

基于 $L_{1/2}$ 正则化的三维人体姿态重构

洪金华¹ 张荣¹ 郭立君¹

摘要 针对从给定 2D 特征点的单目图像中重构对象的 3D 形状问题, 本文在形状空间模型的基础上, 结合 $L_{1/2}$ 正则化和谱范数的性质提出一种基于 $L_{1/2}$ 正则化的凸松弛方法, 将形状空间模型的非凸求解问题通过凸松弛方法转化为凸规划问题; 在采用 ADMM 算法对凸规划问题进行优化求解过程中, 提出谱范数近端梯度算法保证解的正交性与稀疏性. 利用所提的优化方法, 基于形状空间模型和 3D 可变形状模型在卡内基梅隆大学运动捕获数据库上进行 3D 人体姿态重构, 定性和定量对比实验结果表明本文方法均优于现有的优化方法, 验证了所提方法的有效性.

关键词 3D 重构, 稀疏表示, $L_{1/2}$ 正则化, 凸规划

引用格式 洪金华, 张荣, 郭立君. 基于 $L_{1/2}$ 正则化的三维人体姿态重构. 自动化学报, 2018, 44(6): 1086–1095

DOI 10.16383/j.aas.2018.c170199

3D Human Body Pose Reconstruction via $L_{1/2}$ Regularization

HONG Jin-Hua¹ ZHANG Rong¹ GUO Li-Jun¹

Abstract In order to estimate the 3D shape with the given 2D feature points for a single image, we present a convex relaxation approach based on $L_{1/2}$ regularization by using the shape space model and combining $L_{1/2}$ regularization and the properties of spectral norm. Thereby we transform the non-convex optimization problem in a shape space model into a convex programming problem. When optimizing the solution to the convex programming problem by using the ADMM algorithm, we further propose a spectral norm proximal operator to satisfy the orthogonality and sparsity constraint of the solution. Using the proposed optimization algorithm, we conduct experiments on the CMU motion capture dataset for 3D human body pose reconstruction based on the shape space model and 3D deformable shape model. Comparison experimental results show qualitatively and quantitatively that the proposed algorithm outperforms the existing optimization algorithms. The effectiveness of the proposed algorithm is validated.

Key words 3D reconstruction, sparse representation, $L_{1/2}$ regularization, convex program

Citation Hong Jin-Hua, Zhang Rong, Guo Li-Jun. 3D human body pose reconstruction via $L_{1/2}$ regularization. *Acta Automatica Sinica*, 2018, 44(6): 1086–1095

从 2D 图像中识别 3D 对象是计算机视觉的核心任务之一. 近些年来, 研究人员的研究重点已从基于图像初步识别对象 (用边缘盒标记对象) 转向进一步利用图像构建对象的 3D 几何信息 (如形状、姿态等). 3D 对象几何信息的获取不仅能为高层视觉任务如场景理解、增强现实以及人机交互等提供更丰富的判别信息, 而且也有助于提高对象识别性能^[1–2].

基于单目图像重构对象的 3D 形状本身是一个病态问题. 最近几年, 越来越多的科研人员利用不

断增加的在线 3D 模型数据库进行 3D 模型分析并提取出丰富的形状先验信息, 在此基础上, 深入地开展了基于单目图像重构对象的 3D 形状方面研究. 为解决 3D 形状重构过程中存在的类内可变性、非刚体形变和避免穷举所有可能视角等问题, 受主动形状模型 (Active shape model^[3]) 的启示, 有许多工作^[4–5] 是通过采用一种 3D 可变形状模型 (3D deformable shape model) 来表示形状, 在 3D 可变形状模型中, 一个形状被定义为一个有序特征点的集合, 任意一个形状可由预先定义的一组基础形状线性表示. 考虑到稀疏表示^[6] 的思想是通过一组过完备基将输入信号进行线性表示, 若将上述基础形状视作过完备基, 则 3D 可变形状模型类似于一种稀疏表示. 利用 3D 可变形状模型, 基于单目图像中的 2D 特征点重构对象的 3D 形状的问题可视为 2D 图像中的特征点与 3D 可变形状模型匹配^[7] 的问题, 即 3D-to-2D 形状融合问题. 该融合问题本质上是对形状参数 (稀疏系数) 和视点参数 (照相机外部参数) 进行联合估计的问题. 由于只有视点已知的情况下

收稿日期 2017-04-14 录用日期 2018-02-07
Manuscript received April 14, 2017; accepted February 7, 2018
浙江省自然科学基金 (LY17F030002), 浙江省公益技术研究计划 (LGF18F020007) 资助
Supported by Zhejiang Provincial Natural Science Foundation (LY17F030002), Zhejiang Provincial Public Welfare Technology Research Project (LGF18F020007)
本文责任编辑 杨健
Recommended by Associate Editor YANG Jian
1. 宁波大学信息科学与工程学院 宁波 315211
1. Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211

才能将 3D 可变形模型与单目图像中 2D 特征点进行更好的融合; 或者只有已知 3D 形状模型的情况下才能更好的估计视点. 因此, 形状参数与视点参数的联合估计是一个非凸优化问题. 此外, 还需对摄像机旋转矩阵进行正交约束, 这将会使得问题变的更复杂. 对于上述复杂问题的求解, 以前一般采用交替迭代最小化方法实现形状参数与视点参数的交替更新. 然而, 此方法无法保证求得的解全局最优, 且求得的解对初始值敏感. 为解决初始值设置不理想的问题, 有一些启发式方法被提出, 例如: 对初始值进行多次初始化^[8] 或者使用视点感知检测器进行初始化^[5]. 然而, 这些方法依然无法保证求得的解全局最优.

针对上述形状参数与视点参数交替更新的非凸优化问题, 本文在 3D 形状空间模型 (Shape-space model) 的基础上运用 $L_{1/2}$ 正则化和谱范数的性质将其转化为凸规划问题. 在利用 ADMM 算法对凸规划问题进行优化求解的过程中, 进一步提出运用谱范数近端梯度算法保证解的正交性与稀疏性.

1 相关工作

基于单目图像的 3D 对象重构通常是将 3D 可变形模型和 2D 特征点进行融合, 此类方法的一种典型应用是针对不同的任务进行人脸建模, 例如: 人脸识别^[9], 人脸特征提取^[10] 和人脸动画^[11] 等. 近来, Hejrati 等^[4] 用可变形模型对 3D 汽车进行建模, 他们使用改进的部分可变形模型 (Deformable part model^[12]) 生成 2D 特征点, 然后基于此 2D 模型重构出对象的 3D 形状. Lin 等^[13] 提出一种联合优化三维模型拟合与细粒度分类的方法. Zia 等^[5] 提出用一个概率框架同时进行 2D 特征点定位与重构 3D 物体模型. 在进行 3D 物体建模过程中, 对于刚性物体的形变问题, 传统的主动形状模型能够有效的解决. 然而, 其无法有效解决非刚体物体结构形变的问题. 为此, Ramakrishna 等^[14] 提出一种基于稀疏表示的方法实现从一张静态的带有特征点的 2D 图像中重构 3D 人体姿态. Wang 等^[8] 采用一种 2D 人体姿势探测方法实现节点的自动定位并使用一种健壮的重构方法对节点位置定位不准确的问题进行处理. Fan 等^[15] 提出通过在构建人体姿态字典的同时施加局部约束以达到改进 Wang 等^[8] 所提方法性能的目的. Zhou 等^[16] 提出将人体姿态估计视为匹配问题, 将一个时空姿态模型 (Spatio-temporal pose model) 与从视频中提取的特征点进行匹配. Akhter 等^[17] 提出带有角约束的稀疏表示从而降低重构无效的概率.

在上述所提工作中关键的一步是将 3D 可变形模型与相关图像中 2D 特征点进行融合. 正如引

言所述, 这些工作仍是非凸优化问题, 一般采用交替迭代最小化方法对形状参数和视点参数进行交替更新. 由于交替迭代最小化方法求得的解对初始值比较敏感易陷入局部极值. 为使解具有全局唯一性, 本文提出一种凸松弛方法将非凸优化问题转化为凸规划问题予以解决.

本文所提方法不仅能够解决上述交替迭代最小化方法的缺点, 还能为现有的大多数方法提供一个较理想的初始值, 从而达到改进其性能的目的. 在最近关于 3D 形状重构的研究中有一种趋势是基于某一类对象多个实例的单幅图像集重构该对象的 3D 物体模型^[18-21]. 此类研究中, 采用将可变形模型与在 2D 图像中被标注的特征点进行融合. 本文所要研究的问题正是这些工作的一个基础; 另一种趋势是通过在形状集中确定最近邻的方法实现单幅图像中对象的 3D 形状重构^[22-23]. 其详细步骤如下: 首先枚举 3D 数据集中所有实例和视点获取对应的 2D 形状集; 然后, 在 2D 形状集中找出待测 2D 图像的最邻近, 与此 2D 形状对应的 3D 形状和视点可以当作初始值进行设置; 最后根据图像轮廓对视点和非刚体形变的初始值进行优化. 这种方法会产生丰富的重构结果并适用于广泛的对象类别, 其缺点是计算量大且需对图像进行精准的分割. 运用本文所提方法能够快速找到较理想的视点与形状, 将其作为初始值能够有效地提高现有方法的性能.

2 本文算法

2.1 问题描述

本文研究的问题为从给定 2D 特征点的单目图像中重构对象的 3D 形状, 其过程可以用如下公式进行描述:

$$W = \prod S \quad (1)$$

其中, $S \in \mathbf{R}^{3 \times p}$ 表示含有 p 个特征点的 3D 物体, $W \in \mathbf{R}^{2 \times p}$ 表示 S 在 2D 上的投影, $\prod = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \end{bmatrix}$ 为照相机标记矩阵, α 与焦距和物体的距离有关.

式 (1) 对应的 3D 形状重构问题可通过稀疏表示进行描述, 即运用 3D 可变形模型将任意一个 3D 形状用预先定义的一组基础形状线性表示. 其思想来源于主动形状模型^[3]:

$$S = \sum_{i=1}^k c_i B_i \quad (2)$$

其中, $B_i \in \mathbf{R}^{3 \times p}$ 为一组有序特征点表示的基础形状, c_i 为 B_i 的系数, k 为基础形状的个数, $\forall i \in$

[1, k].

在弱透视相机模型下, 从给定 2D 特征点的单目图像中重构对象的 3D 形状的问题可用如下公式进行描述:

$$W = \prod \left(R \sum_{i=1}^k c_i B_i + \mathbf{T} \mathbf{1}^T \right) \quad (3)$$

其中, $R \in \mathbf{R}^{3 \times 3}$ 为旋转矩阵和 $\mathbf{T} \in \mathbf{R}^3$ 为平移向量, 且 R 为属于特殊正交群中的旋转矩阵, 即:

$$SO(3) = \{R \in \mathbf{R}^{3 \times 3} | R^T R = I_3, \det R = 1\} \quad (4)$$

对式 (3) 进行中心化, 我们可以进一步简化为如下:

$$S = \bar{R} \sum_{i=1}^k c_i B_i \quad (5)$$

其中, $\bar{R} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \end{bmatrix} \times R$, R 为上述的旋转矩阵, $\hat{c}_i = \alpha \times c_i$, $\forall i \in [1, k]$, 为了描述简化, 依然用 c_i 来代替 \hat{c}_i . 在不考虑噪声的情况下, 利用稀疏表示理论可将式 (5) 的解转化为求解如下公式:

$$\min_{\mathbf{C}} \|\mathbf{C}\|_0 \quad \text{s.t.} \quad W = \bar{R} \sum_{i=1}^k c_i B_i, \bar{R} \bar{R}^T = I_2 \quad (6)$$

其中, $\mathbf{C} = [c_1, \dots, c_k]$ 为稀疏系数向量, $\|\mathbf{C}\|_0$ 称为 L_0 范数, 代表 \mathbf{C} 中非零元素的个数. 运用拉格朗日乘法将式 (6) 转化为 L_0 正则化问题来求解:

$$\min_{\mathbf{C}, \bar{R}} \left\| W - \bar{R} \sum_{i=1}^k c_i B_i \right\|_F^2 + \lambda \|\mathbf{C}\|_0 \quad \text{s.t.} \quad \bar{R} \bar{R}^T = I_2 \quad (7)$$

其中, λ 是正则化参数, $\|\cdot\|_F$ 代表矩阵 F 范数, $\|\cdot\|^2$ 代表欧几里得范数. 式 (7) 是一个损失函数, 第一项是重投影误差, 第二项是 L_0 正则化. 式 (7) 本质上是一个带有正交性和稀疏性约束的非凸优化问题, 导致式 (7) 非凸的原因有两个: 其一, 式 (7) 中 L_0 正则化是非凸的、不连续的, 也是一个 NP 难问题; 其二, 式 (7) 中存在两个未知变量, 即存在二义性. 为解决上述原因引起的非凸优化问题, 本文采用一种凸松弛方法将其转化为凸规划问题并予以解决.

2.2 基于形状空间模型的凸松弛方法

形状空间模型可用如下公式表示:

$$S = \sum_{i=1}^k c_i R_i B_i \quad (8)$$

式 (8) 表示任意一个形状能够由可旋转的基础形状线性表出. 式 (8) 可以消除式 (5) 的二义性, 通过将

式 (8) 代入式 (1) 得到式 (9):

$$W = \prod \sum_{i=1}^k c_i R_i B_i = \sum_{i=1}^k M_i B_i \quad (9)$$

$$\text{s.t.} \quad M_i M_i^T = c_i^2 I_2, \forall i \in [1, k]$$

其中, $M_i = c_i \bar{R}_i$, \bar{R}_i 为 R_i 的前两行, 再次运用拉格朗日乘法将式 (9) 进一步转化为如下:

$$\min_{\mathbf{C}, M} \frac{1}{2} \left\| W - \sum_{i=1}^k M_i B_i \right\|_F^2 + \lambda \|\mathbf{C}\|_0 \quad (10)$$

$$\text{s.t.} \quad M_i M_i^T = c_i^2 I_2, \forall i \in [1, k]$$

式 (10) 是一个带有正交性与稀疏性约束的 L_0 正则化问题, 也是一个非凸优化问题. 本文利用定理 1 和 $L_{1/2}$ 正则化将非凸优化问题转化为谱范数正则化问题, 转化后的问题将变成一种凸规划问题, 这是本文的主要贡献之一.

定理 1. 给定一个缩放因子 s , $S = \{Y \in \mathbf{R}^{m \times n} | Y^T Y = s^2 I_n\}$ 的凸包等价于半径为 $|s|$ 谱范数球: $\text{con}(S) = \{Y \in \mathbf{R}^{m \times n} | \|Y\|_2 \leq |s|\}$.

定理 1 的具体证明可以参考文献 [24–25]. 由于 $M_i = c_i \bar{R}_i$, 由定理 1 可得出 $\|M_i\|_2 \leq |c_i|$, 在迭代过程中需保证 M_i 具有正交性与稀疏性. 故运用定理 1 和形状空间模型将式 (6) 改写为

$$\min_{\mathbf{C}} \|\mathbf{C}\|_0 \quad (11)$$

$$\text{s.t.} \quad W = \sum_{i=1}^k M_i B_i, \|M_i\|_2 \leq |c_i|, \forall i \in [1, k]$$

其中, $\|\cdot\|_2$ 代表矩阵 2 范数, $|\cdot|$ 代表求绝对值. 在考虑噪声的情况下, 式 (11) 可进一步转化为如下形式:

$$\min_{\mathbf{C}} \|\mathbf{C}\|_0 \quad (12)$$

$$\text{s.t.} \quad \left\| W - \sum_{i=1}^k M_i B_i \right\|_F^2 < \varepsilon, \|M_i\|_2 \leq |c_i|,$$

$$\forall i \in [1, k]$$

其中, ε 为误差界, 式 (12) 中 L_0 范数会导致 NP 难、噪声敏感等问题, 处理起来非常棘手. 对于式 (12) 的求解一般是运用拉格朗日乘法将其转化为 L_0 正则化问题. 然而, 转化后 L_0 正则化问题实质上依然是非凸的、不连续的, 同时也是一个 NP 难问题. 为此, Tibshirani 等^[26] 松弛 L_0 正则化条件将其转化为 L_1 正则化进行求解, Chen 等^[27–28] 提出基匹配追踪算法, 实现了用 L_1 正则化替代 L_0 正则化进

行问题求解, 并且 Elad 等^[29] 和 Donoho 等^[30] 证明了 L_1 与 L_0 在某些条件下具有等价性, 比如: RIP 等. 然而, 在现实生活中 L_1 与 L_0 等价的条件是比较难满足的. 因此, 徐宗本等^[31] 提出用 $L_{1/2}$ 正则化来替代 L_1 正则化, 实现对 L_0 正则化问题的近似求解, 并且徐宗本等证明了 $\|\mathbf{C}\|_{1/2}^{1/2}$ 梯度分量的存在且通过实验证明了 $L_{1/2}$ 正则化求得的解比 L_1 正则化的解更稀疏和更有效. 故本文采用 $L_{1/2}$ 正则化将式 (12) 转化为如下形式:

$$\begin{aligned} & \min_{\mathbf{C}} \|\mathbf{C}\|_{1/2}^{1/2} \\ \text{s. t. } & \left\| W - \sum_{i=1}^k M_i B_i \right\|_F^2 < \varepsilon, \|\mathbf{C}\|_{1/2}^{1/2} = \sum_{i=1}^k |c_i|^{1/2}, \\ & \|M_i\|_2 \leq |c_i|, \forall i \in [1, k] \end{aligned} \quad (13)$$

其中, $\|\mathbf{C}\|_{1/2}$ 称为 \mathbf{C} 的 $L_{1/2}$ 拟范数, 由线性反问题可知, 将式 (13) 进一步改为如下公式:

$$\min_M \sum_{i=1}^k \|M_i\|_2^{1/2} \quad \text{s. t. } \left\| W - \sum_{i=1}^k M_i B_i \right\|_F^2 < \varepsilon \quad (14)$$

利用拉格朗日乘子法将式 (14) 转化为如下形式:

$$\min_{M_1, \dots, M_k} \frac{1}{2} \left\| W - \sum_{i=1}^k M_i B_i \right\|_F^2 + \lambda \|M_i\|_2^{1/2} \quad (15)$$

经过上述一系列的变化, 将从给定 2D 特征点的单目图像中重构对象的 3D 形状问题转化为对式 (15) 进行求解. 该公式是一个带有惩罚约束项的谱范数正则化问题, 本质上是一个带有约束项的凸规划问题. 对于式 (15) 的求解将在第 3 节详细阐述.

2.3 重构算法

2.3.1 基于形状空间模型的 3D 形状重构

通过对式 (15) 的求解, 可以得到解 $M = [M_1, \dots, M_k]$. 为重构对象的 3D 形状, 本文先利用 c_i, R_i, M_i 之间的关系计算 c_i, R_i , 进而实现对象的 3D 形状重构. 同时, 为消除重构的任意性, 可令 $c_i \geq 0$, 基于形状空间模型重构 3D 形状 S 的具体步骤总结为算法 1.

算法 1. 基于形状空间模型的 3D 形状重构

输入. M_1, \dots, M_k

输出. S

步骤.

- 1) for $i = 1 : k$ do
- 2) $c_i = \|M_i\|_2$;
- 3) $\mathbf{r}_i^{(1)} = \mathbf{m}_i^{(1)}/c_i$;

- 4) $\mathbf{r}_i^{(2)} = \mathbf{m}_i^{(2)}/c_i$;
- 5) $\mathbf{r}_i^{(3)} = \mathbf{r}_i^{(1)} \times \mathbf{r}_i^{(2)}$;
- 6) $R_i = [\mathbf{r}_i^{(1)}, \mathbf{r}_i^{(2)}, \mathbf{r}_i^{(3)}]^T$;
- 7) end for;
- 8) $S = \sum_{i=1}^k c_i R_i B_i$;
- 9) 算法结束.

* $\mathbf{m}_i^{(1)}$ 为 M_i 的第一个行向量

2.3.2 基于 3D 可变形模型的 3D 形状重构

本文所提方法不仅能够直接利用形状空间模型进行 3D 形状重构, 而且还能为基于 3D 可变形模型 (式 (2)) 的 3D 重构提供较好的初始值, 达到优化 3D 形状重构的目的. 采用基于流形的图像投影配准算法^[32] 对式 (16) 进行求解, 实现为基于 3D 可变形模型的 3D 形状重构提供较理想的初始值 C, \bar{R} :

$$\min_{c_1, \dots, c_k, \bar{R}} \|M_i - c_i \bar{R}\|_F^2 \quad \text{s. t. } \bar{R} \bar{R}^T = I_2 \quad (16)$$

然后采用交替迭代最小化方法对 3D 形状重构进行优化, 具体重构步骤总结为算法 2.

算法 2. 基于 3D 可变形模型的 3D 形状重构

输入. M_1, \dots, M_k

输出. S

步骤.

- 1) 初始化 \mathbf{C}, \bar{R} ; /* 图像投影配准算法 */
- 2) 优化化 \mathbf{C}, \bar{R} ; /* 交替迭代最小化方法 */
- 3) $S = \sum_{i=1}^k c_i B_i$;
- 4) 算法结束.

3 模型优化求解

式 (15) 是一个带有正交性与稀疏性约束的谱范数正则化问题. 为实现对式 (15) 的优化求解, 本文采用 ADMM^[33] 算法对其进行优化; 在用 ADMM 算法对 M 进行优化求解的过程, 出于保证矩阵 M 的正交性与稀疏性约束的考虑, 本文提出谱范数近端梯度算法予以解决, 这是本文的另一个主要贡献.

3.1 谱范数近端梯度算法

在介绍用 ADMM 算法对式 (15) 进行具体求解之前, 本文将证明定理 2, 其是 ADMM 算法实现对 M_i 矩阵求解中重要一环.

定理 2. 对一般问题

$$\min_{M_i} \left(\frac{1}{2} \|Y - M_i\|_F^2 + \lambda \|M_i\|_2^{1/2} \right) \quad (17)$$

的解可由如下式 (18) 求得:

$$M_i^* = D_{\sqrt{\lambda}}(Y) = U_Y \text{diag} \left\{ \sigma_Y - \sqrt{\lambda} P_{L_1} \left(\frac{\sigma_Y}{\sqrt{\lambda}} \right) \right\} V_Y^T \quad (18)$$

其中, U_Y, V_Y, σ_Y 分别表示 Y 奇异值分解后左边的酉矩阵, 右边的酉矩阵以及特征值向量, P_{L_1} 表示为一个向量在 L_1 范数单位球上的投影.

证明. 式 (17) 本质上是一个关于近端梯度算法的问题, 下面将证明其解可由式 (18) 求得. 为此, 引入谱范数近端梯度算法:

$$\text{prox}_{\lambda F}(Y) = \arg \min_{M_i} \left(\frac{1}{2} \|Y - M_i\|_F^2 + \lambda F(M_i) \right) \quad (19)$$

其中, $F(M_i) = \|M_i\|_2^{1/2} = \|\sigma_{M_i}\|_\infty^{1/2}$, 式 (19) 的解为式 (20)^[34]:

$$\text{prox}_{\lambda F}(Y) = U_Y \text{diag}\{\text{prox}_{\lambda f}(\sigma_Y)\} V_Y^T \quad (20)$$

其中, $f(\sigma_Y) = \|\sigma_Y\|_\infty^{1/2}$, 对于式 (20) 使用莫罗分解^[34] 可得:

$$\begin{aligned} \text{prox}_{\lambda f}(\sigma_Y) &= \sigma_Y - \text{prox}_{(\lambda f)^*}(\sigma_Y) = \\ &= \sigma_Y - \text{prox}_{\lambda f(\cdot/\lambda)^*}(\sigma_Y) = \\ &= \sigma_Y - \text{prox}_{\sqrt{\lambda} f(\cdot/\sqrt{\lambda})^*}(\sigma_Y) = \\ &= \sigma_Y - \text{prox}_{\frac{\sqrt{\lambda}}{\lambda}(\sqrt{\lambda} f)^*}(\sigma_Y) = \\ &= \sigma_Y - \sqrt{\lambda} P_{L_1} \left(\frac{\sigma_Y}{\sqrt{\lambda}} \right) \end{aligned} \quad (21)$$

其中, L_1 范数是 L_∞ 范数的对偶范数, $\text{prox}_{(\lambda f)^*}(\sigma_Y)$ 为 $\text{prox}_{\lambda f}(\sigma_Y)$ 的共轭函数的近端梯度算子. \square

经过上述一系列的证明, 可知式 (17) 的解可由式 (18) 求得.

3.2 基于 ADMM 算法求解

在实际应用中, 图像中 2D 特征点是通过探测得到的, 因此有异常值存在的情况, 需要进行异常值建模^[25], 同时注意到平移向量在中心化的过程是不能被消除的, 故式 (15) 可被转化为如下公式:

$$\begin{aligned} \min_{M_1, \dots, M_k, E, \mathbf{T}} \frac{1}{2} \left\| W - \sum_{i=1}^k M_i B_i - E - \mathbf{T} \mathbf{1}^T \right\|_F^2 + \\ \lambda \sum_{i=1}^k \|M_i\|_2^{1/2} + \beta \|E\|_1 \end{aligned} \quad (22)$$

其中, E 为异常值模拟矩阵, \mathbf{T} 为平移向量. 增加一个辅助变量, 将式 (22) 重写成如下形式:

$$\begin{aligned} \min_{\tilde{M}, \tilde{Z}, E, \mathbf{T}} \frac{1}{2} \left\| W - Z \tilde{B} - E - \mathbf{T} \mathbf{1}^T \right\|_F^2 + \\ \lambda \sum_{i=1}^k \|M_i\|_2^{1/2} + \beta \|E\|_1 \\ \text{s. t. } \tilde{M} = Z \end{aligned} \quad (23)$$

其中, $\tilde{M} = [M_1, \dots, M_k]$, $\tilde{B} = [B_1, \dots, B_k]^T$. 对式 (23) 运用增广拉格朗日乘法, 可得出如下公式:

$$\begin{aligned} L_\mu(\tilde{M}, Z, E, \mathbf{T}, Y) = \\ \frac{1}{2} \|W - Z \tilde{B} - E - \mathbf{T} \mathbf{1}^T\|_F^2 + \lambda \sum_{i=1}^k \|M_i\|_2^{1/2} + \\ \beta \|E\|_1 + \langle Y, \tilde{M} - Z \rangle + \frac{\mu}{2} \|\tilde{M} - Z\|_F^2 \end{aligned} \quad (24)$$

其中, β 是一非负实参数, Y 是对偶变量, μ 是优化过程中控制迭代步长的参数, 通过运用 ADMM 算法进行求解, 其具体求解步骤如下:

$$\tilde{M}^{t+1} = \arg \min_{\tilde{M}} L_\mu(\tilde{M}, Z^t, E^t, \mathbf{T}^t, Y^t) \quad (25)$$

$$Z^{t+1} = \arg \min_Z L_\mu(\tilde{M}^{t+1}, Z, E^t, \mathbf{T}^t, Y^t) \quad (26)$$

$$E^{t+1} = \arg \min_E L_\mu(\tilde{M}^{t+1}, Z^{t+1}, E, \mathbf{T}^t, Y^t) \quad (27)$$

$$\mathbf{T}^{t+1} = \arg \min_{\mathbf{T}} L_\mu(\tilde{M}^{t+1}, Z^{t+1}, E^{t+1}, \mathbf{T}, Y^t) \quad (28)$$

$$Y^{t+1} = Y^t + \mu(\tilde{M}^{t+1} - Z^{t+1}) \quad (29)$$

其中, 上标 $t+1$ 为迭代次数, 对于迭代式 (25) 可转化为式 (30):

$$\begin{aligned} \tilde{M}^{t+1} = \min_{\tilde{M}} \frac{1}{2} \left\| \tilde{M} - Z^t + \frac{1}{\mu} Y^t + E^t + \mathbf{T}^t \mathbf{1}^T \right\|_F^2 + \\ \lambda \sum_{i=1}^k \|M_i\|_2^{1/2} \end{aligned} \quad (30)$$

运用谱范数近端梯度算法, 将式 (30) 转化为式 (31):

$$M_i^{t+1} = D_{\sqrt{\frac{\lambda}{\mu}}}(Q_i^t), \forall i \in [1, k] \quad (31)$$

其中, $[Q_1^t, \dots, Q_k^t] = Z^t - \frac{1}{\mu} Y^t - E^t - \mathbf{T}^t \mathbf{1}^T$, $Q_i^t \in \mathbf{R}^{2 \times 3}, \forall i \in [1, k]$. 将上述式 (26)、式 (27)、式 (28) 依次展开可得出 $Z^{t+1}, E^{t+1}, \mathbf{T}^{t+1}$ 解的形式:

$$\begin{aligned} Z^{t+1} = ((W - E^t - \mathbf{T}^t \mathbf{1}^T) \tilde{B}^T + \mu \tilde{M}^{t+1} + Y^t) \times \\ (\tilde{B} \tilde{B}^T + \mu I)^{-1} \end{aligned} \quad (32)$$

$$E^{t+1} = S_\beta(W - Z^{t+1} B - \mathbf{T}^t \mathbf{1}^T) \quad (33)$$

$$\mathbf{T}^{t+1} = g(W - Z^{t+1} B - E^{t+1}) \quad (34)$$

其中, $S_\beta(X)_{ij} = \text{sign}(X_{ij})(X_{ij} - \beta)_+$, 表示为对每个元素进行软阈值计算, $g(\cdot)$ 为对矩阵按行求均值. 经过 ADMM 算法产生的每个迭代步骤的解与原问题 (22) 的解等价.

4 实验

4.1 字典学习

字典学习是稀疏表示中提取特征的一个重要环节, 字典的好坏会直接影响 3D 形状重构效果以及

稀疏表示的稀疏性. 本文采用矩阵分解和稀疏编码的在线学习方法^[35]学习一组基础形状作为 3D 形状到过完备字典, 本文采用参考文献 [24–25] 中的 15 节点模型来表示基础形状. 学习过程如下:

$$\begin{aligned} \min_{B_1, \dots, B_k, c} & \sum_{j=1}^n \frac{1}{2} \left\| S_j - \sum_{i=1}^k c_{ij} B_i \right\|_F^2 + \alpha \sum_{i,j} c_{ij} \\ \text{s. t. } & c_{ij} \geq 0, \|B_i\|_F \leq 1, \forall i \in [1, k], \forall j \in [1, n] \end{aligned} \quad (35)$$

其中, α 为非负实参数, S_j 为经 Procrustes method 方法对齐后的训练集中的第 j 个 3D 形状, B_i 为字典中第 i 个原子, c_{ij} 为第 j 个训练集中 B_i 的系数, $\|B_i\|_F \leq 1$ 为防止 c_{ij} 、 B_i 任意改变, $c_{ij} \geq 0$ 表示非负稀疏编码.

运用非负稀疏编码, 本文将学习出一个大小为 $B^{45 \times 128}$ 的冗余字典, 其中 128 为原子的个数, 45 为 15 个节点的三维坐标, 参数 α 设为 1.

4.2 3D 人体姿态重构

本文将采用两种重构模型对 3D 人体姿态进行重构, 第一种是基于形状空间模型的 3D 形状重构, 第二种是基于 3D 可变形状模型的 3D 形状重构. 对于每种重构模型本文都进行了定性实验与定量实验, 以验证本文所提方法的性能优于参考文献 [25] 中所提的方法以及交替迭代最小化方法.

对于定性实验, 本文采用参考文献 [25] 中代码所提供的数据进行实验, 通过对单目图像中的 2D 骨架进行 3D 重构, 然后将重构的 3D 骨架经过相同的视角将其投影生成 2D 骨架.

对于定量实验, 本文将从重构误差和稀疏度两个维度进行定量实验分析: 首先, 选取卡内基梅隆大学运动捕获数据库^[36]中的八个 (Walk、Run、Jump、Climb、Box、Dance、Sit、Basketball) 作为定量实验数据, 其与参考文献 [25] 中的实验数据一致; 其次, 2D 骨架是通过正交照相机绕着每个序列旋转 360° 获取的; 最后, 三种方法的对齐方式为: 在相机坐标系下通过平移缩放使之与真实的 3D 进行对齐. 然后, 计算两者之间的欧氏距离 (单位, mm) 和稀疏度 (零元素的个数), 并对重构误差以及稀疏度取均值.

本文所提方法的迭代停止条件为 $\|\hat{M} - \tilde{M}\|_F / \|\tilde{M}\|_F < 10^{-3}$, 与参考文献 [25] 中所提方法的迭代停止条件一样, 其中 \hat{M} 是由 $W = \sum_{i=1}^k M_i B_i$ 估计得出的, \tilde{M} 为本文所提方法求解得出的. 同时, 本文中所有方法的最大迭代次数均设为 1500 次. 只要重构方法达到迭代停止条件或者最大迭代次数时, 结束实验.

4.2.1 基于形状空间模型的重构

本文主要采用三种方法进行对比实验, 第一种为 $L_{1/2} + \text{convex}$ 方法 (本文所提的方法), 第二种为 $L_1 + \text{convex}$ 方法 (参考文献 [25] 所提的 convex 方法), 第三种为 altern 方法 (交替迭代最小化方法). 三种方法的参数与参考文献 [25] 一致. 为验证本文所提方法的有效性, 本文将采用定性实验与定量实验进行对比分析.

三种方法的定性实验效果如图 1 所示. 由左到右依次为图像中 2D 骨架, $L_{1/2} + \text{convex}$ 方法, $L_1 + \text{convex}$ 方法, altern 方法, 右边三列表示重构生成的 3D 骨架经相同的视角投影生成的 2D 骨架.

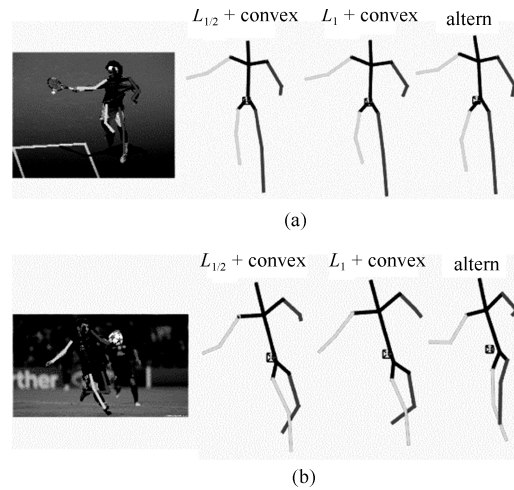


图 1 三种方法的定性实验效果对比图

Fig. 1 The comparison of qualitative experiment results of three methods

从图 1 可清楚地看出本文所提方法取得了较好效果: (a) 中 $L_{1/2} + \text{convex}$ 方法右脚有较明显的弯曲, 其他两种方法均没有, 明显可看出 $L_{1/2} + \text{convex}$ 方法与图像中 2D 骨架更接近; (b) 中 $L_{1/2} + \text{convex}$ 方法的左手比其他两种方法更接近图像中 2D 骨架. 三种方法定量实验中重构误差对比图如图 2 所示.

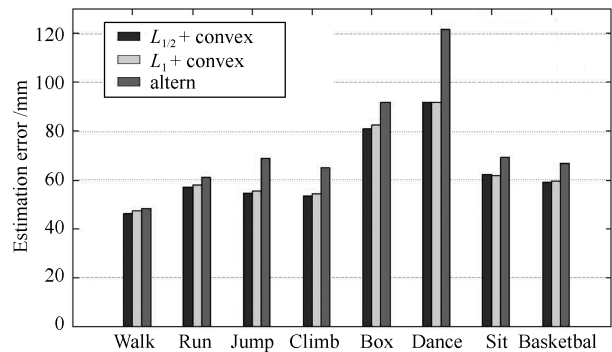


图 2 三种方法的重构误差对比图

Fig. 2 The reconstruction error comparison of three methods

本文采用盒图对三种方法的重构误差的离散分布情况进行统计分析,其实验效果图如图 3 所示.

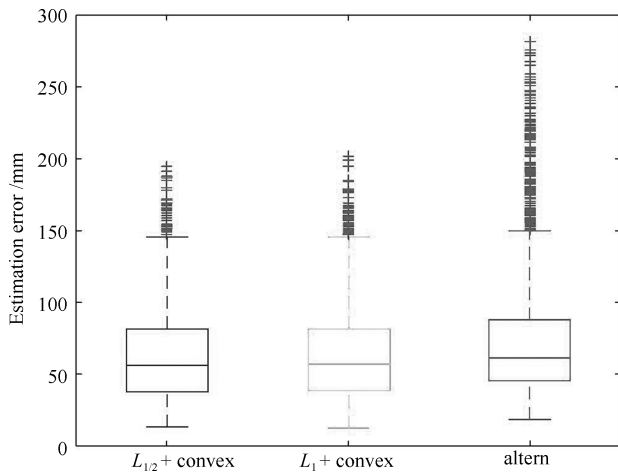


图 3 重构误差的盒图

Fig. 3 The box diagram of reconstruction error

稀疏度是定量实验中另一个重要的评价指标,三种方法在八个序列上稀疏度对比效果图如图 4 所示.

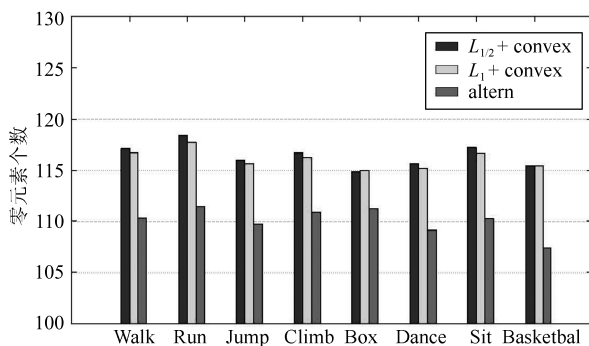


图 4 三种方法的稀疏度对比图

Fig. 4 The sparse contrast graphs of three methods

从图 2、4 可看出,在全部序列上, $L_{1/2} + \text{convex}$ 方法全面超过 altern 方法. 在 Walk、Run、Jump、Climb 序列上, $L_{1/2} + \text{convex}$ 方法在重构误差与稀疏度两个方面都取得了明显优于 $L_1 + \text{convex}$ 方法的实验结果;在 Dance 和 Basketball 序列上, $L_{1/2} + \text{convex}$ 方法略优于 $L_1 + \text{convex}$ 方法;在 Box 序列上, $L_{1/2} + \text{convex}$ 方法的重构误差明显优于 $L_1 + \text{convex}$ 方法,但稀疏度不如 $L_1 + \text{convex}$ 方法;在 Sit 序列上, $L_{1/2} + \text{convex}$ 方法的重构误差稍逊于 $L_1 + \text{convex}$ 方法,而稀疏度则优于 $L_1 + \text{convex}$ 方法. 从图 3 的盒图可以看出, $L_1 + \text{convex}$ 方法的异常点个数明显少于其他两种方法.

综合图 2~4 来看, $L_{1/2} + \text{convex}$ 方法总体取得了较理想的实验效果,验证了本文所提方法的有

效性.但在某些序列上,存在重构误差或者稀疏度不理想的情况,其原因可能是由于不同序列的最优正则化参数不一致所造成的.

4.2.2 基于 3D 可变形状模型的重构

在利用 3D 可变形模型进行重构方面,本文也对比了三种不同方法:第一种为 $L_{1/2} + \text{convex} + \text{refine}$ 方法(本文所提的方法),第二种为 $L_1 + \text{convex} + \text{refine}$ 方法(参考文献 [25] 所提的 $\text{convex} + \text{refine}$ 方法),第三种为 altern 方法(与上述 altern 方法相同). $L_{1/2} + \text{convex} + \text{refine}$ 方法的正则化参数设为 0.8,其他参数与另两种方法相同,均采用参考文献 [25] 所给的参数.我们同样采用定性与定量实验进行对比分析.

三种方法的定性实验效果如图 5 所示.图 5 中从左到右依次为图像中 2D 骨架, $L_{1/2} + \text{convex} + \text{refine}$ 方法, $L_1 + \text{convex} + \text{refine}$ 方法,altern 方法,右边三列表示重构生成的 3D 骨架经相同的视角投影生成的 2D 骨架.

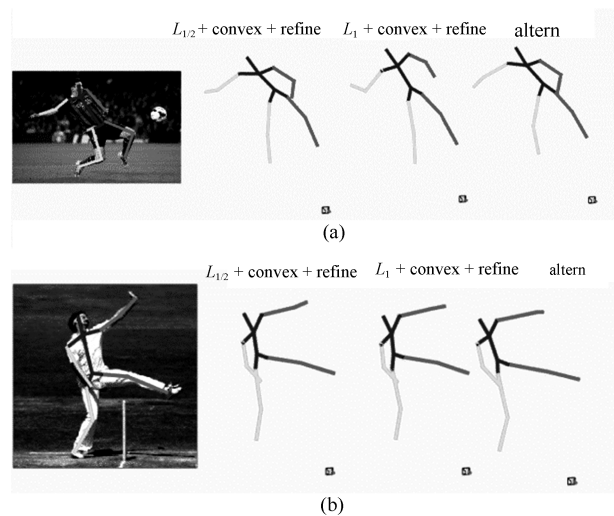


图 5 三种方法的定性实验效果对比图

Fig. 5 The qualitative experiment effect contrast chart of the three methods

从图 5 可清楚地看出本文所提方法取得了较好效果:图 5(a)中 $L_{1/2} + \text{convex} + \text{refine}$ 方法的左手弯曲与 2D 骨架最接近;图 5(b)中 $L_{1/2} + \text{convex} + \text{refine}$ 方法的右手有稍明显的向上弯曲,而其他两种方法均没有.为更好地量化三种方法的对比实验效果,将从重构误差与稀疏度两个方面的定量实验进一步分析.

三种方法的重构误差对比图如图 6 所示,三种方法的重构误差的离散分布情况统计分析的实验效果对比图如图 7 所示,三种方法在八个序列上稀疏度对比图如图 8 所示.从图 6 可看出,在八个序列上, $L_{1/2} + \text{convex} + \text{refine}$ 方

法均取得了较好的重构结果; 从图 8 可看出, 在 Walk、Run、Jump、Climb、Dance、Sit 等 6 个序列上, $L_{1/2} + \text{convex} + \text{refine}$ 方法都取得了最好的稀疏效果; 在 Box、Basketball 两个序列上, $L_{1/2} + \text{convex} + \text{refine}$ 方法的稀疏效果稍逊于 $L_1 + \text{convex} + \text{refine}$ 方法, 但优于 altern 方法. 从图 7 可看出, $L_{1/2} + \text{convex} + \text{refine}$ 方法的异常点个数少于其他两种方法, 其异常点分布相对另两种方法而言, 比较接近盒图的最大观测值. 而且, 在盒图中 $L_{1/2} + \text{convex} + \text{refine}$ 方法的胡须长度最短.

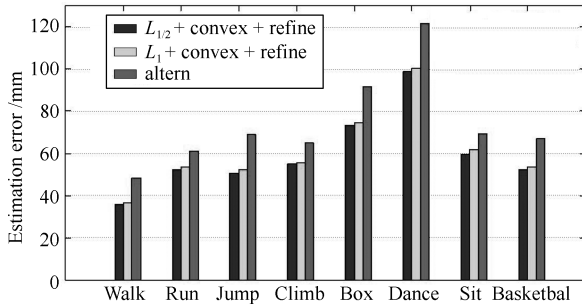


图 6 三种方法的重构误差对比图

Fig. 6 The reconstruction error contrast chart of three methods

析, 最终得出本文所提方法的性能优于其他两种方法.

虽然本文所提方法取得了较理想的实验效果, 但是其收敛速度不够理想, 本文将尝试使用加速近端梯度算法进一步加快算法收敛, 完善本文; 基于深度学习在计算机视觉方面的成功运用, 后续将考虑采用深度学习算法实现 3D 骨架重构^[37-38]; 基于视频具有更加丰富的信息, 将尝试在视频中使用本文所提方法^[39], 后期将在智能制造和机器人领域使用本文所提方法.

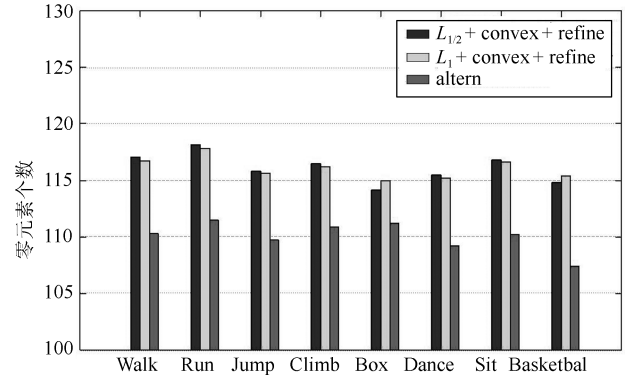


图 8 三种方法的稀疏度对比图

Fig. 8 The sparse contrast graphs of three methods

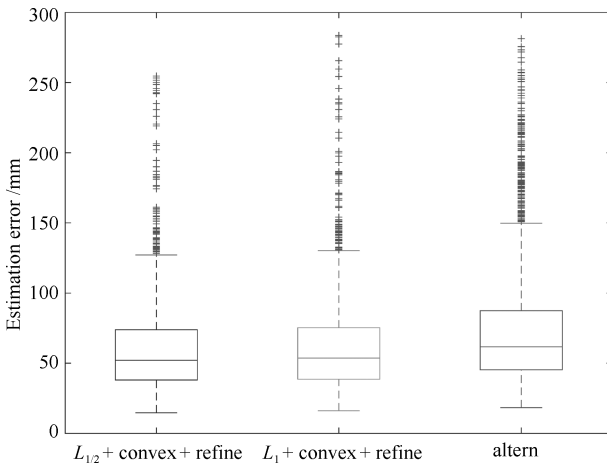


图 7 重构误差的盒图

Fig. 7 The box diagram of reconstruction error

5 结束语

针对从给定 2D 特征点的单目图像中重构对象的 3D 形状问题, 本文在形状空间模型的基础上, 结合 $L_{1/2}$ 正则化和谱范数的性质提出一种基于 $L_{1/2}$ 正则化的凸松弛方法, 将非凸优化问题转化为凸规划问题; 在采用 ADMM 算法对凸规划问题进行优化求解过程中, 为保证求得的解具有正交性与稀疏性, 本文提出谱范数近端梯度算法予以解决. 为验证本文所提方法的有效性, 本文采用两种重构模型进行 3D 重构, 通过定性实验和定量实验进行对比分

References

- Fidler S, Dickinson S, Urtasun R. 3D object detection and viewpoint estimation with a deformable 3D cuboid model. In: Proceedings of the 2012 International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates Inc., 2012. 611-619
- Simo-Serra E, Quattoni A, Torras C, Moreno-Noguer F. A joint model for 2D and 3D pose estimation from a single image. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013. 3634-3641
- Cootes T F, Taylor C J, Cooper D H, Graham J. Active shape models-their training and application. *Computer Vision and Image Understanding*, 1995, **61**(1): 38-59
- Hejrati M, Ramanan D. Analyzing 3D objects in cluttered images. In: Proceeding of the 2012 International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates Inc., 2012. 593-601
- Zia M Z, Stark M, Schiele B, Schindler K. Detailed 3d representations for object recognition and modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(11): 2608-2623
- Fang Hong, Yang Hai-Rong. Greedy algorithms and compressed sensing. *Acta Automatica Sinica*, 2011, **37**(12): 1413-1421
(方红, 杨海蓉. 贪婪算法与压缩感知理论. *自动化学报*, 2011, **37**(12): 1413-1421)

- 7 Zhou Yu, Liu Jun-Tao, Bai Xiang. Research and perspective on shape matching. *Acta Automatica Sinica*, 2012, **38**(6): 889–910
(周瑜, 刘俊涛, 白翔. 形状匹配方法研究与展望. *自动化学报*, 2012, **38**(6): 889–910)
- 8 Wang C Y, Wang Y Z, Lin Z C, Yuille A L, Gao W. Robust estimation of 3D human poses from a single image. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE, 2014. 2369–2376
- 9 Blanz V, Vetter T. Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(9): 1063–1074
- 10 Gu L, Kanade T. 3D alignment of face in a single image. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA: IEEE, 2006. 1305–1312
- 11 Cao C, Weng Y L, Lin S, Zhou K. 3D shape regression for real-time facial animation. *ACM Transactions on Graphics*, 2013, **32**(4): Article No. 41
- 12 Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA: IEEE, 2008. 1–8
- 13 Lin Y L, Morariu V I, Hsu W, Davis L S. Jointly optimizing 3D model fitting and fine-grained classification. In: Proceedings of the 2014 European Conference on Computer Vision, Lecture Notes in Computer Science, Vol. 8692. Heidelberg, Berlin, Germany: Springer, 2014. 466–480
- 14 Ramakrishna V, Kanade T, Sheikh Y. Reconstructing 3D human pose from 2D image landmarks. In: Proceedings of the 2012 European Conference on Computer Vision, Lecture Notes in Computer Science, Vol. 7575. Heidelberg, Berlin, Germany: Springer, 2012. 573–586
- 15 Fan X C, Zheng K, Zhou Y J, Wang S. Pose locality constrained representation for 3D human pose reconstruction. In: Proceedings of the 2014 European Conference on Computer Vision, Lecture Notes in Computer Science, Vol. 8689. Heidelberg, Berlin, Germany: Springer, 2014. 174–188
- 16 Zhou F, de la Torre F. Spatio-temporal Matching for human detection in video. In: Proceedings of the 2014 Computer Vision, Lecture Notes in Computer Science, Vol. 8694. Heidelberg, Berlin, Germany: Springer, 2014. 62–77
- 17 Akhter I, Black M J. Pose-conditioned joint angle limits for 3D human pose reconstruction. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 1446–1455
- 18 Cashman T J, Fitzgibbon A W. What shape are dolphins? Building 3D morphable models from 2D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(1): 232–244
- 19 Vicente S, Carreira J, Agapito L, Batosta J. Reconstructing PASCAL VOC. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014. 41–48
- 20 Carreira J, Kar A, Tulsiani S, Malik J. Virtual view networks for object reconstruction. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 2937–2946
- 21 Kar A, Tulsiani S, Carreira J, Malik J. Category-specific object reconstruction from a single image. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 1966–1974
- 22 Su H, Huang Q X, Mitra N J, Li Y Y, Guibas L. Estimating image depth using shape collections. *ACM Transactions on Graphics*, 2014, **33**(4): Article No. 37
- 23 Huang Q X, Wang H, Koltun V. Single-view reconstruction via joint analysis of image and shape collections. *ACM Transactions on Graphics*, 2015, **34**(4): Article No. 87
- 24 Zhou X W, Leonardos S, Hu X Y, Daniilidis K. 3D shape estimation from 2d landmarks: a convex relaxation approach. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 4447–4455
- 25 Zhou X W, Zhu M L, Leonardos S, Daniilidis K. Sparse representation for 3D shape estimation: a convex relaxation approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(8): 1648–1661
- 26 Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2011, **73**(3): 273–282
- 27 Chen S, Donoho D. Basis pursuit. In: Proceedings of the 1994 Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE, 2002, **1**: 41–44
- 28 Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit. *Siam Review*, 2001, **43**(1): 129–159
- 29 Elad M, Bruckstein A M. A generalized uncertainty principle and sparse representation in pairs of bases. *IEEE Transactions on Information Theory*, 2002, **48**(9): 2558–2567
- 30 Donoho D L, Huo X. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 2001, **47**(7): 2845–2862
- 31 Xu Z B, Zhang H, Wang Y, Change X Y, Liang Y. $L_{1/2}$ regularization. *Science China Information Sciences*, 2010, **53**(6): 1159–1169

- 32 Del Bue A, Xavier J, Agapito L, Paladini M. Bilinear modeling via augmented lagrange multipliers (BALM). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(8): 1496–1508
- 33 Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2010, **3**(1): 1–122
- 34 Parikh N, Boyd S. Proximal algorithms. *Foundations and Trends in Optimization*, 2013, **1**(3): 123–231
- 35 Mairal J, Bach F, Ponce J, Sapiro G. Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 2010, **11**: 19–60
- 36 Mocap: Carnegie Mellon university motion capture database [Online], available: <http://Mocap.cs.cmu.edu/>, March 1, 2017
- 37 Zhu Yu, Zhao Jiang-Kun, Wang Yi-Ning, Zheng Bing-Bing. A review of human action recognition based on deep learning. *Acta Automatica Sinica*, 2016, **42**(6): 848–857
(朱煜, 赵江坤, 王逸宁, 郑兵兵. 基于深度学习的人体行为识别算法综述. *自动化学报*, 2016, **42**(6): 848–857)
- 38 Zhou X W, Zhu M L, Pavlakos G, Leonardos S, Derpanis K G, Daniilidis K. MonoCap: monocular human motion capture using a CNN coupled with a geometric prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, DOI: 10.1109/TPAMI.2018.2816031
- 39 Zhou X W, Zhu M L, Leonardos S, Derpanis K G, Daniilidis K. Sparseness meets deepness: 3D human pose estimation from monocular video. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016. 4966–4975



洪金华 宁波大学信息科学与工程学院计算机应用技术硕士研究生. 主要研究方向为机器学习, 计算机视觉与模式识别.

E-mail: 18892627653@163.com

(**HONG Jin-Hua** Master student at the Faculty of Electrical Engineering and Computer Science, Ningbo University. His research interest covers machine learning, computer vision and pattern recognition.)



张荣 宁波大学信息科学与工程学院副教授. 主要研究方向为计算机视觉, 数字取证与信息安全.

E-mail: zhangrong@nbu.edu.cn

(**ZHANG Rong** Associate professor at the Faculty of Electrical Engineering and Computer Science, Ningbo University. Her research interest covers computer vision, digital forensics and information security.)



郭立君 宁波大学信息科学与工程学院教授. 主要研究方向为机器学习, 计算机视觉与模式识别. 本文通信作者.

E-mail: guolijun@nbu.edu.cn

(**GUO Li-Jun** Professor at the Faculty of Electrical Engineering and Computer Science, Ningbo University. His research interest covers machine learning, computer vision and pattern recognition. Corresponding author of this paper.)