化未知参数的联合分布的负的 \log 似然, 即通过最小化下面的目标函数来学习得到 GPDM 模型的参数:

$$\begin{aligned} L_{\text{GP}} = -\ln p(\mathbf{X}, \alpha, \beta | \mathbf{Y}) = & \\ & \frac{d}{2} \ln |\mathbf{K}_g| + \frac{1}{2} \text{tr}(\mathbf{K}_g^{-1} \mathbf{Y} \mathbf{Y}^T) + \\ & \frac{q}{2} \ln |\mathbf{K}_h| + \frac{1}{2} \text{tr}(\mathbf{K}_h^{-1} \mathbf{X}_{3:m} \mathbf{X}_{3:m}^T) + \\ & \frac{1}{2} \mathbf{x}_{1,2}^T \mathbf{x}_{1,2} + \sum_j \ln \beta_j + \sum_i \ln \alpha_i + \\ & \frac{md}{2} \ln 2\pi + \frac{(m-2)q}{2} \ln 2\pi \quad (9) \end{aligned}$$

2.4 群体动画生成

群体动画的一个特点是群体中的各个个体在执行同一运动时, 个体动画之间具有很大的相似性, 而细节上却存在着差异。因此可以通过生成一系列相似而不同的运动来辅助生成群体动画。

通过人体运动学习模型, 即可获得原始姿势及其隐坐标的联合概率分布, 隐空间中相邻的点在观测空间中也是相似的, 也就是说, 如果用隐空间中的点来重构姿势空间, 那么隐轨迹附近区域的点将具有较高的可信度, 即重构出的姿势与原始运动中的姿势具有较大的相似度。因此可以考虑在隐轨迹附近随机采样, 先获得若干新的随机隐轨迹, 再将它们映射回姿势空间, 即可获得若干个与原运动非常相似, 且又完全不同的新的运动, 以便为群体中的每个个体赋予相似又不同的运动。

在隐空间中, 新的隐状态 \mathbf{x}_i 关于 $\mathbf{x}_i = (\mathbf{x}_{i-1},$

$\mathbf{x}_{i-2})$ 是高斯的, 即

$$\hat{\mathbf{x}}_i \sim \mathcal{N}(\mu_h(\hat{\mathbf{x}}_{i'}), \sigma_h^2(\hat{\mathbf{x}}_{i'}) \mathbf{I}) \quad (10)$$

$$\mu_h(\hat{\mathbf{x}}_{i'}) = \mathbf{X}_{3:m}^T \mathbf{K}_h^{-1} k_h(\hat{\mathbf{x}}_{i'}) \quad (11)$$

$$\sigma_h^2(\hat{\mathbf{x}}_{i'}) = k_h(\hat{\mathbf{x}}_{i'}, \hat{\mathbf{x}}_{i'}) - k_h(\hat{\mathbf{x}}_{i'})^T \mathbf{K}_h^{-1} k_h(\hat{\mathbf{x}}_{i'}) \quad (12)$$

其中, $k_h(\hat{\mathbf{x}}_{i'})$ 是 $m-2$ 维向量, 其第 i 个元素为 $\hat{\mathbf{x}}_{i'}$ 与训练集中向量 \mathbf{x}_i 的相似度的度量, 即 $((k_h(\hat{\mathbf{x}}_{i'}))_i = k_h(\hat{\mathbf{x}}_{i'}, \mathbf{x}_i))$

生成动画时, 首先指定初始隐状态 $(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2)$, 然后进行动态预测, 令当前隐状态 $\hat{\mathbf{x}}_i$ 是将前两步的隐状态 $\hat{\mathbf{x}}_{i'}$ 作为条件的平均隐状态, 即 $\hat{\mathbf{x}}_i = \mu_h(\hat{\mathbf{x}}_{i'})$, 在得到初始的动态预测序列后, 即可通过最大化动态似然, 即最小化 $-\ln p(\mathbf{X} | \alpha)$, 并对其进行优化来得到最接近原始观测姿势的运动序列; 最后采用 Hybrid Monte Carlo 方法^[6]生成随机采样的隐轨迹, 这些隐轨迹满足前面给定的概率分布。

观测空间中新姿势 $\hat{\mathbf{y}}_i$ 关于隐空间中的隐状态 $\hat{\mathbf{x}}_i$ 是高斯的, 即

$$\hat{\mathbf{y}}_i \sim \mathcal{N}(\mu_g(\hat{\mathbf{x}}_i), \sigma_g^2(\hat{\mathbf{x}}_i) \mathbf{I}) \quad (13)$$

$$\mu_g(\hat{\mathbf{x}}_i) = \mathbf{Y}^T \mathbf{K}_g^{-1} k_g(\hat{\mathbf{x}}_i) \quad (14)$$

$$\sigma_g^2(\hat{\mathbf{x}}_i) = k_g(\hat{\mathbf{x}}_i, \hat{\mathbf{x}}_i) - k_g(\hat{\mathbf{x}}_i)^T \mathbf{K}_g^{-1} k_g(\hat{\mathbf{x}}_i) \quad (15)$$

其中, $k_g(\hat{\mathbf{x}}_i)$ 是 m 维向量, 其第 i 个元素为新状态 $\hat{\mathbf{x}}_i$ 与训练集中向量 \mathbf{x}_i 的相似度的度量, 即 $((k_g(\hat{\mathbf{x}}_i))_i = k_g(\hat{\mathbf{x}}_i, \mathbf{x}_i))$ 。同样, 可令新姿势 $\hat{\mathbf{y}}_i = \mu_g(\hat{\mathbf{x}}_i)$ 。

3 实验结果及分析

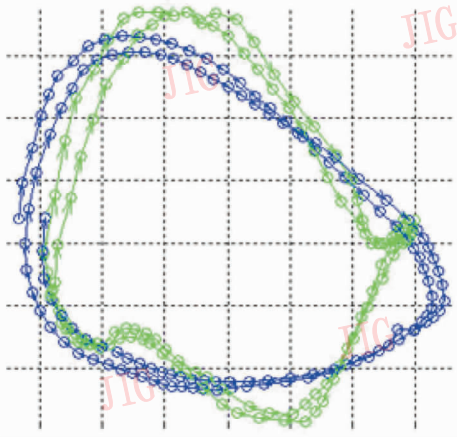
本文采用来自 Carnegie Mellon 大学 Motion Capture Database 中的人体运动捕获数据作为实验数据 (<http://mocap.cs.cmu.edu/>)。实验数据选用了行走(07_01. amc, 07_04. amc(慢走))、踢球(10_01. amc, 10_03. amc(助跑))、棒球击球(124_01. amc)等数十个运动序列, 实验结果良好。

下面以行走运动序列(07_01. amc)和踢球运动序列(10_03. amc)为例给出实验结果。

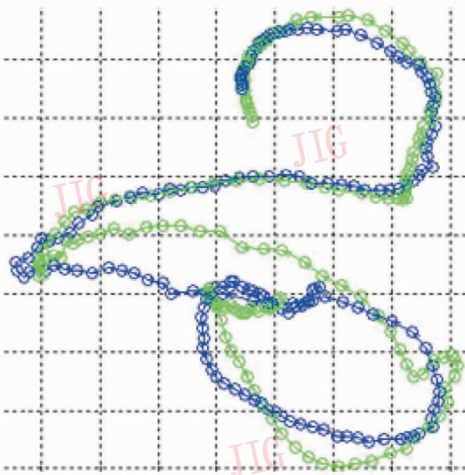
学习的过程实际上是一个数值优化的过程, 本文采用比例共轭梯度(SCG)算法来学习模型参数。在本研究及实验中, 人体模型的自由度简化为 36, 即 $d=36$, 隐空间的维数取为 3, 即 $q=3$ 。

图 2 为实验得到的学习 GPDM 隐空间。由图 2 可见, 蓝色轨迹(GPDM 隐轨迹)要比绿色轨迹(PCA 初值)光滑。

图 2(a)之所以形成一个闭环, 是因为行走运动



(a) 行走运动序列



(b) 踢球运动序列

图 2 学习得到的 GPDM 隐空间(其中绿色轨迹为 PCA 初值,蓝色轨迹为学习得到的隐轨迹)

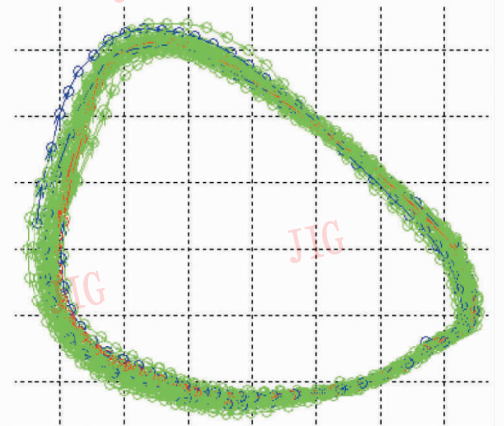
Fig. 2 GPDM latent space obtained by learning (Green trajectory is the initial value of PCA, the blue one is leaned latent trajectory)

是一种周期性的运动,每个步态周期具有很大的相似性所致。由于图 2(a)采用的观测数据为两个步态周期,所以图中隐轨迹为相似的两个闭环,同时也反映了在隐空间中相邻的点在观测空间中的姿势是相似的。图 2(b)之所以有部分闭环,是因为采用的观测数据(踢球运动)先有一个步态周期的起跑运动,然后才踢球。由于一个步态周期中的头尾姿势具有相似性,因此隐空间中的轨迹首先出现一个闭环,然后才是与踢球运动对应的轨迹。由于踢球运动中的姿势不相似,因此不形成环路。

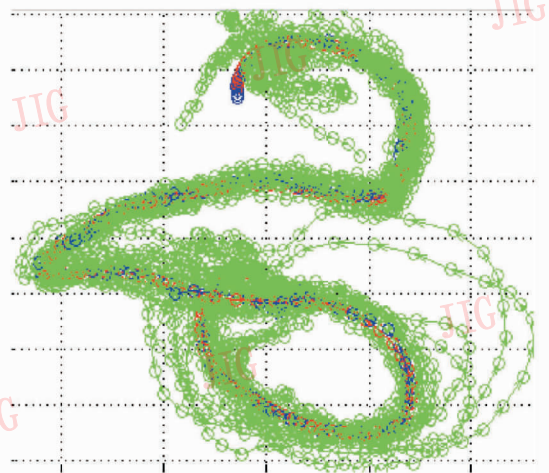
由于学习过程是一个数值优化的过程,即通过调整 x 及所有超参数的值来使得目标函数最小,因此影响计算时间(不论是模型学习还是动作合成的

时间)的主要因素是要学习的动作的数目。一般来说,由于学习时间比较长,对于 300 帧以内的简单运动,如果采用一阶马尔可夫链,通常需要几分钟,而对于复杂运动,如果采用二阶动态模型,则通常需要几十分钟(本文中的踢球序列学习过程大约为 10 ~ 15 min)。同样,采样的过程也比较耗时,这与生成采样的数目相关。所幸的是,学习和采样的过程都是在预处理阶段完成,而重构的速度相对是非常快的,可以实时完成。另外,本文目前的实验代码是用 MATLAB 编写,具体应用时可通过 C++ 来优化程序,可在一定程度上提高算法效率。

图 3 为通过 Hybrid Monte Carlo 采样得到的符合概率分布要求的一系列隐轨迹系列。由图 3 可以看到,所有新生成的隐轨迹都聚集在与观测数据对



(a) 行走运动序列



(b) 踢球运动序列

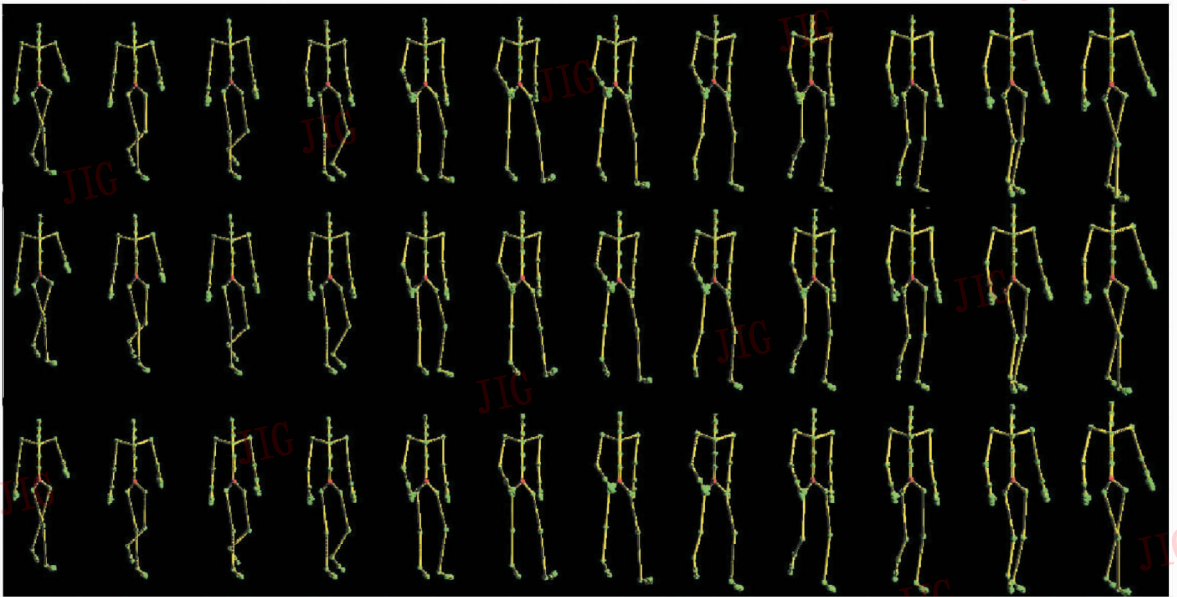
图 3 Hybrid Monte Carlo 采样得到的符合概率分布要求的隐轨迹系列

Fig. 3 A series latent trajectories obtained by Hybrid Monte Carlo sampling and the given probability distribution

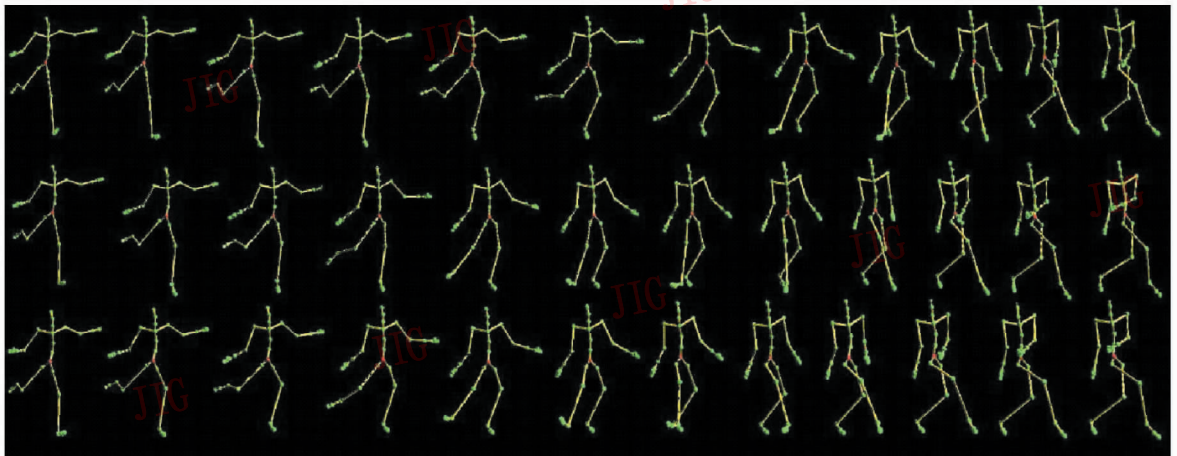
应的隐轨迹附近,而根据本文提出的学习模型,则这些区域具有较高的重构可信度,即重构得到的姿势与原姿势具有较高的相似性。

图 4 为由图 3 中的隐轨迹重构得到的运动序列。图中每一行由 Monte Carlo 采样得到的不同的隐轨迹重构得到,每一列为与不同采样对应的帧。由于篇幅所限,图中只给出了 3 条隐迹重构结果中的部分帧。由该图可以看出,行走运动序列中的 3 次重构结果之所以几乎完全相同,是因为行走运动自身是由两个步态周期构成的相对比较简单的运动,而且运动过程接近匀速运动,使得动态模型的预测结果非常准确,从而使得学习的结果非常好

(图 2(a) 中学习得到的隐轨迹平滑且均匀),采样得到隐轨迹也是紧密地分布在与原观测数据对应的隐轨迹附近。这一方面说明了本文学习模型的可靠性,但是另一方面,对于群体运动来说,由于人们并不希望生成的运动过于相似,因此对于简单运动来说,可适当地放宽正确采样的阈值,以降低采样结果的相似性。相对而言,本文的方法更适合于处理动作相对比较复杂的群体运动,图 4(b) 是踢球运动的三次重构结构。由图 4(b) 可以看出,它们与观测运动具有很大的相似性,但又存在着不同,这符合群体动画的要求。这在图 3(b) 的采样结果中也可以得到印证。



(a) 行走运动序列重构结构的部分帧(隔 5 采样)



(b) 踢球运动序列重构结构的部分帧(隔 3 采样)

图 4 重构的运动序列

Fig. 4 Motion series reconstruction(only part frames was provided for lack of space)

4 结 论

基于机器学习的群体动画生成方法可以克服传统的角色动画中运动参数和物理约束难以计算等缺点,它是从已有的角色动画的运动数据出发,先通过对组成该运动的所有姿势的似然进行建模来得到组成该运动的运动姿势统计模型;然后将复杂的高维运动姿势映射到低维隐空间中,用低维隐轨迹来刻画高维运动。由于隐空间中邻近的点,其映射到观测空间中的姿势是相似的,且隐轨迹中的隐状态的动态演化过程刻画了高维运动姿势之间的时间相关性,因此可以通过隐空间中的动态预测和采样来生成一系列服从给定概率分布的隐轨迹;最后通过姿势重构来生成一系列相似但不相同的运动。实验表明,该方法可自动生成群体动画的运动模型。而现有的大多数群体动画制作,重点都是关注群体中每个个体的位置和分布,很少为每个个体产生不同细节的运动。由此可见,基于学习的角色动画生成方法具有很大的应用潜力,而且对降低角色运动的制作难度和提高3维动画的生产效率,均具有较好的

理论意义和应用价值。

参考文献 (References)

- [1] Badler N I, Phillips C B. Simulating Humans: Computer Graphics, Animation, and Control [M]. Oxford, UK: Oxford University Press, 1992.
- [2] Lawrence N D. Learning for larger datasets with the Gaussian process latent variable model [C]//Proceedings of the 11th International Workshop on Artificial Intelligence and Statistics, Madison, WI, USA: Omnipress, 2007:243-250.
- [3] Lawrence N D. The Gaussian process latent variable model [R]. NO. CS-06-03, Sheffield, UK: The University of Sheffield, 2006.
- [4] Grochow Keith, L Martin Steven, Hertzmann Aaron, et, al. Style-based inverse kinematics [J]. ACM Transactions on Graphics, 2004, 23(3): 522-531.
- [5] Wang J M, Fleet D J, Hertzmann A. Gaussian process dynamical models [C]//Proceedings of Neural Information Processing Systems, Cambridge, MA, USA: The MIT Press, 2005: 1441-1448.
- [6] Neal Radford M. Probabilistic Inference Using Markov Chain Monte Carlo Methods [R]. CRG-TR-93-1. Toronto, Canada: University of Toronto 1993.