

基于加速度平滑约束的非刚体三维重建研究

汪亚明,张 静,郑俊褒

(浙江理工大学信息学院,杭州 310018)

摘 要: 针对单一投影误差重建精度不高的问题,提出了基于加速度残差最小的轨迹平滑约束的非刚体三维重建方法。在运动轨迹连续性理论基础上,通过视频序列相邻帧的连续性进行分析,在三维重建投影误差基础上同时引入加速度平滑约束并建立目标函数,最后推导出闭式最优解析解。由于非刚性的运动复杂多变,提出的加速度连续性约束是对非刚体运动目标轨迹的物理特性的本质约束,更具有一般性和适应性。与四种典型的运动模型进行实验对比,结果表明该重建方法有效提高了重建的精确度,且实现简单、算法运算复杂度更小。

关键词: 非刚体;三维重建;加速度;平滑约束;解析解

中图分类号: TP391.4

文献标志码: A

文章编号: 1673-3851(2017)06-0830-08

0 引 言

从二维图像序列中恢复出目标物体在真实世界场景中的三维结构和相机运动参数,一直以来都是计算机视觉和模式识别领域的一个热点和难点^[1],非刚体的三维重建(non-rigid structure from motion, NRSFM)无论是在科学研究还是在工业应用上都引起了极大的关注^[2]。

在三维重建领域,通常有两种主要的方法,形状基方法和轨迹基方法。自 Bregler 等^[3]第一次提出将因式分解方法应用于形变的非刚体目标,主要思想是将非刚体的形变过程表示成一系列基的线性组合,所提出的低秩形状模型对从刚体场景过渡到非刚体重建有着巨大的意义。然而不同于刚体的重建,非刚体的运动有其固有的自由度和模糊性以至于在实际重建中比较困难。Xiao 等^[4]认为如果仅仅在相机正交约束条件下是不足以唯一的恢复出非刚体的三维结构和相机的旋转矩阵,建议引入各种先验条件来增加新的约束。Gotardo 和 Martinez^[5]也曾提出在没有任何先验知识的情况下计算 NRSFM 依然是一个困难的欠约束问题。但是由于

运动形式的多样性,使得非刚体运动重建的约束条件很难被确定,所以对于形变的非刚体重建,形状基的方法有着明显的局限性,不可能适合所有的非刚体目标。Akhter 等^[6]对形状基和轨迹基之间的对偶性进行了论证并提出了轨迹基模型,该模型的主要优点是重建所使用的轨迹基可以被提前预定义,常用的一般为离散余弦变换基(DCT)或沃什哈达码变换基(WHT),从而减少了重建过程中的未知数,也提高了计算过程中相应的稳定性。然而,对于不同的模型使用的轨迹基并不是越多越好,Wang 等^[7]所提的稀疏逼近方法中就对轨迹基的种类和个数进行了讨论。预定义轨迹基方法虽然有着不可替代的优势,但基的固有模糊性仍然不可忽视。Tao 等^[8]提出了一种流行学习算法,利用流行森林^[9]学习出来的形状基进行形状更新,由非线性最小化二维投影误差迭代出每一帧的相机运动参数和非刚体的结构矩阵,开辟了一个新的思路。非刚体的三维重建在大量学者的研究下取得了极大地进步,但是相比于刚体的研究,其在重建的精确度和鲁棒性以及复杂形变的重建、多目标非刚体的重建依然存在着极大的挑战。而且上述的很多方法都是基于

收稿日期:2017-06-27 网络出版日期:2017-10-10

基金项目:浙江省自然科学基金重点项目(LZ15F020004);浙江理工大学 521 人才资助计划

作者简介:张 静(1991-),女,湖北麻城人,硕士研究生,主要从事图像处理、计算机视觉和模式识别等方面的研究。

通信作者:郑俊褒,E-mail:zhengjunbao@zstu.edu.cn

Bergler 等^[3]的假设,这些方法可以将图像序列的次序任意颠倒,没有考虑非刚性运动的连续性,求解过程复杂,对噪声也比较敏感。

本文受运动轨迹的物理平滑性思想的启发,提出了一个简单有效的重建方法。首先分析了高速拍摄的视频序列的特征,发现视频序列是随着帧数的变化而连续变化的,因此每个特征点的运动轨迹可以表示为一条平滑的三维曲线;接着以加速度作为平滑性能指标,建立了基于单一投影误差和平滑约束的目标函数,并推导出最优闭式解析解,进而求解出非刚体结构矩阵,完成非刚体的三维重建。

1 非刚体三维重建结构模型分析

由因式分解理论可知在正交相机投影模型下,假设有 F 帧图像序列,每一帧有 P 个特征点,非刚体三维重建的任务就是通过分解一个二维图像观测矩阵 $\mathbf{W} \in \mathbf{R}^{2F \times P}$ 来获得相机运动矩阵 $\mathbf{M} \in \mathbf{R}^{2F \times 3F}$ 和非刚体结构矩阵 $\mathbf{S} \in \mathbf{R}^{3F \times P}$ 。观测矩阵是由非刚体的结构矩阵通过相机旋转矩阵的映射得到的,其映射关系如式(1)所示:

$$\begin{aligned} \mathbf{W} &= \begin{bmatrix} \mathbf{X}_{11} & \cdots & \mathbf{X}_{1P} \\ \vdots & & \vdots \\ \mathbf{X}_{F1} & \cdots & \mathbf{X}_{FP} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M}_1 & & \\ & \ddots & \\ & & \mathbf{M}_F \end{bmatrix} \begin{bmatrix} \mathbf{S}_{11} & \cdots & \mathbf{S}_{1P} \\ \vdots & & \vdots \\ \mathbf{S}_{F1} & \cdots & \mathbf{S}_{FP} \end{bmatrix} \\ &= \mathbf{MS} \end{aligned} \quad (1)$$

其中: \mathbf{W} 为二维观测矩阵; $\mathbf{X}_{ij} = [x_{ij}, y_{ij}]^T$ 为非刚体的二维坐标, x_{ij} 代表第 i 帧中第 j 个特征点的 x 方向坐标, y_{ij} 代表第 i 帧中第 j 个特征点的 y 方向坐标; $\mathbf{S}_{ij} = [x_{ij}, y_{ij}, z_{ij}]^T$ 为非刚体的三维坐标, z_{ij} 代表第 i 帧中第 j 个特征点的 z 方向坐标; \mathbf{M}_i 表示第 i 帧时的相机旋转矩阵。为了不失一般性,假设在正交投影模型下的观测矩阵中的所有特征点二维坐标都已经过中心化处理。

2 恢复相机的旋转矩阵

由非刚体重建模型分析可知,重建任务是仅基于二维观测矩阵 \mathbf{W} 来恢复出相机旋转矩阵 \mathbf{M} 和非刚体结构矩阵 \mathbf{S} 。由于非刚体的形状和运动都在随着时间变化使得重建问题是一个欠约束问题,未知数的数量 $(3F + 3FP)$ 要远远大于已知的观测矩阵中的数据 $(2FP)$ 。在轨迹基模型下,非刚体特征点的运动轨迹可以近似为一系列轨迹基和相应的轨迹

系数的线性组合,如式(2)所示:

$$\mathbf{S}_{3F \times P} = \mathbf{\Theta}_{3F \times 3K} \mathbf{A}_{3K \times P} \quad (2)$$

$$\mathbf{\Theta} = \begin{bmatrix} \boldsymbol{\theta}_{11} \mathbf{I} & \cdots & \boldsymbol{\theta}_{1K} \mathbf{I} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{F1} \mathbf{I} & \cdots & \boldsymbol{\theta}_{FK} \mathbf{I} \end{bmatrix} \quad (3)$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1(1) & \cdots & \mathbf{a}_1(P) \\ \vdots & \ddots & \vdots \\ \mathbf{a}_K(1) & \cdots & \mathbf{a}_K(P) \end{bmatrix} \quad (4)$$

其中: $\mathbf{\Theta}$ 表示预定义的轨迹基; \mathbf{A} 表示相对应的轨迹系数; \mathbf{I} 是一个 3×3 的单位矩阵; $\mathbf{a}_{ij} = [a_{ix}(j), a_{iy}(j), a_{iz}(j)]^T$ 。由非刚体重建模型可知,测量观测矩阵 \mathbf{W} 通过奇异值分解表示为: $\mathbf{W} = \mathbf{MS} = \mathbf{M}\mathbf{\Theta}\mathbf{A} = \mathbf{U}\mathbf{A}$, $\mathbf{U} = \mathbf{M}\mathbf{\Theta}$ 是一个 $2F \times 3K$ 的矩阵, K 为轨迹基的个数,

$$\mathbf{U} = \begin{bmatrix} \boldsymbol{\theta}_{11} \mathbf{M}_1 & \cdots & \boldsymbol{\theta}_{1K} \mathbf{M}_1 \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}_{F1} \mathbf{M}_F & \cdots & \boldsymbol{\theta}_{FK} \mathbf{M}_F \end{bmatrix} \quad (5)$$

然而由于奇异值分解固有的模糊性, \mathbf{W} 还可以分解为 $\mathbf{W} = \tilde{\mathbf{U}}\tilde{\mathbf{A}}$, 且 $\mathbf{U} \neq \tilde{\mathbf{U}}, \mathbf{A} \neq \tilde{\mathbf{A}}$, 任何一个非奇异正交矩阵 $\mathbf{Q} \in \mathbf{R}^{3K \times 3K}$ 都可以存在于 $\tilde{\mathbf{U}}$ 和 $\tilde{\mathbf{A}}$ 之间得到一个新的有效分解: $\mathbf{W} = \mathbf{U}\mathbf{A} = \tilde{\mathbf{U}}\mathbf{Q}\mathbf{Q}^{-1}\tilde{\mathbf{A}}$ 。为了消除奇异值分解(SVD)不唯一分解的固有模糊性得到矫正矩阵 \mathbf{Q} , 我们只需要估计 \mathbf{Q} 的三列。记 \mathbf{Q}_k 表示 \mathbf{Q} 矩阵的第 1 列、第 $K+1$ 列、第 $2K+1$ 列, 有:

$$\tilde{\mathbf{U}}\mathbf{Q}_k = \begin{bmatrix} \boldsymbol{\theta}_{11} \mathbf{M}_1 \\ \cdots \\ \boldsymbol{\theta}_{F1} \mathbf{M}_F \end{bmatrix} \quad (6)$$

由于旋转矩阵满足正交约束, 即:

$$\begin{aligned} \tilde{\mathbf{U}}_{2i-1:2i} \mathbf{Q}_k (\tilde{\mathbf{U}}_{2i-1:2i} \mathbf{Q}_k)^T &= \tilde{\mathbf{U}}_{2i-1:2i} \mathbf{Q}_k \mathbf{Q}_k^T \tilde{\mathbf{U}}_{2i-1:2i}^T = \\ &= \boldsymbol{\theta}_{i,1}^2 \mathbf{I}_{2 \times 2}, i=1, 2, \cdots, F \end{aligned} \quad (7)$$

其中: $\tilde{\mathbf{U}}_{2i-1}$ 表示矩阵 $\tilde{\mathbf{U}}$ 中 $2i-1$ 和 $2i$ 连续的两行。式(7)可以提供如下四个约束:

$$\tilde{\mathbf{U}}_{2i-1} \mathbf{Q}_k \mathbf{Q}_k^T \tilde{\mathbf{U}}_{2i-1}^T = \boldsymbol{\theta}_{i,1}^2 \quad (8)$$

$$\tilde{\mathbf{U}}_{2i} \mathbf{Q}_k \mathbf{Q}_k^T \tilde{\mathbf{U}}_{2i}^T = \boldsymbol{\theta}_{i,1}^2 \quad (9)$$

$$\tilde{\mathbf{U}}_{2i-1} \mathbf{Q}_k \mathbf{Q}_k^T \tilde{\mathbf{U}}_{2i}^T = 0 \quad (10)$$

$$\tilde{\mathbf{U}}_{2i} \mathbf{Q}_k \mathbf{Q}_k^T \tilde{\mathbf{U}}_{2i-1}^T = 0 \quad (11)$$

\mathbf{Q}_k 是矫正矩阵 \mathbf{Q} 的三列, 记 $\mathbf{G} = \mathbf{Q}_k \mathbf{Q}_k^T$ 为格拉姆矩阵且 \mathbf{G} 是半正定的, 即 $\mathbf{G} \geq 0$ 。此外 $\text{rank}(\mathbf{G}) = \text{rank}(\mathbf{Q}_k) = 3$, 而秩函数本身的数值稳定性不是很好并且在一般情况下矩阵秩的最小化是一个 NP 难问题^[10]。因此进一步放宽为 \mathbf{G} 的核范数^[11]最小化问题 $\min \|\mathbf{G}\|_*$, 而 $\mathbf{G} = \mathbf{Q}_k \mathbf{Q}_k^T$ 是一个对称正定阵, 所以 \mathbf{G} 的核范数就等于它的迹, 即 $\|\mathbf{G}\|_* = \text{trace}(\mathbf{G})$ 。

上述格拉姆矩阵 \mathbf{G} 的最小化问题是一个标准的半正定规划问题^[12] (SDP), 可以选择半正定规划工具包求解。一旦格拉姆矩阵 \mathbf{G} 被求出, 利用 Cholesky 分解就可以求出矫正矩阵 \mathbf{Q} , 进而恢复出相机的旋转矩阵 \mathbf{M} 。

3 基于加速度平滑约束(Smooth)的非刚体三维结构重建

非刚体三维重建研究的最终目的是恢复出非刚体运动目标的三维结构。由于非刚性的运动复杂多变, 在基于形状基方法重建时都需要对图像序列中的每一帧进行估计新的形状基, 而且对于不同的运动也需要重新定义不同的形状基来进行三维结构的线性组合, 极大地增加了计算量, 导致重建效率和精度都比较低。由轨迹空间理论可知, 可以把非刚体每个特征点的重建转化为随着时间变化的每一帧的重建, 同时从物理学角度, 结合非刚体在高速摄像机拍摄的图像序列下运动的物理连续性^[13], 相机拍摄速度很高的时候相邻帧特征点位移变化很小, 速度和加速度的变化也很小。通过这样的物理连续性在单一投影误差的基础上构造目标函数, 以相邻连续帧的加速度平方差积分的最小值作为平滑性能指标, 引入一个平滑矩阵将运动物理性约束转化为重建约束项。

假设提取特征点的图像序列的时间间隔为 t , 每一帧的图像采集时间间隔是个常数, 则根据物理运动定律:

$$v_i = \frac{\Delta S}{t} = \frac{S_{i+1} - S_i}{t} \quad (12)$$

$$a_i = \frac{v_{i+1} - v_i}{t} \quad (13)$$

一段时间内物体的位移差与时间 t 的比值反映这段时间内物体运动的平均速度, 平均速度之差反映的是加速度变化, 对于一个拍摄的图像序列来说这样的平均速度变化和加速度变化都是非常小的。将 F 帧连续累加起来为:

$$\min \sum_{i=2}^{F-1} (\mathbf{S}_{i+1} + \mathbf{S}_{i-1} - 2\mathbf{S}_i) \quad (14)$$

引入平滑约束后, 非刚体轨迹三维重建的目标函数为:

$$\min \|\mathbf{W} - \mathbf{M}\mathbf{S}\|_F^2 + \lambda_1 \sum_{i=2}^{F-1} (\mathbf{S}_{i+1} + \mathbf{S}_{i-1} - 2\mathbf{S}_i) \quad (15)$$

定义平滑矩阵 $\mathbf{V} \in \mathbf{R}^{3F \times 3F}$ 为:

$$\mathbf{V} = \begin{cases} 1, & j=i, j=i+6, \quad i=1, 2, \dots, 3(F-2) \\ -2, & j=i+3, \quad i=1, 2, \dots, 3(F-2) \\ 0, & \text{其它} \end{cases} \quad (16)$$

即式(15)转化为:

$$\min \frac{1}{2} \|\mathbf{W} - \mathbf{M}\mathbf{S}\|_F^2 + \frac{\lambda_1}{2} \|\mathbf{V}\mathbf{S}\|_F^2 \quad (17)$$

其中 $\lambda_1 \geq 0$ 为约束项的平衡参数, 反映了重建模型拟合能力与轨迹平滑性之间的平衡。由式(17)可知, 目标函数式(17)的无约束最优化模型为二次函数, 具有全局最优解析解。首先计算式(17)关于非刚体目标三维轨迹结构矩阵 \mathbf{S} 的梯度:

$$\frac{\partial f(\mathbf{S})}{\partial \mathbf{S}} = -\mathbf{M}^T (\mathbf{W} - \mathbf{M}\mathbf{S}) + \lambda_1 \mathbf{V}^T \mathbf{V} \mathbf{S} \quad (18)$$

令式(18)为 0 得到:

$$(\mathbf{M}^T \mathbf{M} + \lambda_1 \mathbf{V}^T \mathbf{V}) \mathbf{S} = \mathbf{M}^T \mathbf{W} \quad (19)$$

从而得到解析解:

$$\mathbf{S} = (\mathbf{M}^T \mathbf{M} + \lambda_1 \mathbf{V}^T \mathbf{V})^+ \mathbf{M}^T \mathbf{W} \quad (20)$$

其中 $(\mathbf{M}^T + \lambda_1 \mathbf{V}^T \mathbf{V})^+$ 表示矩阵 $(\mathbf{M}^T + \lambda_1 \mathbf{V}^T \mathbf{V})$ 的 Moore-Penrose 逆, 当矩阵 $(\mathbf{M}^T + \lambda_1 \mathbf{V}^T \mathbf{V})$ 可逆时, 式(17)为精确解, 当矩阵 $(\mathbf{M}^T + \lambda_1 \mathbf{V}^T \mathbf{V})$ 不可逆时, 式(20)给出的 Moore-Penrose 逆解给出式(19)的极小范数的最小二乘解。而在实际中, 矩阵 $(\mathbf{M}^T + \lambda_1 \mathbf{V}^T \mathbf{V})$ 一般为可逆矩阵且 $\lambda_1 \geq 0$, 这个矩阵为半正定矩阵, 显然存在极小值, 即式(20)为目标函数的解析解。

4 实验与分析

4.1 重建评估准则

对非刚体进行三维重建, 需要定义重建效果误差评估标准来判断重建的优劣, 也即是重建后的结构矩阵与真实的三维结构之间的误差, 其评估公式定义 e_{3D} ^[14] 为如式(21):

$$e_{3D} = \frac{1}{\sigma F P} \sum_{t=1}^F \sum_{j=1}^P e_{tj}, \sigma = \frac{1}{3F} \sum_{t=1}^F (\sigma_{tx} + \sigma_{ty} + \sigma_{tz}), \quad (21)$$

$$e_{tj} = \sqrt{L_{tjx}^2 + L_{tjy}^2 + L_{tjz}^2}$$

其中 $L_{tjx} = \|\mathbf{S}_r(3t-2, j) - \mathbf{S}_0(3t-2, j)\|$; $L_{tjy} = \|\mathbf{S}_r(3t-1, j) - \mathbf{S}_0(3t-1, j)\|$; $L_{tjz} = \|\mathbf{S}_r(3t, j) - \mathbf{S}_0(3t, j)\|$; $t=1, 2, \dots, F, j=1, 2, \dots, P, \sigma_{tx}, \sigma_{ty}, \sigma_{tz}$ 分别为三维结构中第 t 帧对应的所有结构点 x, y, z 方向坐标的标准差; e_{tj} 表示第 t 帧第 j 个三维结构点的误差, 即重构点与实际点间的欧氏距离; \mathbf{S}_r 代表重建出来的三维结构矩阵, \mathbf{S}_0 代表实际的空间结构矩阵, 二者的维度和排列方式完全相同。三维点结构

误差作为非刚体三维重建的评价标准具有代表性,它反映的是算法恢复的平均误差。而从轨迹空间的角度分析,重建结果不仅要在全局进行考虑,而且还应从时间和特征点的角度进行分析。依据上面三维点结构误差的定义,得到三维点的帧平均误差 e_{3D-F} 和点平均误差 e_{3D-P} 分别为式(22)—(23):

$$e_{3D-F} = \frac{1}{\sigma P} \sum_{i=1}^P e_{ij}, \sigma = \frac{1}{3F} \sum_{i=1}^F (\sigma_{tx} + \sigma_{ty} + \sigma_{tz}) \quad (22)$$

$$e_{3D-P} = \frac{1}{\sigma F} \sum_{i=1}^P e_{ij}, \sigma = \frac{1}{3F} \sum_{i=1}^F (\sigma_{tx} + \sigma_{ty} + \sigma_{tz}) \quad (23)$$

4.2 实验方法

本文所采用的实验数据来源于卡耐基梅隆大学的运动捕捉数据(<http://cvlab/lums.edu.pk/nrsfm>),选取了四种典型的人体运动,如表 1 所示。

表 1 4 种典型的非刚体运动模型

运动模型	帧数/帧	特征点数/个
Pickup	357	41
Yoga	307	41
Dance	264	75
Stretch	370	41

为了验证本文方法的有效性和可靠性,通过编程实验对上述 4 种非刚性运动进行三维重建。实验所用硬件配置为处理器 Inter(R)@3.40 GHz,内存 4 GB;软件环境为 Window7, Matlab2010b。通过编程实验恢复出人体三维结构矩阵,然后利用式(21)—(23)计算重建误差,并绘制不同方法下三维重建效果图来对本文方法进行评估和有效性分析。在实验中,将本文方法与其他 4 种研究算法进行比较,对比的算法包括:EM 方法^[14]、PTA 方法^[14](选用 DCT 基)和 Sparse 方法^[7],且以上 4 种方法所用实验数据与本文方法实验数据来源于同一网站。

4.3 实验结果与分析

4.3.1 Pickup 运动模型重建结果分析

首先从视觉角度分析,采用 PTA 方法和平滑约束方法(Smooth)对 Pickup 运动模型进行重建,随机抽取该图像序列的第 30 帧、第 120 帧、第 250 帧和第 300 帧的重建效果与 PTA 方法(选用 DCT 基且选取最优的 $K=12$)的重建效果进行比较。从正视、侧视、俯视三个角度更加全面的观察重建效果。该效果图主要通过观察真实三维点与重构三维点之间的距离的吻合情况来验证算法的有效性,真实点与重建点之间越吻合说明该算法的重建效果越

好。其中“•”表示运动目标的真实三维点,“◦”表示 PTA 方法和本文方法重建出的非刚体三维坐标点。其三视图如图 1 所示。

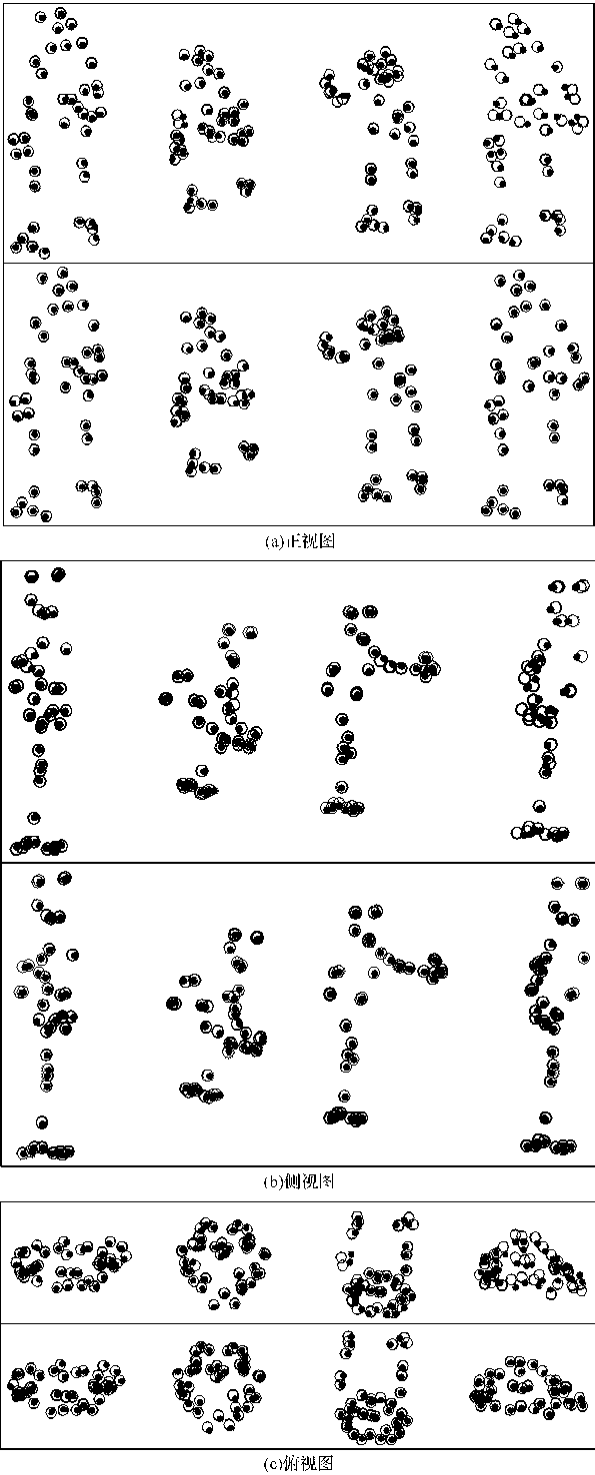


图 1 Pickup 运动三视图

注:子图中第一行为 PTA 方法重建结果,第二行为 Smooth 方法重建结果。

从图 1(a)—(c)可以看出,无论是从正视、侧视和俯视角度观察,Smooth 方法下的三维重建效果都

是优于 PTA 方法下的重建效果。Pickup 运动的运动幅度相对比较大,手臂、腿部和头部等多部位均有摆动,在用 PTA 轨迹基方法进行重建时,不仅最优轨迹基的个数很难确定而且该方法也没有考虑非刚性运动的物理平滑性,因而重建效果并不理想。而本文方法考虑运动平滑性的本质约束,适用性强,对

Pickup 这种运动比较复杂的序列也有着比较好的重建效果,验证了本文方法的适用性。

仅仅从视觉效果图上判断本文方法的优势仍不明显,下面绘制出 Pickup 运动在 PTA 方法($K=12$)和 Smooth 方法下的重构帧均误差曲线图和点均误差曲线图,如图 2(a)、(b)所示。

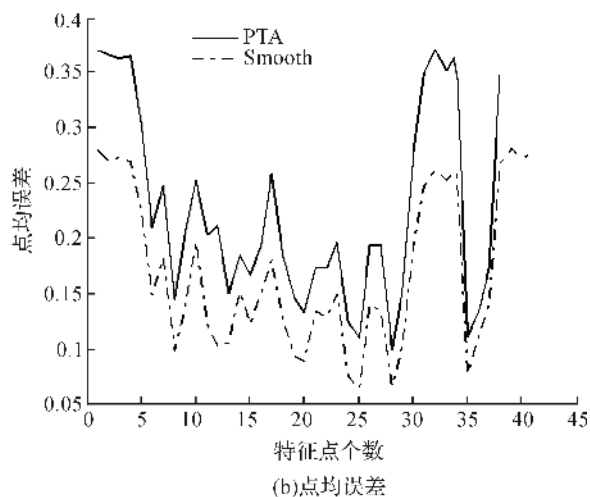
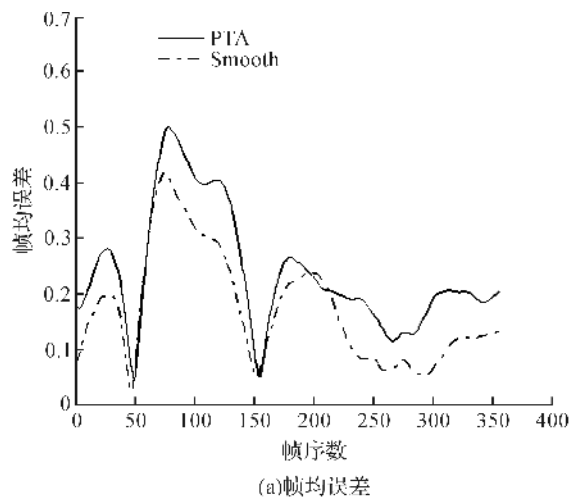


图2 Pickup 运动的帧均误差和点均误差曲线对比图

从图 2(a)—(b)可以看出,随着帧数的变化本文方法下的帧均误差和点均误差基本都在 PTA 方法的误差曲线之下,表明本文方法的重建效果是明显优于 PTA 方法。

为了更加细致地说明这种差异,图 3 画出了 PTA 方法和本文方法在对 Pickup 运动中随机选取的第 3 个和第 32 个特征点进行重建后的结构误差曲线图。

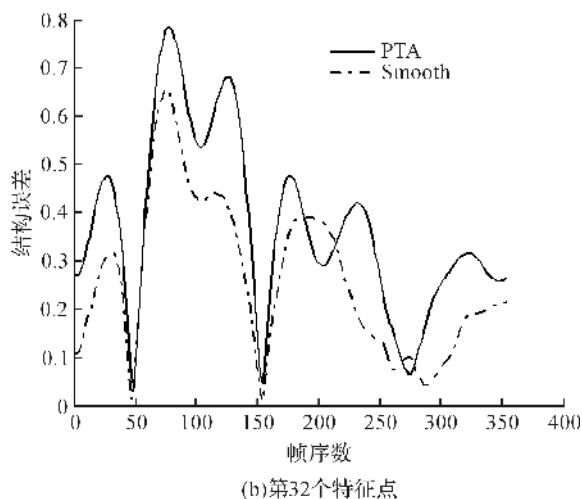
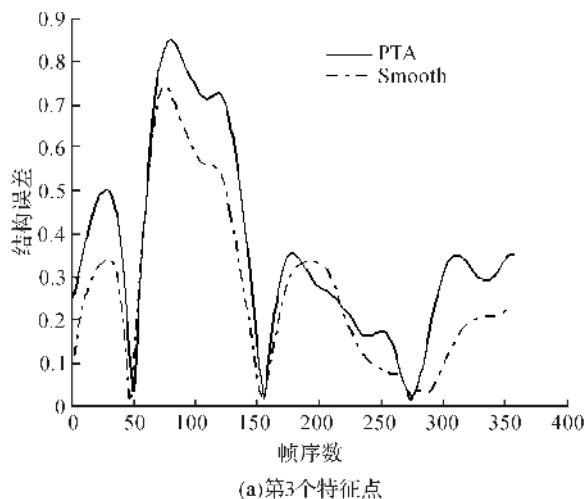


图3 第3个特征点和第32个特征点的结构误差曲线对比

通过曲线可以看出,在对单个特征点的运动重建进行分析时,Smooth 方法下的重建误差在绝大部分帧数下都是小于 PTA 方法的,表明本文方法能够更好的恢复出非刚体的运动结构。而 PTA 方法重建时不仅跟选择的轨迹基的关系很紧密,有很大的局限性,而且直接通过伪逆法求解轨迹系数后得到结构矩阵,也没有考虑运动的平滑性,其重建精度

比较低。

上面的实验结果,说明 Smooth 方法相对于 PTA 方法在对 Pickup 运动进行三维运动重建时有着较大的优势。

4.3.2 Yoga 序列重建的结果和分析

为了进一步验证本文方法的适用性和有效性,对 Yoga 运动模型使用 Sparse 方法和 Smooth 方法

进行三维运动重建。随机选取 Yoga 运动序列中的第 55 帧、第 120 帧、第 170 帧和第 230 帧,图 4 为两

种方法下的重建视觉效果图。

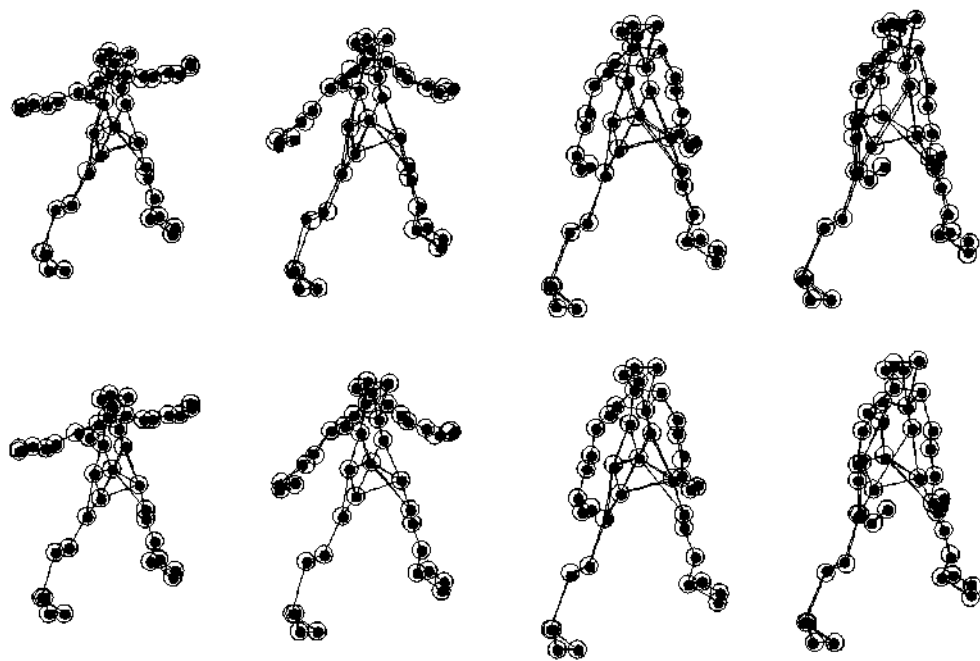


图 4 Yoga 序列重建效果

注:第一行为 Sparse 方法,第二行为 Smooth 方法。

从图 4 中可以看出,相对于 Sparse 方法,本文方法下 Yoga 运动的头部、手臂和腿部的重建效果均有不同程度的提高。Sparse 方法可以自动的选择合适的轨迹基对重建目标运动轨迹进行重建,但 Sparse 方法进行重建时所用的基追踪算法相对复杂,而且没有

考虑到矩阵的稀疏表示会破坏数据的内在连续性,因此它的重建效果要略差于 Smooth 方法的重建效果的。

为了进一步说明这种差异,yoga 模型在 Sparse 方法和 Smooth 方法下的帧均误差和点均误差图,如图 5 所示。

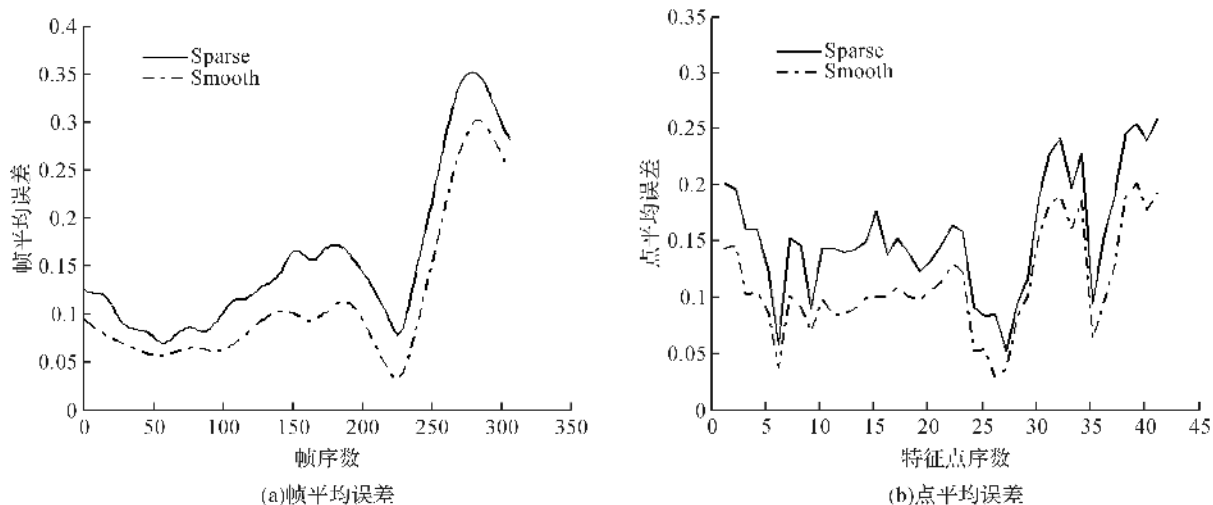


图 5 Yoga 运动的帧均误差和点均误差曲线对比

从图 5 中可以明显看出在 Yoga 模型下,Smooth 方法下的帧平均误差和点平均误差曲线都在 Sparse 方法重建误差的曲线之下,表明 Smooth 方法更好的提高了 Yoga 运动的三维重建精度。

同样,为了更加详细并且直观的看出针对每个特征点本文方法的有效性,图 6 给出了随机选取的第 13 个特征点和第 40 个特征点在随着运动序列帧数变化的结构误差对比图。

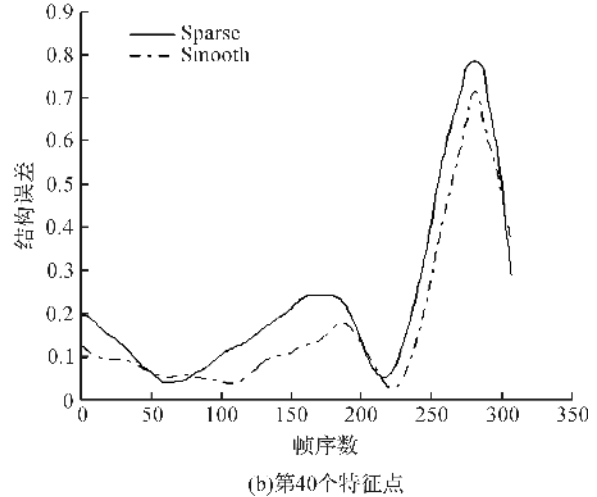
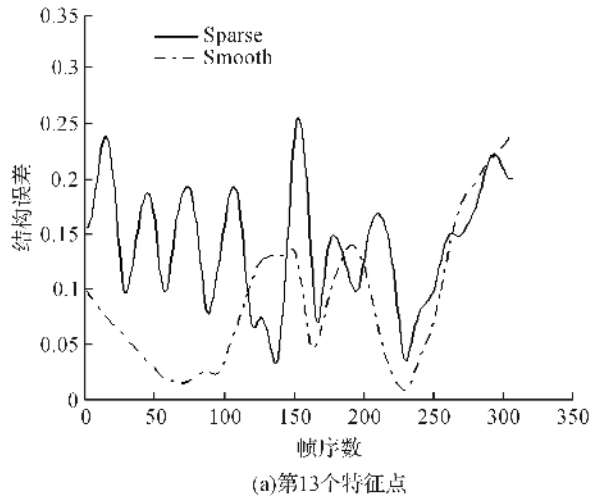


图6 第13个特征点和第40个特征点的结构误差曲线对比

由图6可知,随着帧数的变化,第13个特征点和第40个特征点的结构误差在绝大部分帧下都是减小的,整体的重建精度仍然是有着明显的提高。

4.3.3 不同运动模型在不同方法下重建的结果与分析

为了验证本文算法对不同运动序列重建都具有一定程度上的优势,图7绘制了四种不同的运动分别在MP方法、PTA方法(括号中的数字为该重建算法下重建效果最优时的轨迹基个数)、Sparse方法和Smooth方法下的重建误差对比柱状图,表2给出了重构误差对比表,表中数据越小说明其重建误差越小,重建精度越高。

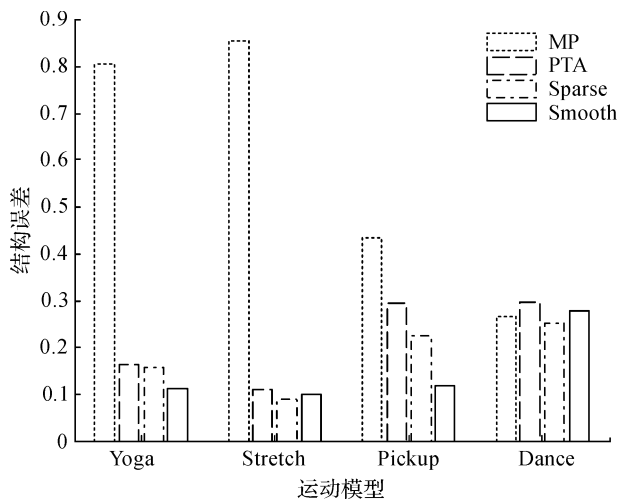


图7 不同运动模型在4种不同方法下的重构误差柱状图

表2 4种典型运动在不同方法下的重构误差对比表

实验模型	MP	PTA	Sparse	Smooth
Yoga	0.8039	0.1622(11)	0.1558	0.1119
Pickup	0.4322	0.2369(12)	0.2245	0.1710
Dance	0.2639	0.2958(5)	0.2519	0.2788
Stretch	0.8549	0.1088(12)	0.0893	0.0976

由图7、表2可知,本文方法下的四组非刚体运动的重建中,Yoga运动、Pickup运动在Smooth方法下的重建误差均小于前面的几种重建方法,重建精度有着明显的提高。Dance运动和Stretch运动下本文方法结果虽然优于PTA方法,但重建精度略微高于Sparse方法,表明Smooth方法的鲁棒性欠佳,还有待进一步的提高。

5 结论

本文方法中首先利用正交约束和迹约束求解出相机旋转矩阵,然后在单一投影误差重建基础上引入非刚体运动轨迹加速度的连续平滑性约束,通过对目标函数求导得到了其最优解析解,求解出非刚体结构矩阵。该方法的重建过程减小了误差积累,并且与其他约束相比,运动目标的平滑连续性约束更具有一般性和通用性,能够更好的拟合运动轨迹。最后由选取的四种不同的非刚体运动模型进行试验,并具体以Pickup运动和Yoga运动为例进行了分析,实验结果表明本文方法有效的提升了非刚体三维结构的重建精度。

对于非刚体的三维重建,在进一步的研究中可以考虑重建过程鲁棒性的提升、角速度约束等因素,也可以尝试从单个非刚体目标的重建扩展到多个非刚体目标的重建上。

参考文献:

- [1] ZHANG Y, GIBSON G M, HAY R, et al. A fast 3D reconstruction system with a low-cost camera accessory [J]. Scientific Reports, 2015, 5: 10909.
- [2] ALLDIECK T, KASSUBECK M, MAGNOR M. Optical flow-based 3D human motion estimation from monocular

- video[EB/OL]. (2017-03-01)[2017-06-27]. <https://arxiv.org/pdf/1703.00177>.
- [3] BREGLER C, HERTZMANN A, BIEERMANN H. Recovering non-rigid 3D shape from image streams[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Los Alamitos, IEEE, 2000:690-696.
- [4] XIAO J, CHAI J, KANADE T. A closed-form solution to non-rigid shape and motion recovery[J]. International Journal of Computer Vision, 2004, 67(2):573-587.
- [5] GOTARDO P F U, MARTINEZ A M. Kernel non-rigid structure from motion[C]//Proceedings of the IEEE International Conference on Computer Vision, IEEE, 2011:802-809.
- [6] AKHTER I, SIMON T, KHAN S, et al. Bilinear spatiotemporal basis models[J]. ACM Transactions on Graphics, 2012, 31(2):1-12.
- [7] WANG Y, YAN X, ZHENG J, et al. Sparse approximation for nonrigid structure from motion[J]. Journal of Robotics, 2015(3):1-8.
- [8] TAO L, MATUSZEWSKI B J. Robust deformable shape reconstruction from monocular video with manifold forests[J]. Machine Vision and Applications, 2016, 27(6):1-19.
- [9] TAO L, MATUSZEWSKI B J. Deformable shape reconstruction from monocular video with manifold forests[J]. Computer Analysis of Images and Patterns, Springer Berlin Heidelberg, 2013:28-36.
- [10] ZHOU Z, SHI F, XIAO J, et al. Non-rigid structure-from-motion on degenerate deformations with low-rank shape deformation model[J]. IEEE Transactions on Multimedia, 2015, 17(2):171-185.
- [11] RECHT B, FAZEL M, PARRILO P. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization[EB/OL]. SIAM review, 2010, 52(3):471-501.
- [12] DAI Y C, LI H D, HE M Y. A simple prior-free method for non-rigid structure-from-motion factorization[J]. International Journal of Computer Vision, 2014, 107(2):101-122.
- [13] 刘松国, 朱世强, 吴文祥. 具有运动时间约束的机械手最优平滑轨迹规划[J]. 电机与控制学报, 2009, 13(6):897-902.
- [14] WANG Y, TONG L, JIANG M. Non-rigid structure estimation in trajectory space from monocular vision[J]. Sensors, 2015, 15(10):25730-25745.

Research on 3D Non-Rigid Reconstruction Based on Acceleration Smoothness Constraints

WANG Yaming, ZHANG Jing, ZHENG Junbao

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: To overcome the defect that the reconstruction method of single projection error is of low precision, a 3D non-rigid reconstruction method of trajectory smoothness constraints based on minimum acceleration residue is proposed. According to the theory of movement trajectory continuity, acceleration smoothness constraints are introduced based on the 3D reconstruction projection error to build an objective function, to obtain the optimal closed analytical solution. Due to the complexity of the non-rigid motion, the proposed acceleration continuity constraints are the constraints on the physical properties of non-rigid motion target trajectories, which are of universality and adaptability. The comparison with four typical motion models show that this reconstruction method improves the accuracy of the reconstruction largely, and realizes algorithm in a simple way.

Key words: non-rigid; 3D reconstruction; acceleration; smoothness constraints; analytical solution

(责任编辑: 康 锋)