

基于多核稀疏编码的三维人体姿态估计

余家林, 孙季丰, 李万益

(华南理工大学电子与信息学院, 广东广州 510641)

摘要: 为了准确有效的重构多视角图像中的三维人体姿态, 该文提出一种基于多核稀疏编码的人体姿态估计算法. 首先, 针对连续帧姿态估计的歧义问题, 该文设计了一种用于表达多视角图像的 HA-SIFT 描述子, 其中, 人体局部拓扑、肢体相对位置及外观信息被同时编码; 然后, 在多核学习框架下建立同时考虑特征空间内在流形结构与姿态空间几何信息的目标函数, 并在希尔伯特空间优化目标函数以更新稀疏编码、过完备字典与多核权重; 最后, 利用姿态字典原子的线性组合来估计对应未知输入的三维人体姿态. 实验结果表明, 与核稀疏编码、Laplace 稀疏编码及 Bayesian 稀疏编码相比, 文本方法具有更高的估计精度.

关键词: 人体姿态估计; 多视角图像; 多核学习; 稀疏编码; 字典学习

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2016)08-1899-10

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.08.019

3D Human Pose Estimation Based on Multi-kernel Sparse Coding

YU Jia-lin, SUN Ji-feng, LI Wan-yi

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China)

Abstract: In order to reconstruct 3D human pose from multi-view images accurately and effectively, a novel human pose estimation algorithm based on multi-kernel sparse coding is proposed. First, for the ambiguity of human pose estimation between the consecutive frames, we describe multi-view images using a special HA-SIFT descriptor, in which the human body local topology, relative coordinates and appearance information are encoded simultaneously; then, an objective function is established within the framework of multi-kernel learning, it takes both intrinsic manifold structure of the feature space and geometrical information of the pose space into consideration. The sparse coding, over-complete dictionary and multi-kernel weight are updated by optimizing the objective function iteratively in the Hilbert space; finally, the corresponding 3D human pose of the unknown input image is estimated by a linear combination of the bases of the human pose dictionary. The experimental results show that our proposed method provides higher accuracy of human pose estimation compared with kernel sparse coding, Laplace sparse coding and Bayesian sparse coding.

Key words: human pose estimation; multi-view images; multi-kernel learning; sparse coding; dictionary learning

1 引言

从多视角序列估计含三维空间位置信息的人体姿态在计算机视觉领域有广泛应用, 譬如手势识别^[1]、行为识别^[2]、运动捕捉^[3]和人机交互^[4]等. 该研究虽已取得重大突破, 但仍有许多具有挑战性和亟待解决的难题. 首先, 复杂的三维人体运动与二维图像之间存在语义鸿沟, 深度信息的缺乏导致估计的人体姿态存在歧义; 其次, 人体外观与轮廓在帧间差异较大, 给肢体的定位带来困难; 最后, 肢体的遮挡、姿态数据的高维及场景

光线的变化等, 都使人体姿态估计成为一项艰巨的任务. 常见方法分3类: 基于模型的方法, 其依赖一种基于先验知识的人体模型, 通过优化目标函数来估计人体姿态, 但计算量较大^[5]; 基于学习的方法, 其直接学习从特征空间到姿态空间的映射^[6], 但它对姿态的判别是以庞大的训练数据为前提; 基于样本的方法, 其在训练数据库中检索与输入最相似的数据, 并利用该数据插值来获取结果, 该方法同样依赖庞大的训练数据, 且要求训练数据要覆盖姿态空间尽可能多的自由度^[7].

最近, 在机器学习和模式识别领域中兴起了一种

被称为稀疏编码^[8]的算法,譬如人脸识别^[8]、目标分类^[9]和人体姿态估计^[10]等. Behnam 等^[11]在 Bayesian 框架下提出一种稀疏编码算法 (Bayesian Sparse Coding, BSC),通过学习两种过完备字典克服了小样本问题所导致的过拟合. Gao 等^[12]提出一种 Laplace 稀疏编码算法 (Laplace Sparse Coding, LSC),解决了数据位置和相似度信息丢失问题,但原始空间的特征往往带噪声,在该空间构造的正则项,未必能精确反映数据的内在流形. 2013 年, Gao 等^[13]提出一种能捕捉特征非线性相似度的核稀疏编码算法 (Kernel Sparse Coding, KSC),打破了仅在原始空间编码的模式. 但对不同类型的输入,采用单一核函数来处理并不合理,且面临核函数及参数选择难题.

人体姿态估计中的样本是高维非线性的,尽管样本的近邻点可在线性距离上被找到,但由它们所构造的近邻图并不能精确反映数据的内在流形,而该流形在许多应用中又非常重要. 核技巧把原始数据隐式映

射至希尔伯特空间克服了该问题,但又面临核函数及参数的选择难题,交叉验证方法虽可解决该难题,但计算量过大. 综合考虑,本文提出一种多核稀疏编码算法 (Multi-Kernel Sparse Coding, MKSC),通过引入多核学习,既解决了姿态数据的“维数灾难”问题^[14],又可应对样本的非线性. 其中,最优核由核函数集中核函数的线性组合导出,从而不存在核函数及参数的选择难题.

2 基于多核稀疏编码的人体姿态估计

本文算法框架如图 1 所示,它从三方面提升了姿态估计精度. 首先,将含肢体局部拓扑、位置及外观信息的 SIFT 算子作用于图像兴趣点,并借助词袋模型^[15]描述图像,打破了以往仅编码单目轮廓的模式;其次,用多核函数将数据隐式映射至希尔伯特空间,使得构造的近邻图能够精确反映数据内在流形;最后,在稀疏编码过程中考虑姿态特征几何信息,使肢体的局部拓扑在运动过程中被有效保持.

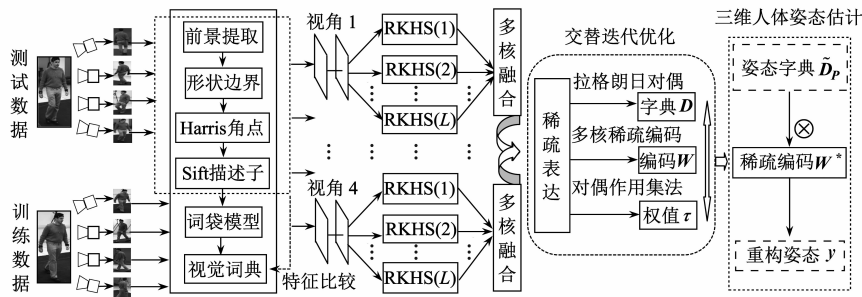


图1 基于多核稀疏编码的人体姿态估计算法框架

2.1 多视角特征融合

人体姿态估计精度高度依赖于图像的表达,人体轮廓特征虽为姿态估计提供强有力的形状线索,且对人体外观及光线具有不变性,但轮廓特征丢失了外观信息,而该信息对连续帧姿态之间微小的变化又十分敏感,这在连续帧姿态估计中容易造成歧义. 为解决该问题,本文通过如下四步设计一种特殊的描述子:

(1)前景提取:通过背景差分获取多视角序列的人体形状边界窗 body.

(2)提取 Harris 兴趣角点:在人体形状边界窗 body 内提取 Harris 兴趣点.

(3)SIFT 算子:将肢体局部拓扑、位置及外观信息编码至 SIFT 算子,并作用于兴趣点得 p .

(4)HA-SIFT 描述子:找到兴趣点的相对位置 (u, v) ,生成 HA-SIFT 描述子: $s = (u, v, p)^T$;

如图 2 所示,本文借助词袋模型来描述输入图像. 首先,找到所有训练样本的 HA-SIFT 特征描述子: $S = (s_1, \dots, s_d)$,并对 S 进行 K -Means 聚类,令基数 K 等于 n ,获取视觉词典: $C = (c_1, \dots, c_n)$,其中, $\{c_i\}_{i=1}^n$ 表示 C

的视觉单词,即聚类中心,从而把输入图像用一个 n 维特征矢量来描述. 在姿态空间中,将人体用一种运动学树^[16]参数化,得到三维人体姿态特征矢量: $y = [t, \vartheta_1, \theta_1, \dots, \theta_l]$,其中, t 表示树根, ϑ_1 表示在世界坐标系中 t 的方向,相对关节角度集 $\{\theta_i\}_{i=1}^l$ 表示各肢体方向.

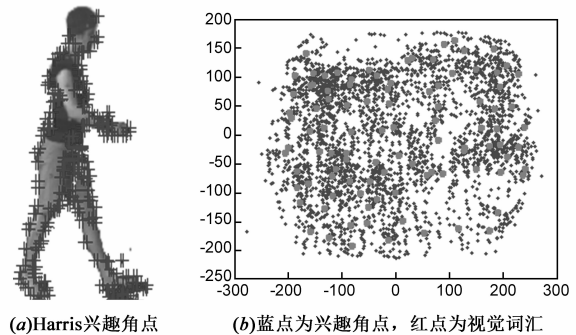


图2 多视角特征融合

2.2 多核稀疏编码

2.2.1 建立目标函数

假设有 m 个训练样本对, $\xi = [(x_1, y_1), \dots, (x_m,$

$y_m)$], 其中, $[\mathbf{x}_1, \dots, \mathbf{x}_m] = \mathbf{X}$, 表示输入特征矢量集, 且 $\mathbf{x}_i = [x_{i1}, \dots, x_{in}]^T \in \mathbb{R}^n$; $[\mathbf{y}_1, \dots, \mathbf{y}_m] = \mathbf{Y}$, 表示与 \mathbf{X} 对应的 m 个姿态特征矢量集, 且 $\mathbf{y}_i = [y_{i1}, \dots, y_{ir}]^T \in \mathbb{R}^r$. 在多核学习^[14]框架下, 利用核函数将原始输入空间隐式映射至高维再生希尔伯特空间 (Reproducing Kernel Hilbert Space, RKHS) $H^{(l)}$, 该空间拥有非线性映射函数 $\phi: \mathbf{X} \rightarrow H$ 的内积 $\langle \cdot, \cdot \rangle$, 即: $\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle = K(\mathbf{x}, \mathbf{y})$. 本文假设选用 L 个核函数 $\{K^{(l)}\}_{l=1}^L$, 从而对应 L 个希尔伯特空间 $\{H^{(l)}\}_{l=1}^L$, 但通常并不明确选用哪个空间, 最合理的做法是级联这 L 个空间, 从而形成增广再生希尔伯特空间: $\tilde{H} = \bigoplus_{l=1}^L H^{(l)}$, 并通过迭代优化空间 $H^{(l)}$ 的权值 τ_l ($\sum_{l=1}^L \tau_l = 1, \tau_l \geq 0, \forall l$), 得最适合输入的空间 \tilde{H}^* . 假设映射函数 $\phi: \xi_i \rightarrow \phi_l(\xi_i), l=1, \dots, L, \phi_l$ 为 $H^{(l)}$ 的映射函数, 级联这 L 个函数: $\phi^T(\xi_i) = [\tau_1 \phi_1(\xi_i), \dots, \tau_L \phi_L(\xi_i)]^T$, 其中, $\{\tau_l\}_{l=1}^L$ 等同于 $H^{(l)}$ 的权值, 确保级联后的映射函数 $\phi^T(\xi_i)$ 也达到最优, 且 $\sum_{l=1}^L \tau_l = 1, \tau_l \geq 0, \forall l$, 则样本对 (ξ_i, ξ_j) 的核函数可表示为如下式子:

$$\begin{aligned} K^T(\xi_i, \xi_j) &= \phi^T(\xi_i)^T \cdot \phi^T(\xi_j) = \sum_{l=1}^L \tau_l^2 \phi_l(\xi_i)^T \cdot \phi_l(\xi_j) \\ &= \sum_{l=1}^L \tau_l^2 K_l(\xi_i, \xi_j) \end{aligned} \quad (1)$$

当原始空间的特征被映射至希尔伯特空间后, 将样本对 $(\mathbf{x}_i, \mathbf{y}_i)$ 用 $(\phi^T(\mathbf{x}_i), \phi^T(\mathbf{y}_i))$ 代替, 输入特征字典 $\mathbf{D}_F \in \mathbb{R}^{n \times k}$ 和人体姿态字典 $\mathbf{D}_P \in \mathbb{R}^{r \times k}$ 分别用 $\phi^T(\mathbf{D}_F)$ 和 $\phi^T(\mathbf{D}_P)$ 代替. 但核函数 $K^T(\mathbf{D}_P, \mathbf{y}_i), K^T(\mathbf{D}_P, \mathbf{D}_P), K^T(\mathbf{D}_F, \mathbf{x}_i)$ 和 $K^T(\mathbf{D}_F, \mathbf{D}_F)$ 的计算量过大. 为解决该问题, 假设输入特征和人体姿态特征的基元跨越在 $\phi^T(\mathbf{X}) = [\phi^T(\mathbf{x}_1), \dots, \phi^T(\mathbf{x}_m)]$ 和 $\phi^T(\mathbf{Y}) = [\phi^T(\mathbf{y}_1), \dots, \phi^T(\mathbf{y}_m)]$ 的列空间中, 输入特征字典和姿态字典分别表示为: $\mathbf{D}_F = \phi^T(\mathbf{X}) \cdot \mathbf{A}$ 和 $\mathbf{D}_P = \phi^T(\mathbf{Y}) \cdot \mathbf{B}$, 其中, $\mathbf{A} \in \mathbb{R}^{m \times k}$ 和 $\mathbf{B} \in \mathbb{R}^{m \times k}$ 为跨越参数矩阵. 目标函数如下所示:

$$\min_{\mathbf{A}, \mathbf{B}, \mathbf{W}, \tau} \left(\|\phi(\mathbf{X}) - \phi^T(\mathbf{X}) \cdot \mathbf{A}\mathbf{W}\|_F^2 + \lambda_2 \|\phi(\mathbf{H}) - \phi^T(\mathbf{Y}) \cdot \mathbf{B}\mathbf{W}\|_F^2 + \phi(\mathbf{W}) \right) \quad (2)$$

式(2)中, 前边两项分别表示输入特征矢量与人体姿态特征矢量的重构误差; 第三项表示为: $\phi(\mathbf{W}) = \lambda_1 \phi_1(\mathbf{W}) + \lambda_3 \phi_2(\mathbf{W})$, $\phi_1(\mathbf{W}) = \|\mathbf{W}\|_1$ 用于控制稀疏矩阵 \mathbf{W} 的稀疏度; $\phi_2(\mathbf{W}) = \sum_{i,j=1}^m \|\mathbf{w}_i - \mathbf{w}_j\|_F^2 S_{ij}^\tau$ 表示 τ -权值希尔伯特空间的流形正则项^[12], 当 $\phi^T(\mathbf{x}_j)$ 为 $\phi^T(\mathbf{x}_i)$ 的最近邻时, 满足: $S_{ij}^\tau = K^T(\mathbf{x}_i, \mathbf{x}_j)$, 否则, $S_{ij}^\tau = 0$. 另外, λ_1, λ_2 和 λ_3 为正则化参数, \mathbf{H} 表示与对应输入间有最小测地距离 (Geodesic Distance) 的姿态特征矩阵. 为了避免权值

τ 出现负值或过度收缩, 要求: $\sum_{l=1}^L \tau_l = 1, \tau_l \geq 0$. 将 $\phi(\mathbf{W})$ 代入式(2)得到如下式子:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{B}, \mathbf{W}, \tau} & \left(\|\phi(\mathbf{X}) - \phi^T(\mathbf{X}) \cdot \mathbf{A}\mathbf{W}\|_F^2 + \lambda_1 \|\mathbf{W}\|_1 + \right. \\ & \lambda_2 \|\phi(\mathbf{H}) - \phi^T(\mathbf{Y}) \cdot \mathbf{B}\mathbf{W}\|_F^2 + \\ & \left. \lambda_3 \sum_{i,j=1}^m \|\mathbf{w}_i - \mathbf{w}_j\|_F^2 S_{ij}^\tau \right) \end{aligned} \quad (3)$$

2.2.2 人体姿态空间转换

姿态字典 $\mathbf{D}_P \in \mathbb{R}^{r \times k}$ 中含关节角度数据 $\{\theta_i\}_{i=1}^k \in [-180^\circ, 180^\circ]$, 使重构误差不能用 l_2 范数度量. 图3显示了两种人体姿态, 它们手部空间位置很接近, 但在角度空间中差异却很大. 因此, 本文采用一种映射将姿态矢量 $\phi_i \in \mathbf{D}_P$ 转换至一个 \mathbb{R}^{2r} 的单位球面空间, 记为:

$$\tilde{\phi}_i = \frac{1}{\sqrt{r}} [\cos(\phi_i), \sin(\phi_i)], i=1, \dots, k \quad (4)$$

式(4)中, $\cos(\phi_i) = [\cos(\phi_{i1}), \dots, \cos(\phi_{ir})]$ 和 $\sin(\phi_i) = [\sin(\phi_{i1}), \dots, \sin(\phi_{ir})]$, 该映射使输入特征和姿态特征在优化过程中位于同一空间, 使目标函数中关于姿态特征的正则项有意义; 并且可用 l_2 范数来度量输入特征和姿态特征的重构误差; 估计出的人体姿态矢量 $\hat{\mathbf{y}}$ 可用 $\tilde{\mathbf{D}}_P = [\tilde{\phi}_1, \dots, \tilde{\phi}_k]$ 中原子线性组合来表示. 此时, 人体姿态字典就表示为: $\tilde{\mathbf{D}}_P = \phi^T(\tilde{\mathbf{Y}}) \cdot \mathbf{B}$.

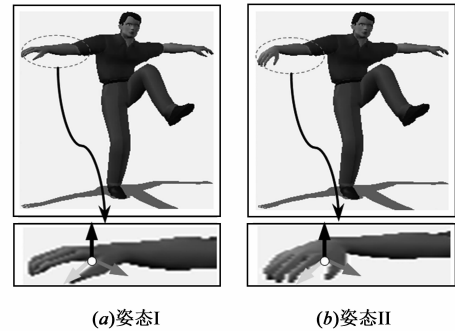


图3 姿态I和姿态II的角度空间对比

2.3 优化目标函数

2.3.1 更新稀疏编码 \mathbf{W}

为了独立更新稀疏编码矢量 \mathbf{w}_i , 将目标函数改写为矢量形式, 记为:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{B}, \{\mathbf{w}_i\}, \tau} & \sum_{i=1}^m \|\phi(\mathbf{x}_i) - \phi^T(\mathbf{X}) \cdot \mathbf{A}\mathbf{w}_i\|_F^2 + \lambda_1 \sum_{i=1}^m \|\mathbf{w}_i\|_1 \\ & + \lambda_2 \sum_{i=1}^m \|\phi(\mathbf{h}_i) - \phi^T(\tilde{\mathbf{Y}}) \cdot \mathbf{B}\mathbf{w}_i\|_F^2 \\ & + \lambda_3 \sum_{i,j=1}^m \|\mathbf{w}_i - \mathbf{w}_j\|_F^2 S_{ij}^\tau \end{aligned} \quad (5)$$

除 \mathbf{w}_i 以外的所有矢量 $\{\mathbf{w}_j\}_{j \neq i}$ 都是固定的, 可将式(5)进一步改写为:

$$\min_{\{\mathbf{w}_i\}} \|\phi(\mathbf{x}_i) - \phi^T(\mathbf{X}) \cdot \mathbf{A}\mathbf{w}_i\|_F^2$$

$$\begin{aligned}
& + \lambda_1 \sum_{j=1}^k |w_{ij}| + \lambda_2 \|\phi(\mathbf{h}_i) - \phi^\tau(\hat{\mathbf{Y}}) \cdot \mathbf{B}\mathbf{w}_i\|_F^2 \\
& + \lambda_3 \sum_{i,j=1}^m \|\mathbf{w}_i - \mathbf{w}_j\|_F^2 S_{ij}^\tau \quad (6)
\end{aligned}$$

式(6)中, $|w_{ij}|$ 为稀疏矢量 \mathbf{w}_i 中第 j 个系数的绝对值。

当更新 \mathbf{w}_i 时, 假设目标函数 $f_{obj}(\mathbf{w}_i)$ 取得最优解时 w_{ij} 符号是已知的, 这样就可计算出 $|w_{ij}|$ 的值, 从而避免求解 l_1 范数, 即是当 $w_{ij} > 0$ 时, $|w_{ij}| = w_{ij}$; 当 $w_{ij} < 0$ 时, $|w_{ij}| = -w_{ij}$; 当 $w_{ij} = 0$ 时, $|w_{ij}| = 0$ 。从而将式(6)转化为一个无约束的 (Quadratic Programming, QP) 问题^[17], 本文设计算法 1 来实现该类问题的优化。

算法 1 多核稀疏编码 (Multi-kernel Sparse Coding)

输入 ① 一个包含 U 个测试样本的数据集 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_U]$, 近邻字典 \mathbf{D}_Z (该字典的第 i 列表示训练样本中与第 i 个测试样本具有最小测地距离的人体姿态特征矢量), 正则参数 λ_1 与 λ_2 , 输入特征字典 \mathbf{D}_F 与对应的三维人体姿态字典 $\hat{\mathbf{D}}_p$ (即已知跨越参数矩阵 \mathbf{A} 和 \mathbf{B}); ② 多核权值 τ ; ③ 特征矢量对的核函数: K_{XX}, K_{YY} 和 K_{YH} 。

输出 测试样本集 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_U]$ 对应的稀疏编码矩阵 $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_U]$ 。

开始 针对每个输入样本 $\mathbf{x}_i \in \mathbf{X}$ 和姿态特征矢量 $\mathbf{h}_i \in \mathbf{D}_Z$ 执行下列五个步骤:

1. **初始化**: 已知 $\mathbf{w}_i = \mathbf{0}$, 符号集 $\mathbf{F} = \mathbf{0}$ 及作用集 $\Pi = \emptyset, \mathbf{F}^{(j)} \in \{-1, 0, 1\}$ 为 $\mathbf{w}_i^{(j)}$ (即 \mathbf{w}_i 中第 j 个系数) 的符号。
2. **激活步**: 从 \mathbf{w}_i 中的零系数开始, 选择系数索引 j , 通过下式计算:

$$\begin{aligned}
& \operatorname{argmax}_j |\nabla_i^{(j)} f_{obj}(\mathbf{w}_i)| = \operatorname{argmax}_j \\
& \left| \frac{\partial (\|\phi(\mathbf{x}_i) - \phi^\tau(\mathbf{X}) \cdot \mathbf{A}\mathbf{w}_i\|_F^2 + \lambda_2 \|\phi(\mathbf{h}_i) - \phi^\tau(\hat{\mathbf{Y}}) \cdot \mathbf{B}\mathbf{w}_i\|_F^2 + \lambda_3 \sum_{i,j=1}^m \|\mathbf{w}_i - \mathbf{w}_j\|_F^2 S_{ij}^\tau)}{\partial w_{ij}} \right| \quad (7)
\end{aligned}$$

当 $\mathbf{w}_i^{(j)}$ 局部优化目标函数时, 激活系数 $\mathbf{w}_i^{(j)}$, 即有:

- ① **负标志** 当 $\nabla_i^{(j)} f_{obj}(\mathbf{w}_i) > \lambda_1$ 时, 设置 $\mathbf{F}^{(j)} = -1, \Pi = \{j\} \cup \Pi$;
- ② **正标志** 当 $\nabla_i^{(j)} f_{obj}(\mathbf{w}_i) < -\lambda_1$ 时, 设置 $\mathbf{F}^{(j)} = 1, \Pi = \{j\} \cup \Pi$ 。

3. 特征符号搜索 (Feature-sign Search):

① 假设 $\hat{\mathbf{D}}_F = \phi^\tau(\hat{\mathbf{X}}) \cdot \hat{\mathbf{A}}$ 和 $\hat{\mathbf{D}}_p = \phi^\tau(\hat{\mathbf{Y}}) \cdot \hat{\mathbf{B}}$ 分别表示 \mathbf{D}_F 和 $\hat{\mathbf{D}}_p$ 对应作用集 Π 的子矩阵;

② 假设 $\hat{\mathbf{w}}_i$ 和 $\hat{\mathbf{F}}$ 分别表示 \mathbf{w}_i 和 \mathbf{F} 对应作用集 Π 的子矢量;

③ 通过解决一个无约束的 QP 问题来获取解析解:

$$\begin{aligned}
& \min_{\hat{\mathbf{w}}_i} \|\phi(\mathbf{x}_i) - \phi^\tau(\hat{\mathbf{X}}) \cdot \hat{\mathbf{A}}\hat{\mathbf{w}}_i\|_F \\
& + \lambda_1 \hat{\mathbf{F}}^\top \hat{\mathbf{w}}_i + \lambda_2 \|\phi(\mathbf{h}_i) - \phi^\tau(\hat{\mathbf{Y}}) \cdot \hat{\mathbf{B}}\hat{\mathbf{w}}_i\|_F \\
& + \lambda_3 \left(\sum_{j \neq i} \|\hat{\mathbf{w}}_i - \hat{\mathbf{w}}_j\|_F^2 S_{ij}^\tau + \sum_{j \neq i} \|\hat{\mathbf{w}}_j - \hat{\mathbf{w}}_i\|_F^2 S_{ji}^\tau + \|\hat{\mathbf{w}}_i - \hat{\mathbf{w}}_i\|_F^2 S_{ii}^\tau \right) \quad (8)
\end{aligned}$$

将式(8)中展开式的函数内积用核函数代替并对 $\hat{\mathbf{w}}_i$ 求导, 令 $\partial q(\hat{\mathbf{w}}_i) / \partial \hat{\mathbf{w}}_i = 0$, 得到在作用集 Π 下的最优解:

$$\begin{aligned}
& \hat{\mathbf{w}}_i^{\text{new}} = (\hat{\mathbf{A}}^\top K^\tau(\hat{\mathbf{X}}, \hat{\mathbf{X}}) \hat{\mathbf{A}} + \hat{\mathbf{B}}^\top K^\tau(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}) \hat{\mathbf{B}} \\
& + \lambda_3 \mathbf{I} \sum_{j \neq i} (S_{ij}^\tau + S_{ji}^\tau))^{-1} (\hat{\mathbf{A}}^\top K^\tau(\hat{\mathbf{X}}, \mathbf{x}_i)
\end{aligned}$$

$$\begin{aligned}
& - \lambda_1 \hat{\mathbf{F}}^\top / 2 + \lambda_2 \hat{\mathbf{B}}^\top K^\tau(\hat{\mathbf{Y}}, \mathbf{h}_i) \\
& + \lambda_3 \sum_{j \neq i} \hat{\mathbf{w}}_j (S_{ij}^\tau + S_{ji}^\tau) \quad (9)
\end{aligned}$$

式(9)中, \mathbf{I} 表示一个单位矩阵;

④ 在 $\hat{\mathbf{w}}_i$ 到 $\hat{\mathbf{w}}_i^{\text{new}}$ 之间的一条封闭线段上执行离散线性搜索: 找到目标函数在 $\hat{\mathbf{w}}_i^{\text{new}}$ 和其它所有使系数 $\mathbf{w}_i^{(j)}$ 符号反转的点处的值, 并将 $\hat{\mathbf{w}}_i$ 更新到使目标函数取得最小值的点处;

⑤ 移除矢量 $\hat{\mathbf{w}}_i$ 中的零系数并更新它的符号 $\mathbf{F} = \operatorname{sign}(\mathbf{w}_i)$ 。

4. 判别优化条件:

I. 对 \mathbf{w}_i 中非零系数, $\nabla_i^{(j)} f_{obj}(\mathbf{w}_i) + \lambda_1 \operatorname{sign}(\mathbf{w}_i^{(j)}) = 0, \forall \mathbf{w}_i^{(j)} \neq 0$; 如果不满足, 回步骤 3, 否则检查 II;

II. 对 \mathbf{w}_i 中零系数, $\nabla_i^{(j)} f_{obj}(\mathbf{w}_i) \leq \lambda_1, \forall \mathbf{w}_i^{(j)} = 0$; 如果 II 不满足, 回步骤 2, 否则得最优解 \mathbf{w}_i^* 。

5. **返回最优解 \mathbf{W}^*** : 当条件 I 与 II 都满足时, 返回最优稀疏编码矩阵 $\mathbf{W}^* = [\mathbf{w}_1^*, \dots, \mathbf{w}_U^*]$ 。

2.3.2 更新跨越参数矩阵 \mathbf{A} 和 \mathbf{B}

为了更新跨越参数矩阵 \mathbf{A} , 现固定稀疏编码 \mathbf{W} , 跨越参数矩阵 \mathbf{B} 与多核权值 τ , 移除目标函数中的无关项, 得如下式子:

$$\begin{aligned}
& \min_{\mathbf{A}} \|\phi(\mathbf{X}) - \phi^\tau(\mathbf{X}) \cdot \mathbf{A}\mathbf{W}\|_F^2, \quad (10) \\
& \text{s. t. } \|\mathbf{a}_i\|^2 \leq c, i = 1, \dots, k
\end{aligned}$$

式(10)是一个带二次约束的最小二乘问题, 可采用一般的基于迭代投影的梯度下降法求解。为提高算法性能, 本文提出一种更加高效的基于拉格朗日对偶方法。

假设 $\boldsymbol{\delta} = [\delta_1, \dots, \delta_k]$, δ_i 表示不等式 $\|\mathbf{a}_i\|^2 \leq c$ 的拉格朗日乘子, 式(10)的拉格朗日对偶函数表示为:

$$\begin{aligned}
& g(\boldsymbol{\delta}) = \inf_{\mathbf{D}_F} L(\mathbf{D}_F, \boldsymbol{\delta}) \\
& = \inf_{\mathbf{A}} \left(\|\phi(\mathbf{X}) - \phi^\tau(\mathbf{X}) \cdot \mathbf{A}\mathbf{W}\|_F^2 + \sum_{i=1}^k \delta_i (\|\mathbf{a}_i\|^2 - c) \right) \quad (11)
\end{aligned}$$

假设 \mathbf{A} 为一个 $k \times k$ 的对角阵, 且 $\mathbf{A}_{ii} = \delta_i$, 这样 $L(\mathbf{D}_F, \boldsymbol{\delta})$ 可表示为:

$$\begin{aligned}
& L(\mathbf{D}_F, \boldsymbol{\delta}) = \|\phi(\mathbf{X}) - \phi^\tau(\mathbf{X}) \cdot \mathbf{A}\mathbf{W}\|_F^2 \\
& + \operatorname{Tr}(\mathbf{A}^\top \phi(\mathbf{X})^\top \phi(\mathbf{X}) \mathbf{A} \mathbf{A}) - c \operatorname{Tr}(\mathbf{A}) \quad (12)
\end{aligned}$$

对于式(12)而言, 令 $\partial L(\mathbf{D}_F, \boldsymbol{\delta}) / \partial \mathbf{A} = 0$ 可得出最优解 \mathbf{A}' , 即:

$$\begin{aligned}
& \partial L(\mathbf{D}_F, \boldsymbol{\delta}) / \partial \mathbf{A} = -2\mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W}^\top + 2\mathbf{W}^\top \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W} \mathbf{A} \\
& + 2\mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{A} \mathbf{A} \\
& = 0 \quad (13)
\end{aligned}$$

$$\mathbf{A}' = \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W}^\top \cdot (\mathbf{W}^\top \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W} + \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{A})^{-1} \quad (14)$$

将式(14)代入式(12), 拉格朗日对偶函数变为:

$$\begin{aligned}
& g(\boldsymbol{\delta}) = \operatorname{Tr}(\mathbf{K}^\tau(\mathbf{X}, \mathbf{X})) - \operatorname{Tr}(\mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W}^\top (\mathbf{W} \mathbf{W}^\top \\
& + \mathbf{A})^{-1} \mathbf{W}) - c \operatorname{Tr}(\mathbf{A}) \quad (15)
\end{aligned}$$

通过解决如下一个拉格朗日对偶问题计算最优解 \mathbf{A}^* :

$$\min_{\mathbf{A}} \operatorname{Tr}(\mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \cdot \mathbf{W}^\top (\mathbf{W} \mathbf{W}^\top + \mathbf{A})^{-1} \mathbf{W}) + c \operatorname{Tr}(\mathbf{A}),$$

$$\text{s. t. } \delta_i \geq 0, i = 1, \dots, k \quad (16)$$

对于式(16)而言,可以直接采用牛顿法或梯度下降法计算最优解 \mathbf{A}^* ,代入式(14)得到最优的跨越参数矩阵 $\mathbf{A}^* = \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W}^\tau \cdot (\mathbf{W}^\tau \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{W} + \mathbf{K}^\tau(\mathbf{X}, \mathbf{X}) \mathbf{A}^*)^{-1}$.

为了更新跨越参数矩阵 \mathbf{B} ,通过固定稀疏编码 \mathbf{W} 、跨越参数矩阵 \mathbf{A} 和多核权值 $\boldsymbol{\tau}$,同理可得最优的跨越参数矩阵 $\mathbf{B}^* = \lambda_2 \mathbf{K}^\tau(\bar{\mathbf{Y}}, \mathbf{H}) \mathbf{W}^\tau (\mathbf{W}^\tau \mathbf{K}^\tau(\bar{\mathbf{Y}}, \bar{\mathbf{Y}}) \mathbf{W} + \mathbf{K}^\tau(\bar{\mathbf{Y}}, \bar{\mathbf{Y}}) \mathbf{A}^*)^{-1}$.

2.3.3 更新多核权值 $\boldsymbol{\tau}$

学习权值 $\boldsymbol{\tau}$ 是为从核函数集中构造最优核来处理不同类型的输入,以提高系统的泛化性能. 现固定稀疏编码 \mathbf{W} 、跨越参数矩阵 \mathbf{A} 和 \mathbf{B} , 移除目标函数的无关项,得如下式子:

$$\min_{\boldsymbol{\tau}} \sum_{i=1}^m \|\phi(\mathbf{x}_i) - \phi^\tau(\mathbf{X}) \cdot \mathbf{A} \mathbf{w}_i\|_F^2 + \lambda_2 \sum_{i=1}^m \|\phi(\mathbf{h}_i) - \phi^\tau(\bar{\mathbf{Y}}) \cdot \mathbf{B} \mathbf{w}_i\|_F^2 \quad (17)$$

式(17)可被视为一个关于权值 $\boldsymbol{\tau}$ 的带约束 QP 问题,即:

$$\sum_{i=1}^L \tau_i^2 e_f + \lambda_2 \sum_{i=1}^L \tau_i^2 e_p, \quad \text{s. t. } \sum_{i=1}^L \tau_i = 1, \tau_i \geq 0 \quad (18)$$

式(18)中, e_f 和 e_p 可分别表示为:

$$\left. \begin{aligned} e_f &= \text{Tr}(K_l(\mathbf{X}, \mathbf{X})) - 2\text{Tr}(K_l(\mathbf{X}, \mathbf{X}) \mathbf{A} \mathbf{W}) \\ &\quad + \text{Tr}(\mathbf{W}^\tau \mathbf{A}^\tau K_l(\mathbf{X}, \mathbf{X}) \mathbf{A} \mathbf{W}) \\ e_p &= \text{Tr}(K_l(\mathbf{H}, \mathbf{H})) - 2\text{Tr}(K_l(\mathbf{H}, \bar{\mathbf{Y}}) \mathbf{B} \mathbf{W}) \\ &\quad + \text{Tr}(\mathbf{W}^\tau \mathbf{B}^\tau K_l(\bar{\mathbf{Y}}, \bar{\mathbf{Y}}) \mathbf{B} \mathbf{W}) \end{aligned} \right\} \quad (19)$$

将式(19)代入式(18),求解一个标准的 QP 问题来更新权值 $\boldsymbol{\tau}$. 为提高算法效率,本文采用一种对偶作用集算法^[18]在可行域内搜索解析解来实现该类问题的优化.

2.4 三维人体姿态估计

当输入一幅测试图像时,首先,借助 HA-SIFT 将其转换为一个 n 维特征矢量;然后,利用特征字典 \mathbf{D}_F 中原子的线性组合来表示该矢量,得最优稀疏编码 \mathbf{W}^* ;最后,联立姿态字典 $\bar{\mathbf{D}}_P$ 与最优稀疏编码来估计对应测试输入的三维人体姿态. 具体见算法 2.

算法 2 三维人体姿态估计(3D Human Pose Estimation)

输入 输入特征矢量集 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_m]$, 其中, $\mathbf{x}_i \in \mathbb{R}^n$; 人体姿态特征矢量集 $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_m]$, 其中, $\mathbf{y}_i \in \mathbb{R}^r$. 一个包含 U 个测试样本的数据集 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_U]$, $\mathbf{x}_i \in \mathbb{R}^n$;

输出 对应测试输入的三维人体姿态特征矢量 $\hat{\mathbf{y}}$;

开始 设计如下四步以估计对应测试输入的三维人体姿态:

① 利用输入特征矢量集 \mathbf{X} 与人体姿态特征矢量集 \mathbf{Y} , 由拉格朗日对偶得到对应的输入特征字典 \mathbf{D}_F 与人体姿态字典 $\bar{\mathbf{D}}_P$;

其中, \mathbf{D}_F 经过映射: $\tilde{\boldsymbol{\varphi}}_i = \frac{1}{\sqrt{r}} [\cos(\boldsymbol{\varphi}_i), \sin(\boldsymbol{\varphi}_i)]$ 后得到 $\bar{\mathbf{D}}_P$;

② 借助最近邻图构造方法, 建立一个近邻字典 \mathbf{D}_Z ;

③ 当给定 \mathbf{D}_F , $\bar{\mathbf{D}}_P$ 及 \mathbf{D}_Z 时, 利用算法 1 学习出最优稀疏编码

\mathbf{W}^* , 即: $\mathbf{x} = \mathbf{W}^* \cdot \mathbf{D}_F$;

④ 联立学习出的最优稀疏编码 \mathbf{W} 与人体姿态字典 $\bar{\mathbf{D}}_P$, 即可推断对应测试输入的三维人体姿态 $\hat{\mathbf{y}} = \mathbf{W}^* \cdot \bar{\mathbf{D}}_P$;

返回 估计出的三维人体姿态特征矢量 $\hat{\mathbf{y}}$.

3 实验分析

3.1 多视角人体姿态估计

为了验证算法的有效性, 本文在布朗大学的 Brown 与 HumanEva 数据库^[19]上实验. 其中, Brown 库中含 1184 帧 Lee 环绕行走的 C1 ~ C4 四视角灰度图像序列; HumanEva 库中含 (S1, S2, S3) 的人体环绕行走姿态共 2950 帧、慢跑运动 2345 帧、拳击运动 2486 帧及其手势动作 2850 帧. 由 C1 ~ C4 灰度摄像机和 C1 ~ C3 彩色摄像机采集而来. 两数据库中图像尺寸分别为: 644 × 484 和 656 × 490.

实验中取 HumanEva-I 的 C1 ~ C3 三个视角、Brown 与 HumanEva-II 的 C1 ~ C4 四个视角. 并分别选取 10 与 15 肢体人体模型, 包括: 头部、躯干、左右上下手臂、左右大小腿, 后者增加了骨盆、左右手及左右脚. 本文实验环境为: CPU: SU3500, 2G 内存, 用 Matlab 2012a 仿真实验. 本文分别从两方面验证算法有效性, 一比较多核学习框架下与一般非监督学习框架下稀疏编码对人体姿态估计的影响; 二比较 HA-SIFT、SIFT 及轮廓特征对算法性能的影响.

图 4 为从 C1 ~ C4 四视角观察 Brown 库中环绕行走姿态估计结果. 并将本文方法 (MKSC) 与 Bayesian 稀疏编码^[11] (BSC)、Laplace 稀疏编码^[12] (LSC) 及核稀疏编码^[13] (KSC) 进行比较.

图 5 为从 HumanEva 中挑选出的部分实验结果: S1 (Gesture)、S2 (Walking) 及 S3 (Jogging). 在原始的 HumanEva-I 数据库中, 是采用 10 肢体的人体模型, 本文将其扩展至 15 肢体.

针对图 4 的实验结果, 对包含本文算法在内的四种算法进行对比分析:

(1) MKSC 算法利用多核将输入特征隐式映射至希尔伯特空间, 在该空间构造的近邻图能精确反映数据内在流形, 并解决了样本的“维数灾难”及非线性问题. 该算法采用多核融合导出的最优核, 能适应各种类型的输入, 该方法估计出的人体姿态效果最佳;

(2) KSC 算法针对多种类型的输入仅采用单一核函数来处理, 存在核函数及参数选择难题. 人们往往是依据输入特征分布来确定最适宜的核函数. 但当难以确定输入特征分布时, 核函数的选取存在很大的随机性, 导致 KSC 算法的泛化性能较差, 估计误差也较大;

(3) LSC 算法是直接于原始空间构造近邻图, 没有考虑输入空间的噪声及样本的非线性, 使构造的近邻

图没能精确反映数据的内在流形. 另外, LSC 的解偏向于一个常数, Laplace 矩阵的嵌入并不能很好保存肢体的局部拓扑, 外推能力较弱, 姿态估计误差较大.

(4) BSC 算法未考虑人体姿态数据的局部几何信息, 而这些信息又是人体姿态估计中不可或缺的, 这往

往会导致肢体之间的局部拓扑在人体运动中不能被有效保持. 实验结果显示, 该算法估计出的人体右臂出现严重畸变, 且各肢体间的连接不协调甚至不成人形.

由此得出结论, 在多视角人体姿态估计问题上, 本文算法比现有最新算法更具优势.

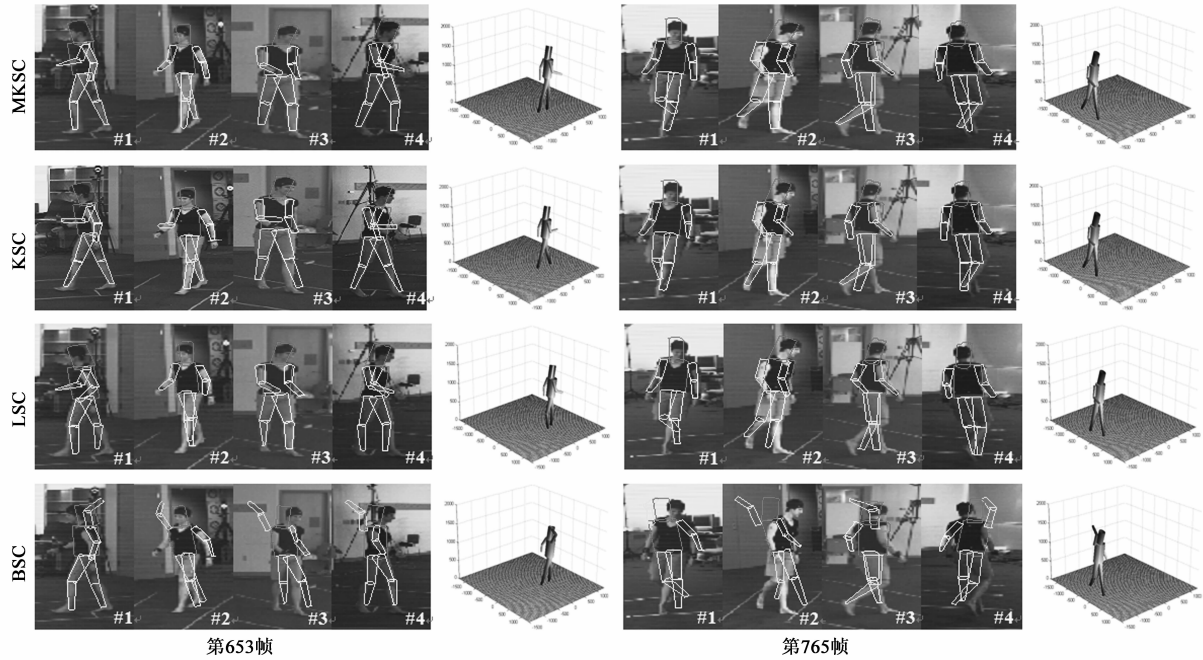


图4 Brown 数据库中环绕行走人体姿态估计结果

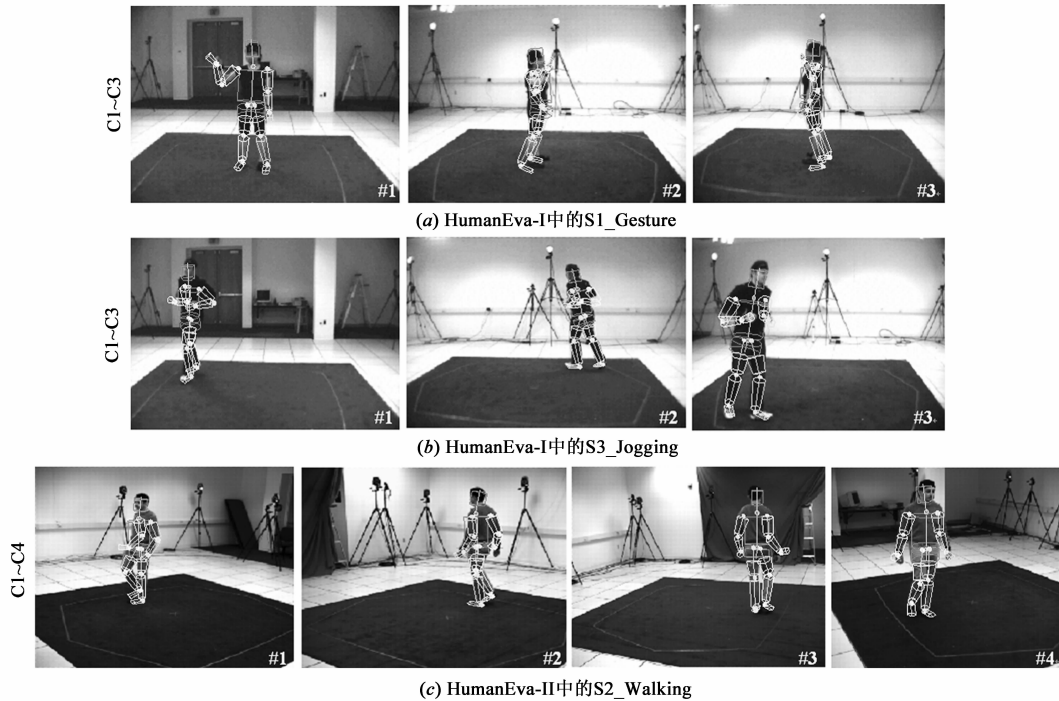


图5 HumanEva-I与HumanEva-II中部分运动姿态估计结果

图 6 为 MKSC 分别基于三类特征的邻帧姿态估计结果,本文采用 15 肢体人体模型在 HumanEva- II 的 S2 (Walking) 上实验. 实验结果表明,HA-SIFT 克服了因自

遮挡所导致的歧义(正常:左肢黄右肢青,歧义:左肢青右肢黄),且估计精度也明显提高.

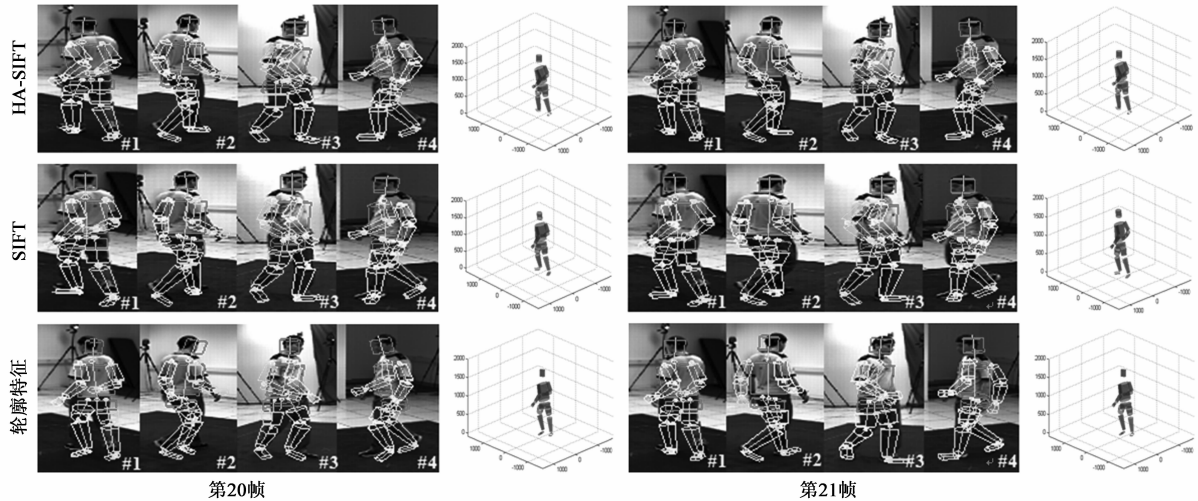


图6 基于三类特征的邻帧姿态估计结果

3.2 人体姿态估计误差

表 1 列出了本文算法与其它三种算法分别基于 HA-SIFT, SIFT 及轮廓特征的关节角度的均方根误差^[5] (RMS), 在行走 (Walking)、拳击 (Boxing)、慢跑 (Jogging) 及手势 (Gesture) 四类数据集上实验. 估计姿态 y 与真实数据 y' 的 RMS 误差 $D(y, y')$ 如下所示:

$$D(y, y') = \frac{1}{m} \sum_{i=1}^m |(y_i - y'_i) \text{ modulo } \pm 180^\circ| \quad (20)$$

图 7 显示了在四种算法作用下, 分别基于三类特

征在 HumanEva- I 的 S3 (Jogging) 数据集上实验所得到的对约定角度变量 $\text{Angle} \in [1, \dots, 20]$ 的关节角度相对误差.

实验结果表明, 在约定角度 $\text{Angle} \in [1, \dots, 20]$ 内, 本文算法对多种类型的输入泛化性能强, 并且本文的 HA-SIFT 特征在人体姿态估计问题上效果最佳, 相对误差在 3 种特征中是最小的. 图 8 为行走过程中本文算法估计的膝盖弯曲角度与标记数据之间的偏差.

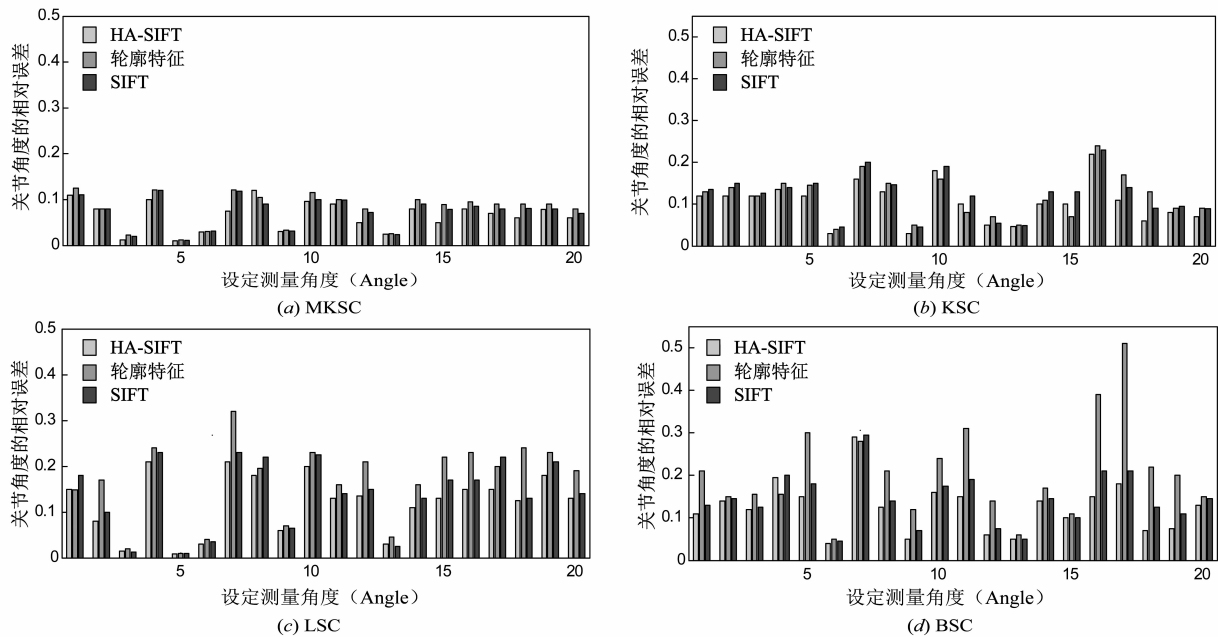


图7 四种算法下S3(Jogging)的关节角度相对误差

表 1 行走、拳击、慢跑及手势四类数据集上的关节角度 RMS 误差(单位:度)

	HA-SIFT 特征				SIFT 特征				轮廓特征			
	行走	拳击	慢跑	手势	行走	拳击	慢跑	手势	行走	拳击	慢跑	手势
MKSC	3.8452	3.0301	2.3531	2.7630	4.2102	3.6536	2.7013	2.8183	4.7193	4.8537	3.6217	3.2021
KSC	5.6627	4.3612	4.1921	3.9636	6.0863	5.6572	5.7143	5.8911	6.2131	6.1646	5.8661	6.1129
LSC	7.1944	6.2561	5.2971	5.7450	7.3640	6.3312	5.6239	6.8535	7.4445	6.9784	6.3255	6.3614
BSC	8.3709	7.2145	7.2915	6.7142	8.8135	8.2785	6.8618	7.2715	9.1467	8.4661	7.5497	7.4610

为了评价估计姿态的空间位置误差,参照文献[19]的方法,用肢体中各关节误差的均值作为每帧的估计误差.图9显示了前300帧中人体头部、骨盆、上下

手臂及大小腿的估计误差.图10表示在单视角与四视角下本文算法与其它三种算法的行走姿态估计误差.

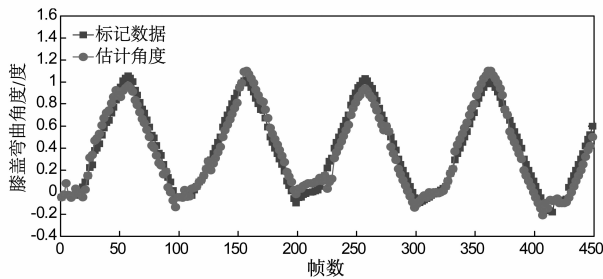


图8 行走中估计的膝盖弯曲角度与标记数据之间的比较

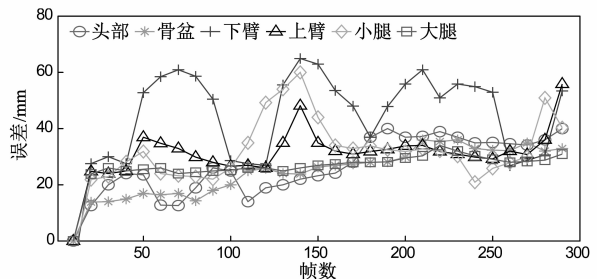
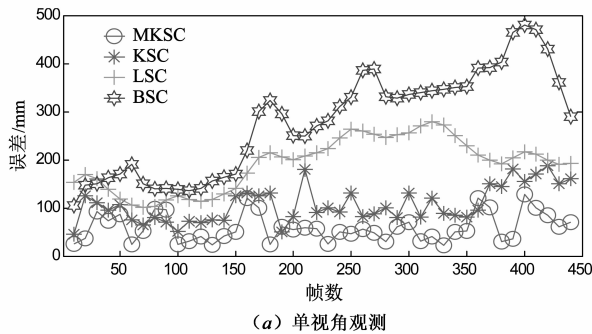
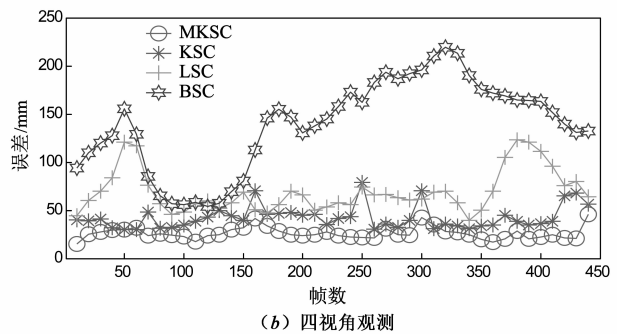


图9 行走中人体部分肢体的估计误差



(a) 单视角观测



(b) 四视角观测

图10 单视角与四视角行走姿态估计误差比较

为了测试本文算法的抗噪能力.在实验中分别设定6种椒盐噪声密度:0%,10%,20%,30%,40%,50%.图11为在6种噪声强度干扰下的外观图像.图12为前300帧在6种噪声密度下行走姿态估计误差.

随噪声密度由0%增大至40%,估计误差逐渐增大,但从整体观察,仍处于较低水平.当噪声密度增大至50%时,估计误差异常偏大,表明算法已失效,本文

算法在一定噪声强度范围内具有较强的鲁棒性.

3.3 算法的复杂性分析

本节主要讨论多核稀疏编码算法的计算复杂度与姿态估计精度之间的关系.假设共有 m 个训练样本对, $\xi = [(x_1, y_1), \dots, (x_m, y_m)]$, 其中, $x_i \in \mathbb{R}^n, y_i \in \mathbb{R}^l$. 在权值特征空间找到 k 个最近邻的计算复杂度为: $O(m^2)$, 其中, k 等于 m ; 固定矩阵 A, B 及权值 τ , 更新 W

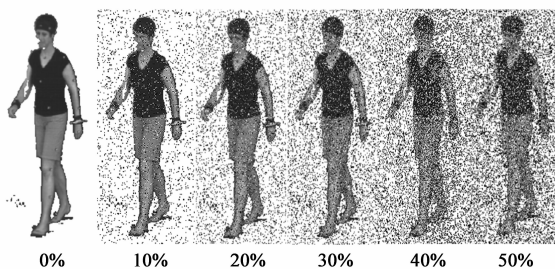


图11 在6种椒盐噪声密度下的人体外观图像

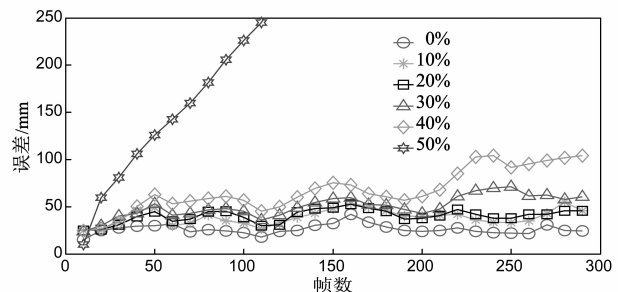


图12 在6种噪声密度下的行走姿态估计误差

的计算复杂度为: $O(nm^2J)$, J 为很小的正整数, 表示 W 中非零元数目, 则可近似为: $O(nm^2)$; 固定矩阵 A, B 及 W , 更新权值 τ 的计算复杂度为: $O(k^3Lm)$, 其中, L 表示核函数的个数; 固定 W 及权值 τ , 更新矩阵 A 与 B 的计算复杂度为: $O(n^3)$ 与 $O(r^3)$. 那么, 知整个算法的计

算复杂度为: $O(m^2 + t(nm^2 + Lm^4 + n^3 + r^3)) \propto m, n$. 其中, t 为迭代次数, r 为姿态特征的维度. 图 13 为在不同给定特征维度 n 下, 手势姿态估计精度随训练样本数 m 的变化趋势.

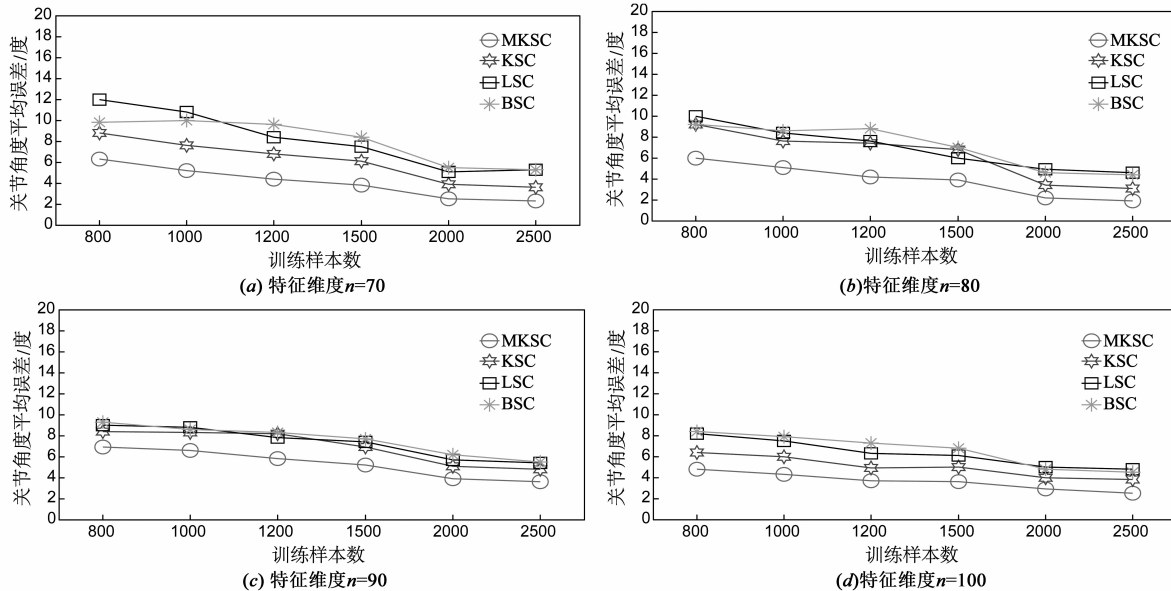


图 13 不同特征维度下手势姿态估计精度随样本数变化趋势

4 结论

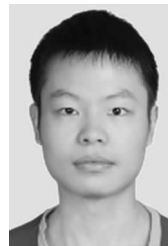
为了准确有效的重构多视角图像中的三维人体姿态, 本文提出一种基于多核稀疏编码的三维人体姿态估计算法. 该算法结合 Harris 兴趣角点与 SIFT 算子来生成本文的 HA-SIFT 描述子, 并借助词袋模型来描述输入图像, 打破了以往仅编码单目轮廓特征的模式; 由于数据可能为非线性分布, 直接在原始空间构造的流形正则项并不可靠, 本文利用核函数将原始数据隐式映射至希尔伯特空间, 使该空间的流形正则项能精确反映数据的内在流形结构. 与目前最新算法相比, 本文算法具有更高的估计精度. 但由于目前仍采用刚性的人体模型, 在估计复杂姿态时效果并不理想. 在接下来的研究中, 将试图采用点云配准方法代替基于能量的方法, 用点云数据建模人体, 有望进一步提高估计精度.

参考文献

- [1] Zhu Ren, Yuan Junsong, Meng Jingjing, et al. Robust part-based hand gesture recognition using kinect sensor [J]. IEEE Transaction on Multimedia, 2013, 15 (5): 1110 - 1120.
- [2] Lin Weiyao, Chen Yuanzhe, Wu Jianxin, et al. A new network-based algorithm for human activity recognition in videos [J]. IEEE Transaction on Circuit and System for Video Technology, 2014, 24 (5): 826 - 841.
- [3] S Vantigodi, W B Radhakrishnan. Action recognition from motion capture data using meta-cognitive RBF network classifier [A]. Proceedings of IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) [C]. Singapore: IEEE Computer Society, 2014. 1 - 6.
- [4] G Palmas, M Bachynskyi, A Oulasvirta, et al. MovExp: A versatile visualization tool for human-computer interaction studies with 3D performance and biomechanical data [J]. IEEE Transaction on Visualization and Computer Graphics, 2014, 20 (12): 2359 - 2368.
- [5] A Agarwal, B Triggs. Recovering 3D human pose from monocular images [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2006, 28 (1): 44 - 58.
- [6] J Shotton, R Girshick, A Fitzgibbon, et al. Efficient human pose estimation from single depth images [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2013, 35 (12): 2821 - 2839.
- [7] N C Tang, Chiou-Ting Hsu, Weng Mingfang, et al. Example-based human motion extrapolation and motion repairing using contour manifold [J]. IEEE Transaction on Multimedia, 2014, 16 (1): 47 - 59.
- [8] Wang Jing, Lu Canyi, Wang Meng, et al. Robust face rec-

- ognition via adaptive sparse representation[J]. IEEE Transaction on Cybernetics, 2014, 44(12): 2368 – 2378.
- [9] 王瑞, 杜林峰, 孙督等. 复杂场景下结合 SIFT 与核稀疏表示的交通目标分类识别[J]. 电子学报, 2014, 42(11): 2129 – 2134.
WANG Rui, DU Lin-feng, SUN Du, et al. Traffic object recognition in complex scenes based on SIFT and kernel sparse representation[J]. Acta Electronica Sinica, 2014, 42(11): 2129 – 2134. (in Chinese)
- [10] Zhou Liuyang, Lu Zhiwu, Howard Leung, et al. Spatial temporal pyramid matching using temporal sparse representation for human motion retrieval[A]. Proceedings of International Conference on Computer Graphics International (CGI) [C]. Sydney, Australia; Springer, Berlin, 2014. 845 – 854.
- [11] B M Behnam, A Jourabloo, A Zarghami, et al. A Bayesian framework for sparse representation-based 3D human pose estimation[J]. IEEE Signal Processing Letters, 2014, 21(3): 297 – 300.
- [12] Gao Shenghua, I W Tsang, Chia Liang-Tien. Laplace sparse coding, Hypergraph laplacian sparse coding, and application[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2013, 35(1): 92 – 101.
- [13] Gao Shenghua, I W Tsang, Chia Liang-Tien. Sparse representation with kernels[J]. IEEE Transaction on Image Processing, 2013, 22(2): 423 – 434.
- [14] Hong Zeng, Yiu-ming Cheung. Feature selection and kernel learning for local learning-based clustering[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1532 – 1546.
- [15] 赵宏伟, 李清亮, 刘萍萍. 基于分级显著信息的空间编码方法[J]. 电子学报, 2014, 42(9): 1863 – 1867.
- ZHAO Hong-wei, LI Qing-liang, LIU Ping-ping. Spatial encoding based on hierarchical salient information [J]. Acta Electronica Sinica, 2014, 42(9): 1863 – 1867. (in Chinese)
- [16] Edgar, Simo-Serra, A Perez-Gracia. Kinematic synthesis using tree topologies[J]. Mechanism and Machine Theory, 2014, 72(2): 94 – 113.
- [17] Y Xia, S Changyin. A novel neural dynamical approach to convex quadratic program and its efficient application [J]. Neural Network, 2009, 10(22): 1463 – 1470.
- [18] Fan Qibin, Jiao Yuling, Lu Xiliang. A primal dual active algorithm with continuation for compressed sensing[J]. IEEE Transaction on Signal Processing, 2014, 62(23): 6274 – 6284.
- [19] L Sigal, and M J Black. Humaneva: synchronized video and motion capture dataset for evaluation of articulated human motion[R]. Report of Brown University, Providence: 2006.

作者简介



余家林 (通信作者) 男, 1989 年生于贵州镇远. 现为华南理工大学电信学院信息与通信工程专业博士研究生. 研究方向为计算机视觉、人体运动形态分析、图像与视频处理等.
E-mail: yu. jialin@mail. scut. edu. cn



孙季丰 男, 1962 年生于广东揭阳, 现为华南理工大学电信学院教授, 博士生导师. 研究方向包括智能信号处理、图像与视频处理、自组织通信网等.