

基于高斯增量降维与流形 Boltzmann 优化的人体运动形态估计

李万益^{1,2}, 孙季丰¹

(1. 华南理工大学电子与信息学院, 广东广州 510641;
2. 广东第二师范学院计算机科学系, 广东广州 510303)

摘 要: 为了从多视角轮廓图像估计出含空间位置信息的三维人体运动形态, 该文提出高斯增量降维与流形 Boltzmann 优化(GIDRMBO)算法. 该算法把表示三维人体运动形态的高维数据分成表示空间位置信息和姿态信息两段子向量后, 用高斯增量降维模型(GIDRM)分别对其样本进行降维, 建立相应的低维空间及映射关系, 然后在相应的低维空间使用流形 Boltzmann 优化算法来对轮廓匹配目标函数进行优化, 从而实现估计. 其中, 所提算法分别利用了两段子向量样本的低维数据作为先验信息, 可较好的避免陷入局部最优区域进行搜索, 最终生成与各视角原始运动图像匹配且含空间位置信息的三维人体运动形态. 经仿真实验验证, 所提算法与常用粒子滤波算法相比, 其估计误差小, 并且还能起到消除轮廓数据歧义和克服短时遮挡的作用.

关键词: 高斯增量降维模型; 流形 Boltzmann 优化; 人体运动形态; 轮廓图像; 子向量

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2017)09-3060-10

10 电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.12.032

Human Motion Estimation Based on Gaussian Incremental Dimension Reduction and Manifold Boltzmann Optimization

LI Wan-yi^{1,2}, SUN Ji-feng¹

(1. School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China;
2. Department of Computer Science, Guangdong University of Education, Guangzhou, Guangdong 510303, China)

Abstract: In the interest of estimating the human motion in 3D (3 Dimensions, 3D) and its spatial position from multi-view silhouettes, an algorithm called Gaussian incremental dimension reduction and manifold Boltzmann optimization (GIDRMBO) is proposed. After the high dimensional data denoting human motion in 3D is divided into two subvectors, the two subvectors denote information of spatial position and pose respectively. The proposed algorithm takes the advantage of Gaussian incremental dimension reduction model (GIDRM) to reduce the dimension of samples of the two subvectors respectively, so that the low dimensional spaces and relevant mappings can be built, and the optimization of objective function matching silhouettes can be carried out to achieve the estimation through the manifold Boltzmann optimization in the low dimensional space. The proposed algorithm can utilize the prior informations from low dimensional data of the samples of the two subvectors respectively. It can skip the area of local optimization better during searching, and generate human motion in 3D and its spatial position which can match the original motion images of multi-view. By the comparison of some traditional particle filters in the experiments, the proposed algorithm has better performance in lowering estimation error, disambiguating the silhouettes and overcoming the transient occlusion.

Key words: Gaussian incremental dimension reduction model (GIDRM); manifold Boltzmann optimization; human motion; silhouettes; subvector

1 引言

三维人体运动形态估计已成功应用于多个领域^[1,2],其具有较高研究价值,如 3D(3 Dimensions,3D)电影动画中的立体人物制作,3D 立体投影仪的人体成像,人体运动三维重构,实现对机器人的操作和控制等等都涉及到该研究.近些年,较多学者研究如何从有限视角的轮廓图像估计三维人体运动形态,提出了些可行的方法,但是其也有不少局限性.首先,一些的启发式算法如进化算法等^[3]经改进后来求解所建立复杂的目标函数,计算效率很低,搜索速度很慢,很难搜索到最优目标数据.其次,文献[4,5]中用标准或者改进的粒子滤波算法进行估计,容易出现空间位置及肢体凌乱的现象,计算时间长,不能有效消除轮廓数据歧义和克服短时遮挡进行估计,并且非常依赖上一时刻估计的高维数据.再次,文献[6]提出退火粒子群优化进行估计,用 PCA(Principal Component Analysis,PCA)降维所得到的低维数据还原高维数据后,所还原的高维数据和原始相应的高维数据差别较大,计算时间也长,效果也不理想,并且没有考虑三维人体运动形态所对应轮廓图像的空间位置信息.最后,文献[7]提出用子空间高斯过程隐变量模型(Subspace Gaussian Process Latent Variable Model,SGPLVM)对轮廓图像和相应姿态样本的高维数据进行降维,在其降维后无法区分出歧义的低维数据点,降维过程中的计算较复杂,效果多数不好,可降维的数据样本范围很有限,也没有考虑对轮廓图像中相应的空间位置信息进行估计.综上所述,估计三维人体运动形态关键在于获取较好的先验信息,利用该先验信息进行搜索才能实现较高效的估计.

由于三维人体运动形态由用高维数据描述的姿态样本表示,所以本文结合该高维数据特点,针对上述的问题而提出高斯增量降维与流形 Boltzmann 优化(Gaussian Incremental Dimension Reduction and Manifold Boltzmann Optimization,GIDRMBO)算法,用其来估计与各视角原始运动图像匹配且含空间位置信息的三维人体运动形态.

本文所提 GIDRMBO 算法理论中主要贡献如下.首先,高斯过程动态模型^[8](Gaussian Process Dynamical Model,GPDM)不能实现增量降维,本文建立增量降维映射关系模型来进行完善,并以文献[9]为基础进行推导,给出该模型的学习算法,即多输出正交最小二乘学习(Orthogonal Least Squares Learning of Multiple Outputs,OLSLMO)算法,使其在学习后可以实现增量降维来对低维空间进行搜索定位,改进后称其为高斯增量降维模型(Gaussian Incremental Dimension Reduction Model,GIDRM).其次,受退火粒子群优化算法^[6]启发,本文还

提出了流形 Boltzmann 优化算法来对 GIDRM 降维后所建立的低维空间进行搜索.通过该算法的搜索,对轮廓匹配目标函数进行优化,最终可以估计出相应含空间位置信息的三维人体运动形态.该算法引入了 Boltzmann 分布函数,结合系统采样法,可以跳出局部最优区域进行全局搜索,还有极值记忆和系统选择数据能力.

以上两点贡献使得所提 GIDRMBO 算法具有低维空间搜索区域定位的准确性,生成运动姿态样本的有效性,系统选择数据区域的全局性.所提算法非常适用于空间位置及姿态信息变化规则的人体运动形态,可以满足日常大多数人体运动形态的估计.因此,GIDRMBO 算法与各种常用粒子滤波算法相比,其估计误差小,且能够起到消除轮廓数据歧义和克服短时遮挡的作用.

2 数据模型

本文先介绍一下数据模型,所用图像数据来自 HumanEva-I 数据库^[10],图 1(a)的模型为人体运动立体模型,该模型描述了含空间位置信息的三维人体运动形态,为要估计的对象,它要和各视角原始运动图像的人体肢体匹配,用高维数据 $\mathbf{y} \in \mathbf{R}^D$ 表示.各视角原始运动图像如图 1(b)所示,经过图像分割算法^[11-13]处理后得到相应视角的轮廓图像,该图像用于与人体运动立体模型在相应视角的二值投影图像进行重叠比较,如图 1(c)所示,白色轮廓为原始运动图像的轮廓图像,灰色轮廓为模型的二值投影图像.已知经过图像分割处理所得多视角人体运动的轮廓图像和人体运动立体模型样本相应视角的二值投影图像,通过对这两种图像的像素重叠比较后,估计出与原始多视角人体运动图像相匹配的人体运动立体模型,从而实现人体运动形态的估计.这里,人体运动立体模型是未知的.总而言之,即从已知的原始多视角人体运动图像,估计出相应与之匹配含有空间位置信息的三维人体运动立体模型.

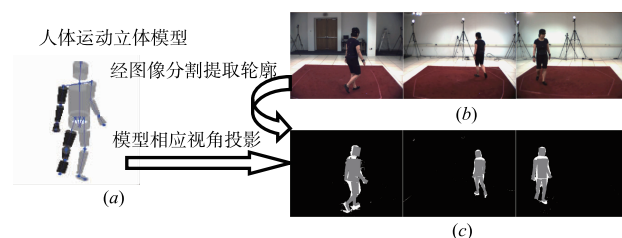


图1 相关数据模型示意图

本文在以上研究的基础上,先做一些定义.其定义如下, t 为图像帧序数, k 为镜头视角序数, α 为权重系数, p 为二维图像的像素点. M_t^i 为轮廓图像的像素点集合, M_t^k 为人体运动立体模型在相应视角投影图像的像素点集合. $M_t^i(p)$ 为像素点 p 是否在轮廓图像里的统计

函数($M_i^f(p) = 1, p \in M_i^f; M_i^f(p) = 0, p \notin M_i^f$), $M_i^b(p)$ 为像素点 p 是否在投影图像里的统计函数($M_i^b(p) = 1, p \in M_i^b; M_i^b(p) = 0, p \notin M_i^b$). 在第 k 个视角, $R_i^k = \sum_p (M_i^f(p)(1 - M_i^b(p)))$ 为像素点 p 只在轮廓图像里而不在投影图像的统计函数值, $B_i^k = \sum_p (M_i^b(p)(1 - M_i^f(p)))$ 为像素点 p 只在投影图像里而不在轮廓图像的统计函数值, $H_i^k = \sum_p (M_i^f(p)M_i^b(p))$ 为像素点 p 同时在轮廓图像和投影图像里的统计函数值. 在 N_v 个视角下可建立轮廓匹配目标函数:

$$F(\mathbf{y}) = F({}^1\mathbf{y}, {}^2\mathbf{y}) = -\frac{S}{N_v} \sum_{k=1}^{N_v} \left((1 - \alpha) \frac{B_i^k}{B_i^k + H_i^k} + \alpha \frac{R_i^k}{R_i^k + H_i^k} \right) \quad (1)$$

其中, 参数 S, α 通常分别取 $S = 100, \alpha = 0.5$, 如参考文献[4,6]的取法. 下面再定义 $\mathbf{y}(n:m) \in \mathbf{R}^{m-n+1}$ 为取 $\mathbf{y} \in \mathbf{R}^{D_r}$ 的第 n 维数据到第 m 维数据的子向量, $m > n, m, n \in \mathbf{N}^+, n:m$ 表示集合 $\{n, n+1, \dots, m\}$. 因此, 式(1)中 ${}^1\mathbf{y} = \mathbf{y}(1:D_1) \in \mathbf{R}^{D_1}$ 和 ${}^2\mathbf{y} = \mathbf{y}(D_1+1:D_r) \in \mathbf{R}^{D_2}$ 分别为 \mathbf{y} 的子向量, $D_2 = D_r - D_1, \mathbf{y} = [{}^1\mathbf{y}^T, {}^2\mathbf{y}^T]^T$. 式(1)与人体运动立体模型有着紧密的关系, 该式的值越大说明模型与各视角运动图像匹配得越好, 可对其进行优化从而实现人体运动立体模型的估计.

3 高斯增量降维与流形 Boltzmann 优化

3.1 GIDRM 的子向量样本降维及子向量关系

对于表示三维人体运动形态的高维数据, 我们采用 GIDRM 分别对其子向量样本进行降维, 所降的维数

为 3, 如图 2 所示. 在人体运动立体模型的高维数据 $\mathbf{y} \in \mathbf{R}^{D_r}$ 的各维数据中, 有部分维(前 D_1 维)数据 ${}^1\mathbf{y} = \mathbf{y}(1:D_1) \in \mathbf{R}^{D_1}$ 表示空间位置信息, 其为人体运动立体模型中心位于三维空间的具体位置, 其余部分维(后 D_2 维, $D_2 = D_r - D_1$)数据 ${}^2\mathbf{y} = \mathbf{y}(D_1+1:D_r) \in \mathbf{R}^{D_2}$ 表示姿态信息, 其为人体运动立体模型各肢体相对于模型中心的具体位置, 这两部分数据正好构成了 $\mathbf{y} \in \mathbf{R}^{D_r}$ 的子向量, 即 $\mathbf{y} = [{}^1\mathbf{y}^T, {}^2\mathbf{y}^T]^T$, 这也是根据这两部分数据的信息含义划分子向量. 然而, 相对于降维后的低维(三维)数据, 子向量 ${}^1\mathbf{y}$ 和 ${}^2\mathbf{y}$ 也为高维数据, 但其维数都比 \mathbf{y} 低, 其样本就称为子向量样本. 子向量样本降维即为用 GIDRM 分别对 ${}^1\mathbf{y}$ 和 ${}^2\mathbf{y}$ 的数据样本进行降维, 分别获取各自的低维数据, 建立相应的低维空间及映射关系. 因为 \mathbf{y} 的子向量样本降维后, 其低维数据构成如图 2 所示的规则近环状流形, 容易看出数据规律并且进行搜索, 如果直接用 \mathbf{y} 的样本降维则其低维数据就不会构成规则的近环状流形, 而是无规律的, 不利于增量降维映射关系的建立与在该低维空间的搜索, 所以需把 \mathbf{y} 分成两段子向量, 并且对其样本降维. 另外, ${}^1\mathbf{y}$ 和 ${}^2\mathbf{y}$ 关系是较密切的. 在使用 HumanEva-I 数据库的数据时, 人体运动立体模型(图 2 所示)里是有关节点标记的, \mathbf{y} 里的各维数据信息就为这些关节节点标记的信息, 其维数 $D_r = 40$, 其中 ${}^1\mathbf{y}$ 的维数 $D_1 = 6, {}^2\mathbf{y}$ 的维数 $D_2 = 34. {}^1\mathbf{y}$ 的信息是人体运动立体模型根节点标记的空间坐标及模型旋转角度数据, ${}^2\mathbf{y}$ 的信息是人体运动立体模型其余关节节点标记的旋转角度数据.

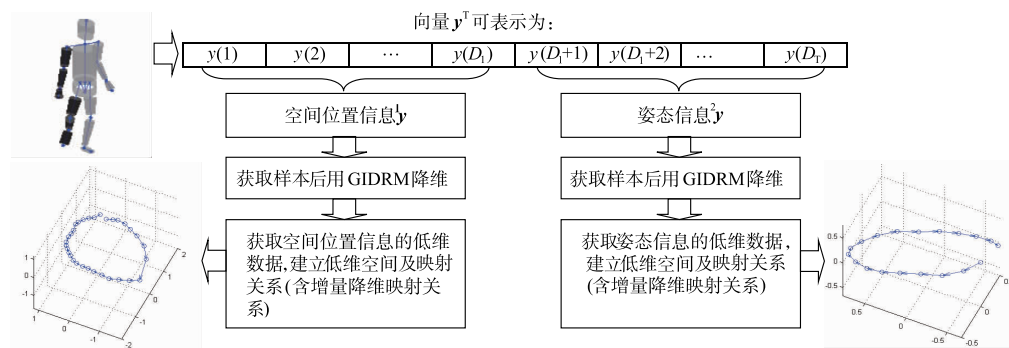


图2 子向量样本降维示意图

3.2 GIDRM 的主要理论与增量降维

由于 GIDRM 具有增量降维映射关系, 该映射关系有利于低维空间的搜索. 该模型部分概率关系采用 GPDM 的概率关系, 其主要概率关系如下^[8]:

$$p(\mathbf{Y}|\mathbf{X}, \bar{\boldsymbol{\beta}}, \mathbf{W}) = \frac{|\mathbf{W}|^N}{\sqrt{(2\pi)^{ND} |\mathbf{K}_Y|^D}} \exp\left(-\frac{1}{2} \text{tr}(\mathbf{K}_Y^{-1} \mathbf{Y} \mathbf{W}^2 \mathbf{Y}^T)\right) \quad (2)$$

$$p(\mathbf{X}|\bar{\boldsymbol{\alpha}}) = \frac{p(\mathbf{x}_1)}{\sqrt{(2\pi)^{(N-1)q} |\mathbf{K}_x|^q}} \exp\left(-\frac{1}{2} \text{tr}(\mathbf{K}_x^{-1} \mathbf{X}_{2:N} \mathbf{X}_{2:N}^T)\right) \quad (3)$$

其中, $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_i, \dots, \mathbf{y}_N]^T \in \mathbf{R}^{N \times D}, \mathbf{y}_i \in \mathbf{R}^D$ 为高维(D 维)数据序列, $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N]^T \in \mathbf{R}^{N \times q}, \mathbf{x}_i \in \mathbf{R}^q$ 为低维(q 维)数据序列, $D > q, \mathbf{K}_Y \in \mathbf{R}^{N \times N}, (\mathbf{K}_Y)_{i,j} =$

$k_Y(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\frac{\beta_1}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2) + \beta_2^{-1} \delta_{xx}$, 为带核参数 $\bar{\beta} = [\beta_1, \beta_2]$ 核矩阵, \mathbf{W} 为 D 维尺度参数且 $p(\mathbf{W}) = \prod_{m=1}^D (\frac{2}{\kappa \sqrt{2\pi}}) \exp(-\frac{w_m^2}{2\kappa^2})$, $w_m > 0, \kappa = 10^{-3}, \mathbf{K}_X \in \mathbf{R}^{(N-1) \times (N-1)}$, $(\mathbf{K}_X)_{i,j} = k_X(\mathbf{x}_i, \mathbf{x}_j) = \alpha_1 \exp(-\frac{\alpha_2}{2} \|\mathbf{x}_i - \mathbf{x}_j\|^2) + \alpha_3 \mathbf{x}_i^T \mathbf{x}_j + \alpha_4^{-1} \delta_{xx}$, 为带核参数 $\bar{\alpha} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$ 核矩阵, 通过 $\mathbf{X}_{1:N-1} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N-1}]^T$ 构造, 然而 $\mathbf{X}_{2:N} = [\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N]^T$, \mathbf{x}_1 服从 q 维高斯先验, 核参数 $\bar{\alpha}, \bar{\beta}$ 满足 $p(\bar{\alpha}) \propto \prod_i \alpha_i^{-1}, p(\bar{\beta}) \propto \prod_i \beta_i^{-1}$. 其条件概率有^[8]:

$$p(\mathbf{Y}^* | \mathbf{X}^*, \mathbf{I}) = \frac{|\mathbf{W}|^M}{\sqrt{(2\pi)^{MD} |\mathbf{K}_Y|^D}} \exp(-\frac{1}{2} \text{tr}(\mathbf{K}_Y^{-1} \mathbf{Z}_Y \mathbf{W}^2 \mathbf{Z}_Y^T)) \quad (4)$$

其中, 式(4)也为高斯分布, $\mathbf{I} = \{\mathbf{X}, \mathbf{Y}, \bar{\alpha}, \bar{\beta}, \mathbf{W}\}$, $\mathbf{Y}^* \in \mathbf{R}^{M \times D}$, $\mathbf{X}^* \in \mathbf{R}^{M \times q}$, $\mathbf{Z}_Y = \mathbf{Y}^* - \mathbf{A}^T \mathbf{K}_Y^{-1} \mathbf{Y}$, $\mathbf{K}_Y = \mathbf{B} - \mathbf{A}^T \mathbf{K}_X^{-1} \mathbf{A}$, 核矩阵 $\mathbf{A} \in \mathbf{R}^{N \times M}$, $\mathbf{A}_{i,j} = k_Y(\mathbf{x}_i, \mathbf{x}_j^*)$, $\mathbf{B} \in \mathbf{R}^{M \times M}$, $\mathbf{B}_{i,j} = k_Y(\mathbf{x}_i^*, \mathbf{x}_j^*)$.

已知高维数据序列 \mathbf{Y} , 有 $p(\mathbf{Y})$ 为常数, $\min(-\ln p(\mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W} | \mathbf{Y}))$ 等价于 $\min(-\ln p(\mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W}, \mathbf{Y}))$, 按下式获取相应低维数据 (如图 2 所示) 和相关参数, 如下^[8]:

$$\begin{aligned} \mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W} &= \arg \min_{\mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W}} (-\ln p(\mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W}, \mathbf{Y})) \\ &= \arg \min_{\mathbf{X}, \bar{\alpha}, \bar{\beta}, \mathbf{W}} (-\ln(p(\mathbf{Y} | \mathbf{X}, \bar{\beta}, \mathbf{W}) \cdot p(\mathbf{X} | \bar{\alpha}) p(\bar{\alpha}) p(\bar{\beta}) p(\mathbf{W}))) \end{aligned} \quad (5)$$

式(5)优化过程为非监督学习过程, 可用尺度共轭梯度 (Scaled Conjugate Gradient, SCG) 算法进行, 该算法可以具体参照文献[8, 14, 15]. 根据式(4), $\mathbf{y}^* \sim \mathbf{N}(\mu_Y(\mathbf{x}^*), \sigma_Y^2(\mathbf{x}^*) \mathbf{I})$, $\mathbf{y}^* \in \mathbf{R}^D$, $\mathbf{x}^* \in \mathbf{R}^q$, 低维数据到高维数据映射关系可通过均值计算给出^[8]:

$$\begin{aligned} \mathbf{y}^* &= f(\mathbf{x}^*) = \mu_Y(\mathbf{x}^*) \\ &= \mathbf{Y}^T \mathbf{K}_Y^{-1} [\mathbf{K}_Y(\mathbf{x}_1, \mathbf{x}^*), k_Y(\mathbf{x}_2, \mathbf{x}^*), \dots, \\ &\quad \mathbf{K}_Y(\mathbf{x}_N, \mathbf{x}^*)]^T = \mathbf{Y}^T \mathbf{K}_Y^{-1} \mathbf{K}_Y(\mathbf{x}^*) \end{aligned} \quad (6)$$

式(1)~式(6)为 GPDM 原有的概率与映射关系, 其是无法获取新高维数据样本 $\mathbf{y}^* \in \mathbf{R}^D$ 的相应低维数据, 从而无法快速确定低维空间的搜索位置. 因此, 其需要建立增量降维映射关系对此进行改进. 该关系我们采用如下模型进行监督学习后来建立:

$$\mathbf{X} = \Phi \mathbf{W}_D \quad (7)$$

其中, $\Phi \in \mathbf{R}^{N \times N_k}$ 为基函数矩阵, $\Phi_{k_1, k_2} = \varphi(\mathbf{y}_{k_1}, \mathbf{c}_{k_2}) = \exp(b \|\mathbf{y}_{k_1} - \mathbf{c}_{k_2}\|^2)$, $\mathbf{W}_D \in \mathbf{R}^{N_k \times q}$ 为权重矩阵, $N_k \leq N$. 在基函数参数 b 为已知设定值的情况下, 获取所有参数

$\mathbf{c}_{k_2}, k_2 = 1, 2, \dots, N_k$ 来确定 Φ, \mathbf{W}_D . 设学习完成后 \mathbf{W}_D 的值为 $\widehat{\mathbf{W}}_D$ (最小二乘估计值), 这时, \mathbf{y}^* 与 \mathbf{x}^* 可建立增量降维映射关系式如下:

$$\mathbf{x}^* = g(\mathbf{y}^*) = (\Phi(\mathbf{y}^*) \widehat{\mathbf{W}}_D)^T = \widehat{\mathbf{W}}_D^T (\Phi(\mathbf{y}^*))^T \quad (8)$$

其中, $\Phi(\mathbf{y}^*) = [\phi(\mathbf{y}^*, \mathbf{c}_1), \phi(\mathbf{y}^*, \mathbf{c}_2), \dots, \phi(\mathbf{y}^*, \mathbf{c}_{N_k})]$. 式(7)的学习为多维数据输入与多维数据输出的学习, 因此, 我们在此给出一种多输出正交最小二乘 (OLSLMO) 算法, 把最小二乘估计值 $\widehat{\mathbf{W}}_D$ 带入式(7), 有:

$$\mathbf{X} = \widehat{\Phi} \widehat{\mathbf{W}}_D + \mathbf{e} \quad (9)$$

其中, $\mathbf{e} \in \mathbf{R}^{N \times N_k}$ 为误差矩阵, 可表示为 $\mathbf{e} = [\mathbf{e}_1, \dots, \mathbf{e}_{k_1}, \dots, \mathbf{e}_N]^T = [\mathbf{e}'_1, \dots, \mathbf{e}'_{v_1}, \dots, \mathbf{e}'_q] \in \mathbf{R}^{N \times q}$. $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{k_1}, \dots, \mathbf{x}_N]^T = [\mathbf{x}'_1, \dots, \mathbf{x}'_{v_1}, \dots, \mathbf{x}'_q] \in \mathbf{R}^{N \times q}$, $\widehat{\mathbf{W}}_D = [\widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_{k_2}, \dots, \widehat{\mathbf{w}}_{N_k}]^T = [\widehat{\mathbf{w}}'_1, \dots, \widehat{\mathbf{w}}'_{v_1}, \dots, \widehat{\mathbf{w}}'_q] \in \mathbf{R}^{N_k \times q}$, 矩阵 $\Phi = [\varphi_1, \dots, \varphi_{k_2}, \dots, \varphi_{N_k}] \in \mathbf{R}^{N \times N_k}$ 可分解为 $\Phi = \widetilde{\mathbf{W}} \widetilde{\mathbf{A}}$, 其中, $\widetilde{\mathbf{W}} = [\widetilde{\mathbf{w}}_1, \dots, \widetilde{\mathbf{w}}_{k_2}, \dots, \widetilde{\mathbf{w}}_{N_k}] \in \mathbf{R}^{N \times N_k}$ 为正交矩阵 ($\widetilde{\mathbf{w}}_{k_2}^T \widetilde{\mathbf{w}}_{k'_2} = 0, k_2 \neq k'_2; \widetilde{\mathbf{w}}_{k_2}^T \widetilde{\mathbf{w}}_{k'_2} \neq 0, k_2 = k'_2$), $\widetilde{\mathbf{A}} \in \mathbf{R}^{N_k \times N_k}$ 为可逆矩阵 ($\widetilde{\mathbf{A}}_{i,j} = a_{i,j}, i < j; \widetilde{\mathbf{A}}_{i,j} = 1, i = j; \widetilde{\mathbf{A}}_{i,j} = 0, i > j$). 然后, 最小二乘估计值 $\widehat{\mathbf{W}}_D = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{X}$, 令 $\widehat{\mathbf{g}} = (\widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}})^{-1} \widetilde{\mathbf{W}}^T \mathbf{X}$, 为关于 $\widetilde{\mathbf{W}}$ 的最小二乘估计值, 可表示为 $\widehat{\mathbf{g}} = [\mathbf{g}_1, \dots, \mathbf{g}_{k_2}, \dots, \mathbf{g}_{N_k}]^T = [\widehat{\mathbf{g}}_1, \dots, \widehat{\mathbf{g}}_{v_1}, \dots, \widehat{\mathbf{g}}_q] \in \mathbf{R}^{N_k \times q}$, 然后推出:

$$\begin{aligned} \Phi \widehat{\mathbf{W}}_D &= \Phi (\Phi^T \Phi)^{-1} \Phi^T \mathbf{X} \\ &= \widetilde{\mathbf{W}} \widetilde{\mathbf{A}} (\widetilde{\mathbf{A}}^T \widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}} \widetilde{\mathbf{A}})^{-1} \widetilde{\mathbf{A}}^T \widetilde{\mathbf{W}}^T \mathbf{X} \\ &= \widetilde{\mathbf{W}} (\widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}})^{-1} \widetilde{\mathbf{W}}^T \mathbf{X} = \widetilde{\mathbf{W}} \widehat{\mathbf{g}} \end{aligned} \quad (10)$$

因此, 可将式(9)改写成下式:

$$\mathbf{X} = \widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \mathbf{e} \quad (11)$$

根据最小二乘的性质, $\widetilde{\mathbf{W}}^T \mathbf{e} = \mathbf{0}, \mathbf{e}^T \widetilde{\mathbf{W}} = \mathbf{0}$, 可得出:

$$\begin{aligned} \mathbf{X}^T \mathbf{X} &= (\widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \mathbf{e})^T (\widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \mathbf{e}) \\ &= \widehat{\mathbf{g}}^T \widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \widehat{\mathbf{g}}^T \widetilde{\mathbf{W}}^T \mathbf{e} + \mathbf{e}^T \widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \mathbf{e}^T \mathbf{e} \\ &= \widehat{\mathbf{g}}^T \widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}} \widehat{\mathbf{g}} + \mathbf{e}^T \mathbf{e} \end{aligned} \quad (12)$$

因此, 可得到 $\mathbf{e}^T \mathbf{e} = \mathbf{X}^T \mathbf{X} - \widehat{\mathbf{g}}^T \widetilde{\mathbf{W}}^T \widetilde{\mathbf{W}} \widehat{\mathbf{g}}$. 学习过程中, 在使用 N_k 个正交向量的情况下, 使得误差矩阵范数 (2 范数) 平方的均值极小化即可, 所以有:

$$\begin{aligned} \min(\frac{\|\mathbf{X} - \Phi \widehat{\mathbf{W}}_D\|^2}{(N \times q)}) &= \min(\|\mathbf{X} - \widetilde{\mathbf{W}} \widehat{\mathbf{g}}\|^2 / (N \times q)) \\ &= \min(\text{tr}(\mathbf{e}^T \mathbf{e}) / (N \times q)) \\ &= \min(\frac{1}{q} \sum_{v_1=1}^q (\frac{1}{N} (\mathbf{e}'_{v_1})^T \mathbf{e}'_{v_1})) \\ &= \min(\frac{1}{q} \sum_{v_1=1}^q (\frac{1}{N} (\mathbf{x}'_{v_1})^T \mathbf{x}'_{v_1} \\ &\quad - \frac{1}{N} \sum_{k_2=1}^{N_k} \widetilde{\mathbf{w}}_{k_2}^T \widetilde{\mathbf{w}}_{k_2} \widehat{\mathbf{g}}_{v_1}^2(k_2))) \end{aligned} \quad (13)$$

极小化式(13)等价于:

$$\max_{S_w} \left(\sum_{v_1=1}^q \left(\sum_{k_2=1}^{N_k} (\tilde{\mathbf{w}}_{k_2}^T \tilde{\mathbf{w}}_{k_2} \hat{\mathbf{g}}_{k_2}^2) / (\mathbf{x}'_{v_1})^T \mathbf{x}'_{v_1} \right) \right), \quad (14)$$

$$S_w \in S'_w, N_k \leq N$$

式(14)中, $S_w = \{\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_{k_2}, \dots, \tilde{\mathbf{w}}_{N_k}\}$ 与 $S'_w = \{\tilde{\mathbf{w}}'_1, \dots, \tilde{\mathbf{w}}'_{k_2}, \dots, \tilde{\mathbf{w}}'_{N_k}\}$ 均为正交向量集合, 即从 S'_w 中选出子集 S_w 使得式(14)极大化. 其中, S'_w 为矩阵 $\Phi' = [\varphi'_1, \dots, \varphi'_{k_2}, \dots, \varphi'_{N_k}] \in \mathbf{R}^{N \times N}$ 中列向量 φ'_{k_i} 的 Schmidt 正交化向量集合, $\Phi'_{k_1, k_2} = \phi(\mathbf{y}_{k_1}, \mathbf{y}_{k_2}) = \exp(\mathbf{b} \|\mathbf{y}_{k_1} - \mathbf{y}_{k_2}\|^2)$. 当满足 $\|\mathbf{X} - \Phi \mathbf{W}_D\|^2 / (N \times q) < \varepsilon_1, \varepsilon_1 > 0$, 学习即可完成, S_w 选择过程中需调节 N_k 的值, 使得 N_k 的值尽可能小的满足上述条件. 下面以为伪代码形式, 给出式(7)的学习(OLSLMO)算法.

算法 1 OLSLMO

```

输入:  $\mathbf{X}, \mathbf{Y}, b$ ;
输出:  $c_{k_2}, k_2 = 1, 2, \dots, N_k$ ;
For  $k = 1 : 1 : N$ 
If  $k = 1$ 
 $\mathbf{w}_k^s = \varphi'_s, s = 1, 2, \dots, N$ ;
Else If  $k \geq 2$ 
 $\alpha_{ik}^s = \tilde{\mathbf{w}}_i^T \varphi'_s / (\tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i), s = 1, 2, \dots, N, s \notin \{1n_1, 1n_2, \dots, 1n_{k-1}\}, 1 \leq i < k$ ;
 $\mathbf{w}_k^s = \varphi'_s - \sum_{i=1}^{k-1} \alpha_{ik}^s \tilde{\mathbf{w}}_i$ ;
End If
 $\mathbf{g}_k^s = (\mathbf{w}_k^s)^T \mathbf{X} / ((\mathbf{w}_k^s)^T \mathbf{w}_k^s)$ ;
 $V^s = \sum_{v_1=1}^q ((\mathbf{w}_k^s)^T \mathbf{w}_k^s (\mathbf{g}_k^s(v_1))^2) / (\mathbf{x}'_{v_1})^T \mathbf{x}'_{v_1}$ ;
 $s^* = \arg \max_s V^s; \tilde{\mathbf{w}}_k = \mathbf{w}_k^{s^*}; 1n_k = s^*$ ;
 $k_2 = k; c_{k_2} = y_{1n_k}; \mathbf{W}_D = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{X}$ ;
If  $\|\mathbf{X} - \Phi \mathbf{W}_D\|^2 / (N \times q) < \varepsilon_1$ 
Break;
End If
End For
 $N_k = k$ ;

```

当 GIDRM 学习完成后, 我们可用其观察新子向量样本的相应低维数据(低维流形), 对低维空间快速而较准确的定位搜索. 两段子向量分别表示空间位置信息和姿态信息, 下面我们把相应帧的真实子向量样本(用于降维的数据样本除外)通过 GIDRM 进行增量降维. 如图 3(a) 和图 3(b) 所示, 我们可以发现相应帧的真实子向量经过 GIDRM 增量降维后, 其相应低维空间中的低维数据(绿色空心圆圈标识)分布在已知低维数据(蓝色实心圆圈标识)附近, 并且围绕排列在相邻两个已知低维数据的相连直线之间. 下面我们再分别对两种搜索方式做比较. 首先, 分别在空间位置信息与姿

态信息的低维空间里以任意一个已知低维数据为起点, 通过非线性搜索后, 所获得的最终高维数据生成人体运动立体模型如图 3(c) 所示; 其次, 我们相应的高维数据作为搜索起点, 直接在低维空间进行非线性搜索, 所获得的最终高维数据生成人体运动立体模型如图 3(d) 所示. 从图 3(c) 和图 3(d) 可以看出, 在低维空间搜索后生成的人体运动立体模型要比在高维空间搜索后生成的更有效. 前者所生成模型的空间位置较正确, 其肢体成人形(数据有效); 而后者所生成模型的空间位置错误, 其肢体凌乱不成人形(数据无效). 由于描述人体运动形态的高维数据各维之间有着复杂的函数关系, 直接在最终高维数据空间搜索较难生成有效数据, 所以用 GIDRM 分别对表示空间位置信息与姿态信息的子向量样本降维后, 获得的低维数据信息可以对我们的搜索提供重要的依据, 并以该数据为先验信息能

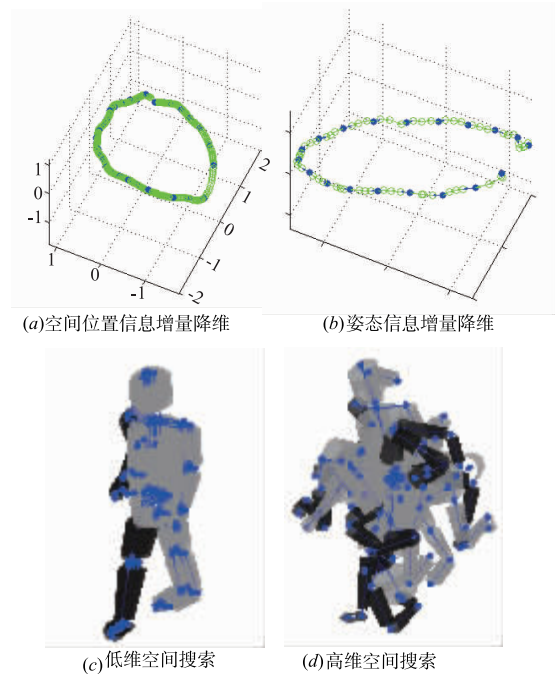


图3 增量降维的优势描述示意图

起到很好辅助作用. 与一些降维算法或模型相比, 采用 GIDRM 进行降维, 其优势是明显的. 例如, 一些算法或模型 PCA^[6], ISOMAP^[16], LLE^[17], SGPLVM^[7] 以及 GPDM^[8], 此外还有一些更早的算法等等, 其都没有增量降维映射关系, 不能对新数据样本进行降维, 部分算法或模型还原的高维数据会和原始相应的高维数据有很大偏差, 会造成高维数据信息损失.

3.3 基于 GIDRM 映射关系的流形 Boltzmann 优化

只有 GIDRM 还不足以估计出所需含空间位置信息的三维人体运动形态, 所以根据上述的情况, 我们可

以推出轮廓匹配的流形 Boltzmann 优化算法,搜索匹配多视角轮廓图像的三维人体运动形态,从而实现估计.该方法借助了 GIDRM 的映射关系,在低维空间进行搜索来实现式(1)的优化.设在已知估计上一时刻的高维数据 $\mathbf{y}_{i-1}^* = [({}^1\mathbf{y}_{i-1}^*)^\top, ({}^2\mathbf{y}_{i-1}^*)^\top]^\top \in \mathbf{R}^{D_i}$ 中, ${}^s\mathbf{y}_{i-1}^* \in \mathbf{R}^D$, $s = 1, 2, D = D_1 - D_2 = D_T - D_1$ 为其相应各段子向量,用式(8)增量降维映射关系获取其低维数据 $\mathbf{x}_i^* \in \mathbf{R}^q$ 后进行聚类,然后在已知低维数据 ${}^s\mathbf{x}_i$, $i = 1, 2, \dots, N_s$, $s = 1, 2$ 中,可得到其低维空间的最近低维数据点:

$$\begin{aligned} i^* &= \arg \min_i \|\mathbf{x}_i^* - {}^s\mathbf{x}_i\| \arg \min_i \|g({}^s\mathbf{y}_{i-1}^*) - {}^s\mathbf{x}_i\|, \\ i &= 1, 2, \dots, N_s \end{aligned} \quad (15)$$

该数据点 ${}^s\mathbf{x}_i$ 代表着与子向量 ${}^s\mathbf{y}_{i-1}^*$ 最近已知的低维特征,在低维数据 \mathbf{x}_i^* 与 ${}^s\mathbf{x}_i$ 之间附近的区域进行搜索,可以找到所需的低维数据.获得该数据后,再转换成相应子向量,以此获取与真实要估计的子向量很接近的子向量,最后再组合成最终高维数据.这样避免了直接在高维空间非线性搜索后产生无效的高维数据,提高了搜索效率与准确率.其低维空间的搜索引入了 Boltzmann 分布函数,采用系统采样搜索,可避免所定位低维空间区域的局部最优,对所定位区域全局搜索,并且搜索过程中可对极值进行记忆.

在获取 \mathbf{x}_i^* 与 ${}^s\mathbf{x}_i$ 后,根据真实数据增量降维后的低维数据分布规律(图 3(a)和图 3(b)所示),我们先进行线性搜索,再进行非线性搜索.现定义已知低维数据之间线性搜索个数为 N_{ls} , $F_{ns} = i^* + ns$, $B_{ns} = i^* - ns$, $ns \in \mathbf{N}^+$, 当 $F_{ns} > N_s$, $F_{ns} = F_{ns} - N_s$; 当 $B_{ns} < 1$, $B_{ns} = B_{ns} + N_s$. 前后分别各取两个最近已知低维数据,进行线性搜索后获得的新低维数据有 $\mathbf{s}_l^1 = {}^s\mathbf{x}_{B_{ns}} + l({}^s\mathbf{x}_{F_{ns}} - {}^s\mathbf{x}_{B_{ns}})/(N_{ls} + 1)$, $\mathbf{s}_l^2 = {}^s\mathbf{x}_{B_{ns}} + l({}^s\mathbf{x}_i - {}^s\mathbf{x}_{B_{ns}})/(N_{ls} + 1)$, $\mathbf{s}_l^3 = {}^s\mathbf{x}_{F_{ns}} + l({}^s\mathbf{x}_{F_{ns}} - {}^s\mathbf{x}_i)/(N_{ls} + 1)$, $\mathbf{s}_l^4 = {}^s\mathbf{x}_{F_{ns}} + l({}^s\mathbf{x}_{F_{ns}} - {}^s\mathbf{x}_{F_{ns}})/(N_{ls} + 1)$, $l = 1, \dots, N_{ls}$, 可获得集合 $S_L = \{{}^s\mathbf{x}_{B_{ns}}, \mathbf{s}_l^1, {}^s\mathbf{x}_{B_{ns}}, \mathbf{s}_l^2, {}^s\mathbf{x}_i, \mathbf{s}_l^3, {}^s\mathbf{x}_{F_{ns}}, \mathbf{s}_l^4, {}^s\mathbf{x}_{F_{ns}} \mid l = 1, 2, \dots, N_{ls}\}$, 加入 \mathbf{x}_i^* 后可得扩充集合 $S_{LE} = \{\mathbf{x}_i^*, S_L\}$, 之后选择低维数据点进行非线性搜索,这时有:

$$\begin{aligned} \mathbf{x}_s^* &= \arg \max_{\mathbf{x}_s} F(\mathbf{y}_s) = \arg \max_{\mathbf{x}_s} F(\mathbf{y}_s(S_1), \mathbf{y}_s(S_2)) \\ &= \arg \max_{\mathbf{x}_s} F(f(\mathbf{x}_s), \mathbf{y}_{i-1}^*(S_2)) \\ &= \arg \max_{\mathbf{x}_s} F(\mathbf{x}_s), \mathbf{x}_s \in S_{LE} \end{aligned} \quad (16)$$

式(16)中, $\mathbf{y}_s = [\mathbf{y}_s(S_1)^\top, \mathbf{y}_s(S_2)^\top]^\top \in \mathbf{R}^{D_i}$, $\mathbf{y}_s(S_1) = f(\mathbf{x}_s) \in \mathbf{R}^D$, $\mathbf{y}_s(S_2) = \mathbf{y}_{i-1}^*(S_2) \in \mathbf{R}^{D_i-D}$, $D = D_1, D_2, D_2 = D_T - D_1$, S_1 和 S_2 为维数序号集合, $S_1 \cup S_2 = \{1, 2, \dots, D_T\}$. 找到 \mathbf{x}_s^* 后,设非线性搜索个数为 N_{lns} , 搜索向量为 $\tilde{\mathbf{u}}_{l'} = \tilde{\eta}_{l'} \tilde{L}(\mathbf{u}_{l'} / \|\mathbf{u}_{l'}\|)$, 所有元素服从均匀分布 $U(-1, 1)$ 的矩阵生成函数为 $\text{rand}(n_1, n_2) \in \mathbf{R}^{n_1 \times n_2}$, 令 $\tilde{\eta}_{l'} = \text{rand}(1, 1)$, $\mathbf{u}_{l'} = \text{rand}(q, 1)$, 其搜索步长 \tilde{L} 为:

$$\tilde{L} = \begin{cases} p_L \|\mathbf{x}_s^* - \mathbf{x}_i^*\|, \mathbf{x}_s^* \neq \mathbf{x}_i^* \\ p_L \|\mathbf{x}_s^* - \tilde{\mathbf{x}}^*\|, \mathbf{x}_s^* = \mathbf{x}_i^*, \\ \tilde{\mathbf{x}}^* = \arg \max_{\tilde{\mathbf{x}}} F(f(\tilde{\mathbf{x}}), \mathbf{y}_{i-1}^*(S_2)), \tilde{\mathbf{x}} \in S_L \end{cases} \quad (17)$$

式(17)中, p_L 为调节参数. 这时,第 r 次(当 $r = 1$ 时)非线性搜索获得的 N_{lns} 个低维数据有 ${}^r\tilde{\mathbf{x}}_{l'} = \mathbf{x}_s^* + \tilde{\mathbf{u}}_{l'}$, $l' = 1, 2, \dots, N_{lns}$, 因此获得集合 $S_{NL} = \{{}^r\tilde{\mathbf{x}}_{l'} \mid l' = 1, 2, \dots, N_{lns}\}$, 根据式(16),获得最大目标函数值的低维数据后,可进行记忆如下:

$${}^r\tilde{\mathbf{x}}^* = \arg \max_{\tilde{\mathbf{x}}} F(\tilde{\mathbf{x}}), {}^r\tilde{\mathbf{x}} \in S_{NL} \quad (18)$$

这时引入温度参数 T_A , 引入 Boltzmann 分布函数,进行系统采样搜索,这时有:

$$\begin{aligned} F_A({}^r\tilde{\mathbf{x}}_{l'}) &= \exp(-(-F(\mathbf{y}_s))/T_A) \\ &= \exp(F(\mathbf{y}_s)/T_A) \\ &= \exp(F(\mathbf{y}_s(S_1), \mathbf{y}_s(S_2))/T_A) \\ &= \exp(F(f({}^r\tilde{\mathbf{x}}_{l'}), \mathbf{y}_s(S_2))/T_A) \\ &= \exp\left(-\frac{S}{N_v} \sum_{k=1}^{N_v} \left((1-\alpha) \frac{B_t^k}{B_t^k + H_t^k} + \alpha \frac{R_t^k}{R_t^k + H_t^k}\right) / T_A\right) \end{aligned} \quad (19)$$

$$\bar{F}_A({}^r\tilde{\mathbf{x}}_{l'}) = F_A({}^r\tilde{\mathbf{x}}_{l'}) / \sum_{l'=1}^{N_{lns}} F_A({}^r\tilde{\mathbf{x}}_{l'}) \quad (20)$$

在原有数据集合 S_{NL} 的基础上,根据式(20)归一化目标值的累加值,对集合 S_{NL} 里的低维数据 ${}^r\tilde{\mathbf{x}}_{l'}$ 进行系统选择后作为 $r+1$ 次搜索的起点,并确定相应的搜索次数,对其进行复制保留.在设定温度参数 T_A 后,该次搜索实现系统选择数据如下式所示:

$$\begin{aligned} \tilde{\mathbf{x}}'_{n'} &= {}^r\tilde{\mathbf{x}}_{l'}, \sum_{l'=1}^{l'-1} \bar{F}_A({}^r\tilde{\mathbf{x}}_{l'}) < r_{n'} \leq \sum_{l'=1}^{l'} \bar{F}_A({}^r\tilde{\mathbf{x}}_{l'}), \\ l', n' &= 1, 2, \dots, N_{lns}, r_{n'} \sim U(0, 1) \end{aligned} \quad (21)$$

然后,令 $l' = n'$,可以得到 $r+1$ 次搜索的低维数据 ${}^{r+1}\tilde{\mathbf{x}}_{l'} = \tilde{\mathbf{x}}'_{l'} + \tilde{\mathbf{u}}_{l'}$, $l' = 1, 2, \dots, N_{lns}$, 调节 T_A 再代入式(18)~式(20)计算,再用式(21)对数据进行系统选择,即可进行 $r+2$ 次搜索,其原理示意图如图 4 所示.

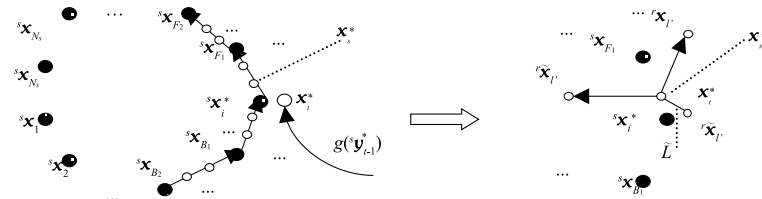
在 N_{lns} 足够大,经过一定次数搜索,即 r 也足够大时,设 $\mathbf{r} = \tilde{\mathbf{r}}$, 可得历史最优低维数据集合 $S_{\text{Best}} = \{\mathbf{x}_s^*, {}^r\tilde{\mathbf{x}}^* \mid r = 1, 2, \dots, \tilde{r}\}$. 由于使用增量降维映射关系 g 获得低维数据后,再使用映射关系 f 生成的子向量还是和原子向量存在一定误差的,所以为了避免搜索后生成的子向量比原来 ${}^s\mathbf{y}_{i-1}^*$ 差,根据式(16),可得最优子向量如下:

$$\mathbf{y}_s^*(S_1) = \arg \max_{\tilde{\mathbf{y}}} F(\tilde{\mathbf{y}}), \tilde{\mathbf{y}} \in \{{}^s\mathbf{y}_{i-1}^*, f(\tilde{\mathbf{x}}^*) \mid \tilde{\mathbf{x}}^* \in S_{\text{Best}}\} \quad (22)$$

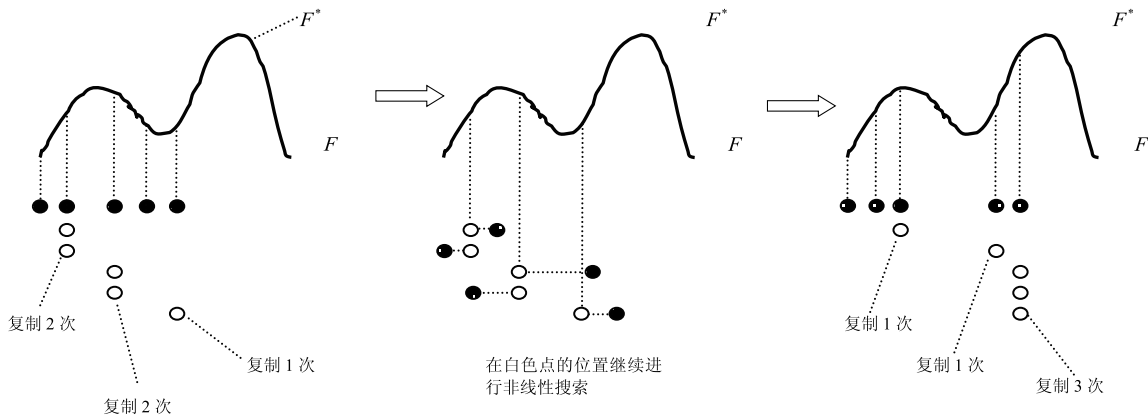
同理,也可获得另一部分最优子向量 $\mathbf{y}_s^*(S_2)$, 从而获得最优最终高维数据 $\mathbf{y}_s^* = [\mathbf{y}_s^*(S_1)^\top, \mathbf{y}_s^*(S_2)^\top]^\top \in \mathbf{R}^{D_i}$.

由于低维空间是对高维空间的概略描述,所以搜索获取的 \mathbf{y}_s^* 需要在其高维空间进行细致微调一下,使之校正后,更能匹配其当前多视角的轮廓图像,因此相应帧的最终高维数据估计值由下式得出:

$$\mathbf{y}_i^* = \arg \max_{\mathbf{y}'_i} F(\mathbf{y}'_i), \mathbf{y}'_i \in S_H \quad (23)$$



(a)线性搜索(以 $\mathbf{x}_i^* = \mathbf{x}_i^*$ 为例)



(b)非线性搜索(以5个低维数据点的搜索,曲线代表函数F的值, F^* 代表函数F最优值为例)

图4 流形Boltzmann优化原理示意图

流形 Boltzmann 优化的搜索过程中,通过温度参数 T_A 的调节来控制搜索的区域. 当 T_A 较大时,本次搜索目标函数值 F 较小的低维数据都有较大概率保留到下一次搜索过程中,搜索到一定程度后,当 T_A 较小时,目标函数值 F 较小的低维数据将会有很大概率被淘汰,这样就在所定位的低维空间区域实现全局到局部的搜索,进而搜索到所定位区域的最优低维数据,再转换成相应子向量,估计人体运动立体模型.

从上述搜索原理可知,若没有式(8)的增量降维映射关系. 光是靠式(6)的低维数据到高维数据映射关系是无法看到图 4(a)和图 4(b)所示的低维数据分布(真实子向量样本在增量降维后的低维数据分布),从而也就无法对低维空间的进行定位后搜索.

3.4 GIDRMBO 算法

我们根据上述章节的讨论,给出简要 GIDRMBO 算法步骤. 在此先定义已知空间位置信息子向量样本为 $\mathbf{Y}_1 = [{}^1\mathbf{y}_1, \dots, {}^1\mathbf{y}_i, \dots, {}^1\mathbf{y}_{N_1}]^T \in \mathbf{R}^{N_1 \times D_1}$ 及姿态信息子向量样本为 $\mathbf{Y}_2 = [{}^2\mathbf{y}_1, \dots, {}^2\mathbf{y}_i, \dots, {}^2\mathbf{y}_{N_2}]^T \in \mathbf{R}^{N_2 \times D_2}$, 用 GIDRM 对其降维后获取的低维数据(低维流形)分别为 $\mathbf{X}_1 = [{}^1\mathbf{x}_1, \dots, {}^1\mathbf{x}_i, \dots, {}^1\mathbf{x}_{N_1}]^T \in \mathbf{R}^{N_1 \times q}$, $\mathbf{X}_2 = [{}^2\mathbf{x}_1, \dots, {}^2\mathbf{x}_i, \dots, {}^2\mathbf{x}_{N_2}]^T \in \mathbf{R}^{N_2 \times q}$, 低维数据到高维数据映射关系分别为 f_1 与 f_2 ($f_s, s = 1, 2$), 以及增量降维映射关系分别为 g_1 与 g_2 ($g_s, s = 1, 2$). 设初始化高维数据为 $\mathbf{y}_0^* \in \mathbf{R}^{D_1}$, 共估计 F_t 帧人体运动立体模型. 当估计第 t 帧的人体运动立体模型时,算法实现步骤如下:

式(23)中,微调样本集合 $S_H = \{\mathbf{y}_s^*, \tilde{\mathbf{y}}_h^* | h = 1, 2, \dots, \bar{h}\}$, $\tilde{\mathbf{y}}_h^* \sim N(\mathbf{y}_s^*, \Sigma)$, Σ 为小幅度值方差矩阵,计算式(23)即为以 \mathbf{y}_s^* 为均值, Σ 为方差获取 \bar{h} 个微调后的样本 $\tilde{\mathbf{y}}_h^*$ 与 \mathbf{y}_s^* 再行比较选择.

输入: \mathbf{y}_{t-1}^* ;
输出: \mathbf{y}_t^* ;
1. 对最终高维数据进行分段后提取子向量, ${}^1\mathbf{y}_{t-1}^* = \mathbf{y}_{t-1}^* (1: D_1)$, ${}^2\mathbf{y}_{t-1}^* = \mathbf{y}_{t-1}^* (D_1 + 1: D_T)$, 令 $s = 1, s' = 2$.
2. 计算 $i^* = \arg \min_i \|g_s({}^s\mathbf{y}_{t-1}^*) - {}^s\mathbf{x}_i\|, i = 1, 2, \dots, N_s$, 找到 \mathbf{x}_{i^*} 后, 用线性搜索获取集合 S_L 与 S_{LE} , 计算 $\mathbf{x}_s^* = \arg \max_{\mathbf{x}_s} F(f_s(\mathbf{x}_s), {}^s\mathbf{y}_{t-1}^*)$, $\mathbf{x}_s \in S_{LE}$ 及搜索步长 \bar{L} . 设置初始温度参数 T_A 及非线性搜索次数 \tilde{r} , 令 $r = 1$, 通过非线性搜索得到 ${}^r\tilde{\mathbf{x}}_r = \mathbf{x}_s^* + \tilde{\mathbf{u}}_r, r' = 1, 2, \dots, N_{ins}$, 获取集合 S_{NL} .
3. 计算 $\tilde{\mathbf{x}}^* = \arg \max_{\tilde{\mathbf{x}}} F(\tilde{\mathbf{x}}), \tilde{\mathbf{x}} \in S_{NL}$, 若 $r > \tilde{r}$, 则执行第 4 步; 否则令 $F_A({}^r\tilde{\mathbf{x}}_{r'}) = \exp(F(f_s({}^r\tilde{\mathbf{x}}_{r'}), {}^s\mathbf{y}_{t-1}^*)/T_A)$, 并计算其归一化目标值 $\bar{F}_A({}^r\tilde{\mathbf{x}}_{r'}) = F_A({}^r\tilde{\mathbf{x}}_{r'}) / \sum_{l'=1}^{N_{ins}} F_A({}^r\tilde{\mathbf{x}}_{l'})$, 计算 $\tilde{\mathbf{x}}'_{r'} = {}^r\tilde{\mathbf{x}}_{r'}, \sum_{r'=1}^{l'-1} \bar{F}_A({}^r\tilde{\mathbf{x}}_{r'}) <$

算法 2 GIDRMBO

输入: \mathbf{y}_{t-1}^* ;
输出: \mathbf{y}_t^* ;
1. 对最终高维数据进行分段后提取子向量, ${}^1\mathbf{y}_{t-1}^* = \mathbf{y}_{t-1}^* (1: D_1)$, ${}^2\mathbf{y}_{t-1}^* = \mathbf{y}_{t-1}^* (D_1 + 1: D_T)$, 令 $s = 1, s' = 2$.
2. 计算 $i^* = \arg \min_i \|g_s({}^s\mathbf{y}_{t-1}^*) - {}^s\mathbf{x}_i\|, i = 1, 2, \dots, N_s$, 找到 \mathbf{x}_{i^*} 后, 用线性搜索获取集合 S_L 与 S_{LE} , 计算 $\mathbf{x}_s^* = \arg \max_{\mathbf{x}_s} F(f_s(\mathbf{x}_s), {}^s\mathbf{y}_{t-1}^*)$, $\mathbf{x}_s \in S_{LE}$ 及搜索步长 \bar{L} . 设置初始温度参数 T_A 及非线性搜索次数 \tilde{r} , 令 $r = 1$, 通过非线性搜索得到 ${}^r\tilde{\mathbf{x}}_r = \mathbf{x}_s^* + \tilde{\mathbf{u}}_r, r' = 1, 2, \dots, N_{ins}$, 获取集合 S_{NL} .
3. 计算 $\tilde{\mathbf{x}}^* = \arg \max_{\tilde{\mathbf{x}}} F(\tilde{\mathbf{x}}), \tilde{\mathbf{x}} \in S_{NL}$, 若 $r > \tilde{r}$, 则执行第 4 步; 否则令 $F_A({}^r\tilde{\mathbf{x}}_{r'}) = \exp(F(f_s({}^r\tilde{\mathbf{x}}_{r'}), {}^s\mathbf{y}_{t-1}^*)/T_A)$, 并计算其归一化目标值 $\bar{F}_A({}^r\tilde{\mathbf{x}}_{r'}) = F_A({}^r\tilde{\mathbf{x}}_{r'}) / \sum_{l'=1}^{N_{ins}} F_A({}^r\tilde{\mathbf{x}}_{l'})$, 计算 $\tilde{\mathbf{x}}'_{r'} = {}^r\tilde{\mathbf{x}}_{r'}, \sum_{r'=1}^{l'-1} \bar{F}_A({}^r\tilde{\mathbf{x}}_{r'}) <$

$r_{n'} \leq \sum_{v=1}^{l'} \bar{F}_A(\tilde{x}_{v'})$, $l', n' = 1, 2, \dots, N_{\text{Ins}}, r_{n'} \sim U(0, 1)$. 令 $l' = n', r = r + 1$, 搜索获取新的数据 $\tilde{x}_{l'} = \tilde{x}_{l'} + \tilde{u}_{l'}, l' = 1, 2, \dots, N_{\text{Ins}}$, 并降低温度参数 T_A , 重复第 3 步.

4. 获得集合 $S_{\text{Best}} = \{ \mathbf{x}_s^*, \tilde{\mathbf{x}}^* \mid r = 1, 2, \dots, \tilde{r} \}$ 后, 计算 $\mathbf{y}_t^* = \arg \max_{\mathbf{y}} F(\tilde{\mathbf{y}})$, $\tilde{\mathbf{y}} \in \{ \mathbf{y}_{t-1}^*, \int_s(\tilde{\mathbf{x}}^*) \mid \tilde{\mathbf{x}}^* \in S_{\text{Best}} \}$, 若 $s=1$, 令 $s=2, s'=1$, 返回第 2 步; 否则执行第 5 步.

5. 在获取两段最优子向量 $\mathbf{y}_t^*, \mathbf{y}_{t+1}^*$ 后, 令 $\mathbf{y}_s^* = [(\mathbf{y}_t^*)^T, (\mathbf{y}_{t+1}^*)^T]^T \in \mathbf{R}^{D_r}$, 获取微调样本集合 $S_{\text{H}} = \{ \mathbf{y}_s^*, \tilde{\mathbf{y}}_h^* \mid h = 1, 2, \dots, \tilde{h} \}$, $\tilde{\mathbf{y}}_h^* \sim N(\mathbf{y}_s^*, \Sigma)$, 最终计算估计值 $\mathbf{y}_t^* = \arg \max_{\mathbf{y}_t'} F(\mathbf{y}_t')$, $\mathbf{y}_t' \in S_{\text{H}}$, 输出第 t 帧的人体运动立体模型. 令 $t = t + 1$, 若 $t \leq F_t$, 返回第 1 步; 否则结束算法.

3.5 GIDRMBO 的算法性能优势分析

GIDRMBO 算法的借助了低维空间数据的先验信息, 可以在低维空间进行高效的搜索, 生成有效的子向量. 常用的粒子滤波方法则都采用直接在高维空间进行搜索. 由于在高维数据中, 各维数据之间有着复杂的函数关系, 其搜索向量又是随机的, 很难生成有效的高维数据进行估计. GIDRMBO 算法采用了由 GPDM 改进的 GIDRM 进行降维, GIDRM 具有 GPDM 的性质. 因此, 在低维空间中, 已知低维数据之间的区域为高置信区域, 即该区域为生成相应高维数据的有效区域, 该区域大小覆盖了已知相邻的低维数据及其附近区域, 控制搜索的空间区域, 再随机生成搜索向量, 就可以在低维空间搜索过程中搜索到有效的低维数据, 随之生成相应有效的子向量, 组成有效最终高维数据, 所提出 GIDRMBO 算法就利用了这个优势进行搜索. 然而, 在高维空间搜索, 其搜索向量要求就很高, 随机生成的高维搜索向量很难满足各维数据的有效范围, 进行搜索后, 所得的最终高维数据很有可能是无效的, 即转换成人体运动立体模型后, 其空间位置与肢体凌乱, 常用的一些算法几乎用的搜索方法都等价于这个方法. 因此, 这也说明了图 4(c) 和图 4(d) 所得实验结果的原因, 同时表明 GIDRMBO 算法比常用粒子滤波算法估计的效果要好, 其估计的人体运动立体模型与各视角原始运动图像更匹配.

4 仿真实验与评价

由于一些粒子滤波算法经常用于该三维人体运动形态估计, 所以本文采用退火粒子滤波 (Annealed Particle Filter, APF^[4]), 高斯粒子滤波 (Gaussian Particle Filter, GPF^[18,19]) 和标准粒子滤波 (Particle Filter, PF^[5,20-22]) 三种算法与所提算法 (GIDRMBO) 进行比较. 各种算法的性能测试主要有估计误差, 消除轮廓数据歧义估计, 及克服短时遮挡估计三个方面, 并且采用

三个视角的运动图像序列进行测试.

首先, 我们用连续的 50 帧多视角运动图像序列 (走路 1~3, 跑步 1, 2) 分别测试各算法的估计误差, 并给出其中几种运动形态的每帧估计误差, 实验测试结果如图 5 所示. 从图 5(a)~图 5(d) 中我们发现, 本文所提算法 (GIDRMBO) 的平均误差, 误差标准差, 最大估计误差 (图 5(b)~图 5(d) 中的算法标识所示) 及每帧估计误差 (计算方法如文献[10]) 相对于其他三种算法 (APF, GPF, PF) 总体较低. 在估计运动形态的前 20 帧时, 各算法每帧估计误差可能较接近, 但随着帧数的增加, 所提算法 (GIDRMBO) 的每帧估计误差总体低于其他三种算法 (APF, GPF, PF). 由于算法概率的随机性, 偶尔会有一些帧的误差很接近, 这属正常现象.

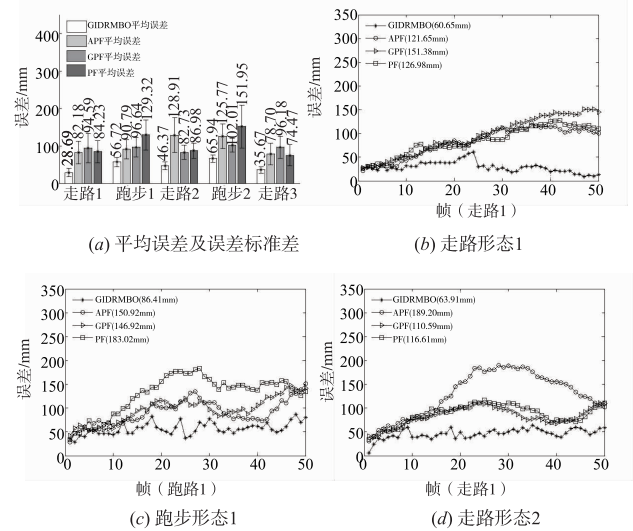


图5 GIDRMBO、APF、GPF及PF算法估计误差比较

其次, 我们测试一下各算法消除轮廓数据歧义估计的性能. 该数据歧义是指一组有限视角的轮廓图像 (如图 1 右第二行的图像所示) 无法准确描述人体运动立体模型各肢体的具体位置, 例如, 在该组轮廓中, 到底是左脚在前还是右脚在前难以确定. 所提出算法以低维数据为先验信息可进行区分, 从而消除轮廓数据歧义进行较准确的估计. 我们下面采用人体走路运动形态进行测试, 测试结果如图 6(a)~图 6(e) 所示, 所提算法 (GIDRMBO) 可以消除轮廓数据歧义进行估计, 所估计的人体运动立体模型与各视角原始运动图像里的人体肢体匹配, 其他的三种算法 (APF, GPF, PF) 则无法克服, 与真实数据相比, 其估计人体运动立体模型的左脚 (灰色) 和右脚 (黑色) 与各视角原始运动图像的人体左脚 (白色) 和右脚 (灰色) 都出现了左右相反的现象.

最后, 我们测试一下各算法克服短时遮挡估计的性能. 这里的克服短时遮挡估计是指各视角的镜头被

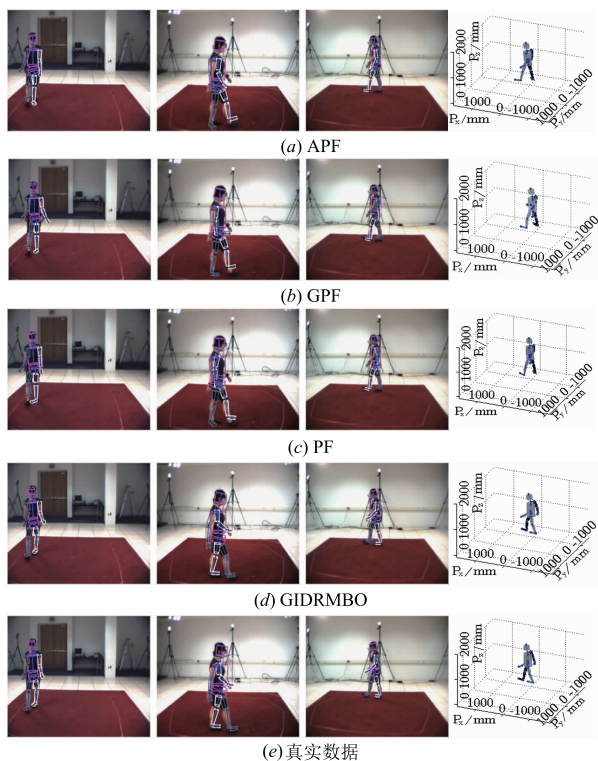


图6 各算法消除轮廓数据歧义估计比较

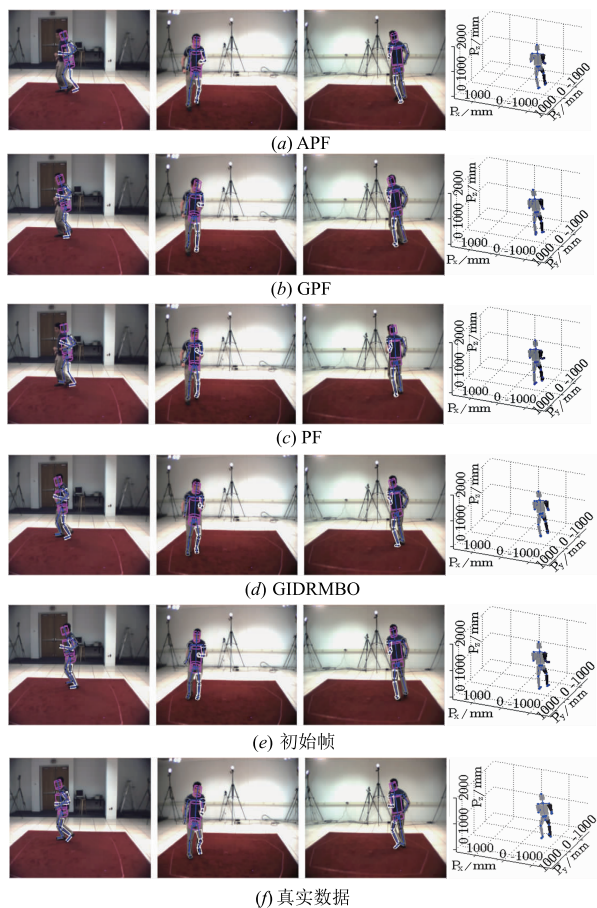


图7 各算法克服短时遮挡估计比较

暂时遮挡后,无法获取其相应被遮挡帧的轮廓图像,因此要估计下一时刻的运动姿态与本时刻估计的运动姿态相差较大,这相当于估计间隔若干帧(例如 10 ~ 15 帧)后各视角轮廓图像相应的人体运动立体模型. 我们采用人体跑步运动形态进行测试,初始帧与被估计帧相间隔 15 帧,各算法测试结果如图 7(a) ~ 图 7(f) 所示,所提算法(GIDRMBO)借助了所获取的低维空间进行搜索后,可以直接估计出初始帧被遮挡以后第 15 帧的人体跑步运动姿态,其估计的人体运动立体模型与各视角原始运动图像匹配较好,而其他三种算法(APF, GPF, PF)与所提算法(GIDRMBO)相比,其估计人体运动立体模型的肢体与各视角原始运动图像的人体肢体匹配较差,大多数肢体的位置明显错位.

5 结束语

本文提出的高斯增量降维与流形 Boltzmann 优化(GIDRMBO)算法可以较好的从多视角轮廓图像估计出含有空间位置信息的三维人体运动形态. 所提算法与一些常用粒子滤波算法(APF, GPF, PF)相比,其借助了子向量样本降维后的低维数据先验信息,具有较低估计误差,并且可以消除轮廓数据歧义和克服短时遮挡进行估计. 对于空间位置信息变化规则,运动姿态为周期运动的三维人体运动形态估计,所提算法可以较完美实现,其基本能满足常用三维人体运动形态的估计. 然而,所提算法在对于空间位置信息和姿态信息变化不规则的三维人体运动形态估计还有所欠缺,初始化数据还需要用真实数据进行初始化,另外,所提算法对于多视角人体运动图像的轮廓提取质量要求较高,否则会影响实验结果. 在今后研究中,我们将会对所提算法继续改进,使其能满足空间位置信息及姿态信息变化不规则的三维人体运动形态估计,且无需真实数据初始化. 那时,无标记人体运动立体模型^[1,3]将会用于描述三维人体运动形态.

参考文献

- [1] Martin Sandau, Henrik Koblauch, Thomas B. Moeslund, et al. Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane[J]. *Medical Engineering & Physics*, 2014, 36(9): 1168 - 1175.
- [2] Xin Zhang, Guoliang Fan, Li-shan Chou. Two-layer dual gait generative models for human motion estimation from a single camera[J]. *Image and Vision Computing*, 2013, 31(6-7): 473 - 486.
- [3] E Yeguas-Bolivar, R Muñoz-Salinas, R Medina-Carnicer, et al. Comparing evolutionary algorithms and particle filters for markerless human motion capture [J]. *Applied Soft Computing*, 2014, 17: 153 - 166.

- [4] Leonid Sigal, Alexandru O Balan, Michael J Black. Humaneva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion [J]. *International Journal of Computer Vision*, 2010, 87 (1-2): 4-27.
- [5] Ming Du, Xiaoming Nan, Ling Guan. Monocular human motion tracking by using de-mc particle filter [J]. *IEEE Transactions on Image Processing*, 2013, 22 (10): 3852-3865.
- [6] 李毅, 孙正兴, 等. 基于退火粒子群优化的单目视频人体姿态分析方法 [J]. *自动化学报*, 2012, 38(5): 732-741.
Li Yi, Sun Zheng-xing, et al. 3D human pose analysis from monocular video by simulated annealed particle swarm optimization [J]. *Acta Automatica Sinica*, 2012, 38(5): 732-741. (in Chinese)
- [7] C H Ek. *Shared Gaussian Process Latent Variables Models* [D]. Oxford, U K: Oxford Brookes University, 2009.
- [8] Jack M Wang, David J Fleet, Aaron Hertzmann. Gaussian process dynamical models for human motion [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(3): 283-298.
- [9] S Chen, C F N Cowan, P M Grant. Orthogonal least squares learning algorithm for radial basis function networks [J]. *IEEE Transactions on Neural Networks*, 1991, 2(2): 302-309.
- [10] L Sigal, M J Black. Humaneva: Synchronized Video and Motion Capture Dataset for Evaluation of Articulated Human Motion [R]. Providence, USA: Brown University, 2006. 1-18.
- [11] Renbin Peng, Pramod K Varshney. A human visual system-driven image segmentation algorithm [J]. *Journal of Visual Communication and Image Representation*, 2015, 26: 66-79.
- [12] Zuoyong Li, Kezong Tang, Yong Cheng, et al. Transition region-based single-object image segmentation [J]. *International Journal of Electronics and Communications*, 2014, 68(12): 1214-1223.
- [13] Xiangli Liao, Hongbo Xu, Yicong Zhou, et al. Automatic image segmentation using salient key point extraction and star shape prior [J]. *Signal Processing*, 2014, 105: 122-136.
- [14] N Andrei. Scaled conjugate gradient algorithms for unconstrained optimization [J]. *Computational Optimization and Applications*, 2007, 38(3): 401-416.
- [15] I T Nabney, *Algorithms for Pattern Recognition* [M]. Berlin, Germany: Springer, 2001. 645-647.
- [16] J B Tenenbaum, V de Silva, J C Langford. A global geometric framework for nonlinear dimensionality reduction [J]. *Science*, 2000, 290: 2319-2323.
- [17] S T Roweis, L K Saul. Nonlinear dimensionality reduction by locally linear embedding [J]. *Science*, 2000, 290: 2323-2326.
- [18] Miodrag Bolić, Akshay Athalye, Sangjin Hong, et al. Study of algorithmic and architectural characteristics of Gaussian particle filters [J]. *Journal of Signal Processing Systems*, 2010, 61(2): 205-218.
- [19] Dilmurat Tursun, Askar Hamdulla. Gaussian particle filter based algorithm for tracking of a dim moving point target in IR image sequences [A]. *Proceedings of 2012 International Conference on Multimedia and Signal Processing* [C]. Heidelberg: Springer Verlag, 2012. 215-223.
- [20] Zhengjiang Zhang, Junhui Chen. Simultaneous data reconciliation and gross error detection for dynamic systems using particle filter and measurement test [J]. *Computers & Chemical Engineering*, 2014, 69: 66-74.
- [21] Yingya Su, Qingjie Zhao, Liu Jun Zhao, et al. Abrupt motion tracking using a visual saliency embedded particle filter [J]. *Pattern Recognition*, 2014, 47(5): 1826-1834.
- [22] I-Cheng Chang, Shih-Yao Lin. 3D human motion tracking based on a progressive particle filter [J]. *Pattern Recognition*, 2010, 43(10): 3621-3635.

作者简介



李万益 (通信作者) 男, 1985 年生于广西南宁. 现为广东第二师范学院计算机科学系讲师, 博士, 主要研究方向为机器学习, 图像处理, 通信系统优化.

E-mail: luther1212@163.com



孙季丰 男, 1962 年生于广东广州, 现为华南理工大学电子与信息学院教授, 博士生导师. 主要研究方向为数字视频处理, 生物信息处理和 DNA 数据建模和压缩, 人体形态分析及建模, 自组织无线网络等.

E-mail: ecjfsun@scut.edu.cn