论 文

单目视频运动目标轨迹三维重建的平滑约束法

戴玉超^{①*}, 何明一^{①*}, LI HongDong^②, 李波^①^① 西北工业大学电子信息学院, 陕西省信息获取与处理重点实验室, 西安 710129^② Research School of Engineering, Australian National University, Canberra ACT 2600, Australia

* 通信作者. E-mail: daiyuchao@gmail.com, myhe@nwpu.edu.cn

收稿日期: 2012-04-19; 接受日期: 2012-05-13

国家自然科学基金 (批准号: 60736007, 61171154) 和陕西省自然科学基金 (批准号: 2010JZ011) 资助项目

摘要 提出了利用单目视频重建运动目标三维轨迹的平滑约束方法. 通过引入平滑约束, 得到重建运动目标三维轨迹的无约束最优化解模型, 进而推导出其闭式最优解. 提出的平滑约束是对运动目标三维轨迹的本质约束, 更具一般性, 可以反映目标运动的连续、渐变与平滑特性, 与基于离散余弦变换基和多项式基等预先定义基约束相比具有更强的适应性, 同时能够直接应用于部分观测数据缺失的情形. 给出了单目视频轨迹重建问题的几何解释, 并对轨迹重建问题中的唯一性进行分析. 仿真和真实单目视频序列上的实验证明了提出方法的有效性和先进性.

关键词 平滑约束 轨迹重建 多视角 运动目标 单目视频 闭式最优解 视线

1 引言

基于视频序列的运动目标轨迹三维重建在目标跟踪、多角度探测、机器人、自然人机交互、动作捕捉、三维电视、月球车巡游和对地观测等方面^[1~5]有着重要应用. 常规的运动目标跟踪与轨迹重建系统由多个精确同步的标定相机构成, 多相机同时获取运动目标不同视角下的多个观测, 利用相机标定信息和多视角观测根据三角原理重建运动目标每一时刻的三维位置^[6], 运动目标的三维轨迹则由目标位置的序列估计组合得到, 如图 1(a) 所示, 图中 c 表示不同位置的静态相机. 在这一多相机运动目标跟踪与轨迹重建框架中, 运动目标跟踪与轨迹重建精度依赖于多相机系统的精确时钟同步与标定, 同时多相机系统成本高且观测视角固定, 对系统的实用性提出挑战. 另外这一框架无法应用于空间非合作目标远距离监测等情形^[3].

本文研究如何从单目视频重建运动目标三维轨迹. 显然, 在没有任何约束的情形下, 利用单目视频只能获取目标相对相机的视线方向而无法获取其距离, 因而无法重建运动目标三维轨迹. 具有运动目标先验信息时, 利用先验信息可以实现从单目视频重建运动目标三维轨迹. 该重建方法可以应用于机动平台多角度测绘、机器人人机交互、月球探测、空间目标监测和对地观测等方面. 在单目视频运动目标轨迹重建问题中, 单相机在不同位置获取运动目标不同时刻不同角度的多个观测, 如图 1(b) 所示, 图中 c 表示不同时刻的同一单目运动相机, 由于观测不满足时钟同步因此不能使用静态重建下的三角原理. 本文研究如何根据单目相机标定信息和多视角 (不同时间、不同角度) 观测重建动态目标的三维轨迹.

引用格式: 戴玉超, 何明一, LI H D, 等. 单目视频运动目标轨迹三维重建的平滑约束法. 中国科学: 信息科学, 2012, 42: 988–1006, doi: 10.1360/112011-84

图 1 多静态相机与单目视频运动目标轨迹重建

Figure 1 Moving target trajectory reconstruction from multiple static cameras and monocular video. (a) Multiple static camera configuration; (b) monocular video configuration

Avidan 等^[7]描述了从单目图像序列重建运动点的“轨迹三角化”问题,提出两种可以重建运动点轨迹的情形:1)目标点沿直线运动;2)目标点沿平面二次曲线运动.其后,Shashua 等^[8]给出重建平面运动目标点的理论.Kaminski 等^[9]采用多项式表示扩展至一般轨迹的重建.Wolf 等^[10]对不同类型的重建问题进行分类,并推广至从 $\mathcal{P}^k(k=3, \dots, 6)$ 到 $\mathcal{P}^2$ 的投影.

Park 等^[1]提出可以表达为基轨迹线性组合的运动目标三维轨迹的重建方法.该方法假设轨迹可以采用离散余弦变换(DCT)基以较少的低频分量表示,从而得到具有最小二乘解的超定系统.Park 等^[11]进一步提出从单目视频重建铰接物体轨迹的方法.Chen 等^[12]研究在 DCT 基描述模型下的线性系统临近欠定时如何进一步引入稀疏先验信息以恢复运动轨迹.于起峰等^[3]假设运动目标轨迹可以用时间的有限次多项式参数化描述,将运动相机观察运动目标的系列视线,与目标的参数化运动轨迹进行交会,从而确定目标的三维轨迹和速度等运动参数.Zhu 等^[13]针对现实世界中相机运动与点运动轨迹高度相关导致可重建性低的问题,提出基于关键帧和应用 L_1 范数施加稀疏性约束的重建方法.

与基于 DCT 基表示和多项式参数化等预先定义基轨迹方法不同,本文引入运动目标三维轨迹的平滑约束先验信息,在约束问题求解的同时没有对于轨迹的过约束,因此不失一般性而且能够以较高精度重建运动目标三维轨迹.进一步推导出平滑约束下的闭式最优解,具有快速性,能够快速实时恢复运动目标三维轨迹.提出的方法不需要运动目标轨迹的基轨迹数量信息,并且可以适用于部分观测缺失的情形.本文进一步讨论了单目视频轨迹重建问题的几何解释,并且分析了轨迹重建问题的典型退化情形以及重建结果的唯一性.仿真和真实单目视频上的实验结果表明了提出方法的有效性和先进性.

2 动态轨迹重建

令 $\mathbf{x}_i = [u_i, v_i]^T$ 表示时刻 i 对于运动目标三维轨迹的图像观测, $\mathbf{P}_i$ 为时刻 i 相机投影矩阵,透视相机模型下 $\mathbf{P}_i = \mathbf{K}_i[\mathbf{R}_i|\mathbf{t}_i] \in \mathbb{R}^{3 \times 4}$, 根据相机成像模型有

$$\lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{P}_i \bar{\mathbf{X}}_i = \mathbf{P}_i^{:,1:3} \mathbf{X}_i + \mathbf{P}_i^{:,4}, \quad (1)$$

其中 $\lambda_i > 0$ 为未知射影深度^[6], $\mathbf{X}_i = [x_i, y_i, z_i]^T$ 为目标三维位置, $\bar{\mathbf{X}}_i = [\mathbf{X}_i^T, 1]^T$ 为目标三维位置的

齐次坐标, $P_i^{:,1:3}$ 为 P_i 的前 3 列而 $P_i^{:,4}$ 为第 4 列. 将未知的射影深度 λ_i 代入方程 (1) 经过整理可得

$$\begin{aligned}(u_i P_i^{3,1:3} - P_i^{1,1:3}) X_i &= P_i^{1,4} - u_i P_i^{3,4}, \\ (v_i P_i^{3,1:3} - P_i^{2,1:3}) X_i &= P_i^{2,4} - v_i P_i^{3,4}.\end{aligned}\quad (2)$$

令

$$Q_i = \begin{bmatrix} u_i P_i^{3,1:3} - P_i^{1,1:3} \\ v_i P_i^{3,1:3} - P_i^{2,1:3} \end{bmatrix}, q_i = \begin{bmatrix} P_i^{1,4} - u_i P_i^{3,4} \\ P_i^{2,4} - v_i P_i^{3,4} \end{bmatrix},$$

则 $Q_i X_i = q_i$. 累积所有图像帧上的观测可以得到如下方程:

$$\begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_F \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_F \end{bmatrix} = \begin{bmatrix} q_1 \\ \vdots \\ q_F \end{bmatrix}. \quad (3)$$

令 $Q = \text{blkdiag}(Q_1, \dots, Q_F) \in \mathbb{R}^{2F \times 3F}$ 表示由 $Q_i, i = 1, \dots, F$ (F 为单目视频序列长度) 组成的对角阵, $X = [X_1^T, \dots, X_F^T]^T$ 为运动目标三维轨迹, $q = [q_1^T, \dots, q_F^T]^T$, 进一步得到如下模型:

$$QX = q. \quad (4)$$

在仿射相机模型下: 令 $W_i = x_i$ 表示时刻 i 对运动目标轨迹的图像观测, $P_i = [R_i | t_i] \in \mathbb{R}^{2 \times 4}$ 表示相机参数矩阵, 则 $W_i = P_i \bar{X}_i = R_i X_i + t_i$, 经过变换可以得到

$$\bar{W}_i = W_i - t_i = R_i X_i. \quad (5)$$

累积所有图像帧下的观测可以得到

$$RX = \bar{W}, \quad (6)$$

其中 $\bar{W} \in \mathbb{R}^{2F \times 1}$ 表示去除平移向量的图像观测向量, $R = \text{blkdiag}(R_1, \dots, R_F) \in \mathbb{R}^{2F \times 3F}$ 为由 $R_i, i = 1, \dots, F$ 组成的对角阵, $X = [X_1^T, \dots, X_F^T]^T \in \mathbb{R}^{3F \times 1}$ 为运动目标三维轨迹.

可见仿射相机模型下的求解模型 (6) 与透视相机模型下的求解模型 (4) 具有相似的形式, 其各项的含义不同. 在不引起误解的情况下, 下文统一采用 $QX = q$ 的形式描述透视相机模型与仿射相机模型下的求解模型.

由于 Q 的行数小于列数, 求解模型 (4) 为欠定线性系统, 其通解形式为 $X = Q^\dagger q + \mathcal{N}(Q)v$, 其中 $\mathcal{N}(Q)$ 表示 Q 的核空间, $Q^\dagger$ 为矩阵 Q 的 Moore-Penrose 伪逆, $v \in \mathbb{R}^{F \times 1}$ 为待定向量, 可见解空间的维数为 $F \times 1$. 易证, $\hat{X} = Q^\dagger q$ 为欠定线性系统 $QX = q$ 的最小 L_2 范数解.

2.1 几何解释

以上从代数角度对单目视频运动目标轨迹三维重建问题进行了描述, 这里我们进一步给出其几何解释. 给定图像观测 $x_i = [u_i, v_i]^T$, 我们首先确定映射至该观测的所有三维点的集合 (即三维视线). 首先, 相机中心点 C_i 和真实目标点位置 X_i 位于该视线上. 透视相机模型下, 相机中心点 C_i 满足 $P_i \bar{C}_i = 0$, 其中 $\bar{C}_i = [C_i^T, 1]^T$ 为齐次坐标形式, 对应的射影深度为 $\lambda_i = 0$. 进一步推导可以得到 $C_i = -(P_i^{:,1:3})^{-1} P_i^{:,4}$. 连接相机中心 C_i 与真实目标点 X_i 的视线可以参数化表达为

$$X_i(\mu_i) = C_i + \mu_i(X_i - C_i) = \mu_i X_i + (1 - \mu_i)C_i, \quad (7)$$

其中 μ_i 为任意实数. 如图 1(b) 所示, 视线描述了从相机中心点观测目标轨迹的方向.

通过引入矩阵形式, 所有满足视线约束的运动目标轨迹 (轨迹重建问题解空间) 可以表达为^[1]

$$\mathbf{X}(\mathbf{A}) = \mathbf{A}\mathbf{X} + (\mathbf{I} - \mathbf{A})\mathbf{C}, \quad (8)$$

其中 $\mathbf{A} = \mathbf{D} \otimes \mathbf{I}_3 \in \mathbb{R}^{3F \times 3F}$, $\mathbf{D}$ 为由 $\mu_1, \mu_2, \dots, \mu_F$ 构成的对角阵, $\otimes$ 表示 Kronecker 乘积, $\mathbf{C} = [\mathbf{C}_1^T, \mathbf{C}_2^T, \dots, \mathbf{C}_F^T]^T$ 为相机运动轨迹. 当所有 $\mu_i = 0$ 时, $\mathbf{X}(\mathbf{A})$ 为相机轨迹 $\mathbf{C}$, 可见相机运动轨迹属于轨迹重建问题的解空间.

2.2 预先定义基轨迹方法

线性系统 $\mathbf{Q}\mathbf{X} = \mathbf{q}$ 的解空间的维数为 $F \times 1$, 因此一般情形下的运动目标轨迹三维重建问题是欠约束的. 为了约束问题, 不同的先验信息被引入. 预先定义基轨迹方法假设运动目标的三维轨迹可以由预先定义的基轨迹的线性组合近似表示^[1~3,12,13], 即 $\mathbf{X} \approx \Theta\beta$, 其中

$$\Theta = \begin{bmatrix} \theta_1^T & & & & & \\ & \theta_1^T & & & & \\ & & \theta_1^T & & & \\ & & & \ddots & & \\ & & & & \theta_F^T & \\ & \theta_F^T & & & & \\ & & \theta_F^T & & & \\ & & & & \theta_F^T & \end{bmatrix} \in \mathbb{R}^{3F \times 3K}, \quad (9)$$

$\theta_i \in \mathbb{R}^K$ 为采样时刻 i 预先定义基的前 K 项, $\beta = [\beta_{x1}, \dots, \beta_{xK}, \beta_{y1}, \dots, \beta_{yK}, \beta_{z1}, \dots, \beta_{zK}]^T \in \mathbb{R}^{3K}$ 为组合系数, K 为基的数量. 利用以上假设可以将轨迹重建模型 (4) 转化为 $\mathbf{Q}\Theta\beta = \mathbf{q}$. 当 $2F \geq 3K$ 时, 系统是超定的, 从而可以通过最小二乘原理求解得到 $\hat{\beta} = (\mathbf{Q}\Theta)^\dagger \mathbf{q}$, $\mathbf{X} = \Theta\hat{\beta}$. 预先定义基轨迹方法可以视为求解轨迹重建问题解空间与预先定义基轨迹 Θ 张成空间的交集.

预先定义基可以选择任意正交基, 如多项式基和 DCT 基等. 其中 DCT 基由于其描述运动轨迹的良好性能得到广泛应用^[1,2,12,13]. 预先定义基轨迹方法依赖于基数量 K 的准确估计. 图 2 给出同一运动目标轨迹重建问题中, 采用 DCT 基方法时不同基数量 K 下重建轨迹 (X 坐标) 与真实轨迹的对比. 从图中可以看出当 K 较小时该方法无法描述复杂的运动轨迹; 当 K 较大时又会产生过拟合现象. 该轨迹重建问题中不同 DCT 基数量 K 下轨迹重建相对误差曲线如图 3 所示, 图中局部采用放大显示. 当估计的基数量接近真实的基数量时, 轨迹重建相对误差较小, 当估计的基数量产生较大偏差时轨迹重建误差急剧增大. 采用多项式基等其他预先定义基描述运动轨迹^[3] 也会产生欠拟合和过拟合问题. 真实的单目视频序列中, 基数量往往难以准确估计, 这给基于 DCT 基和多项式基等预先定义基轨迹方法的实用性带来挑战.

3 平滑约束的动态轨迹三维重建

3.1 平滑约束

基于 DCT 基和多项式基等预先定义基的轨迹重建方法依赖于基数量的准确估计, 难以适用于快

图 2 不同 DCT 基数量 K 下轨迹重建结果 (X 坐标)

Figure 2 Trajectory reconstruction results (X coordinate) under various number of DCT basis K . (a) $K = 2$; (b) $K = 9$; (c) $K = 12$; (d) $K = 15$

图 3 DCT 基轨迹重建相对误差与基数量 K 关系曲线

Figure 3 Relationship between relative error in trajectory reconstruction and number of DCT basis K

速变化的复杂轨迹重建; 同时, 采用 DCT 基和多项式基等难以以较少的基描述任意平滑轨迹, 易产生欠拟合与过拟合现象. 真实运动目标 (人体动作、地面车辆和空间目标等) 的三维轨迹往往是渐变、连续、平滑的, 为了刻画这一本质特性, 与基于 DCT 基和多项式基等预先定义基轨迹方法不同, 我们不引入任何预先定义基轨迹而是直接施加对于运动目标三维轨迹的平滑约束. 为了描述轨迹的渐变、连续与平滑特性, 我们引入如下平滑项 $\|H\mathbf{X}\|_2$, 其中平滑矩阵 $\mathbf{H} \in \mathbb{R}^{3F \times 3F}$ 定义为

$$\mathbf{H}_{ij} = \begin{cases} 1, & j = i, i = 1, \dots, 3(F-1), \\ -1, & j = i+3, i = 1, \dots, 3(F-1), \\ 0, & \text{其他}. \end{cases} \quad (10)$$

$\|\mathbf{H}\mathbf{X}\|_2^2 = \sum_{i=1}^{F-1} ((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2 + (z_i - z_{i+1})^2)$ 描述了运动目标三维轨迹相邻帧之间变化的平方和. $\mathbf{H}$ 可以是一阶平滑矩阵, 也可选择二阶或高阶平滑矩阵. 本文仿真和真实实验中验证了一阶平滑矩阵描述的性能.

引入平滑约束后, 运动目标轨迹三维重建问题需要同时求解模型 (4) 和优化平滑项 $\|\mathbf{H}\mathbf{X}\|_2$.

当图像观测中不含噪声时, 我们可以在求解模型 (4) 的解空间 $\mathbf{X} = \mathbf{Q}^\dagger \mathbf{q} + \mathcal{N}(\mathbf{Q})\mathbf{v}$ 中直接寻求最平滑的解, 即

$$\min_{\mathbf{X}} \|\mathbf{H}\mathbf{X}\|_2^2 = \min_{\mathbf{v}} \|\mathbf{H}\mathbf{Q}^\dagger \mathbf{q} + \mathbf{H}\mathcal{N}(\mathbf{Q})\mathbf{v}\|_2^2, \quad (11)$$

经过推导可以得到 $\hat{\mathbf{v}} = -((\mathbf{H}\mathcal{N}(\mathbf{Q}))^\top (\mathbf{H}\mathcal{N}(\mathbf{Q})))^\dagger (\mathbf{H}\mathcal{N}(\mathbf{Q}))^\top \mathbf{H}\mathbf{Q}^\dagger \mathbf{q}$ 最小化平滑约束, 对应的运动目标三维轨迹为 $\mathbf{X} = \mathbf{Q}^\dagger \mathbf{q} + \mathcal{N}(\mathbf{Q})\hat{\mathbf{v}}$.

实际图像观测中往往存在噪声, 因此 $\mathbf{Q}\mathbf{X} = \mathbf{q}$ 无法精确满足, 我们需要松弛模型约束, 寻求给出最小重建误差的运动轨迹, 即 $\min_{\mathbf{X}} \|\mathbf{q} - \mathbf{Q}\mathbf{X}\|_2^2$. 同时根据平滑约束, 我们寻求最平滑的运动轨迹, 即 $\min_{\mathbf{X}} \|\mathbf{H}\mathbf{X}\|_2^2$. 同时考虑这两种约束, 我们得到求解单目视频运动目标轨迹三维重建问题的无约束最优化模型:

$$\min_{\mathbf{X}} f(\mathbf{X}) = \|\mathbf{q} - \mathbf{Q}\mathbf{X}\|_2^2 + \lambda \|\mathbf{H}\mathbf{X}\|_2^2, \quad (12)$$

其中 $\lambda \geq 0$ 为平衡参数, 反映了模型拟合能力与轨迹平滑性之间的平衡.

3.2 闭式解

求解无约束最优化模型 (12), 我们可以重建运动目标的三维轨迹. 易见, 无约束最优化模型 (12) 为二次函数, 具有全局最优的闭式解.

首先计算代价函数 (12) 关于运动目标三维轨迹 $\mathbf{X}$ 的梯度

$$\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} = 2\mathbf{Q}^\top (\mathbf{Q}\mathbf{X} - \mathbf{q}) + 2\lambda \mathbf{H}^\top \mathbf{H}\mathbf{X}. \quad (13)$$

令 $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{0}$, 可以得到

$$(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})\mathbf{X} = \mathbf{Q}^\top \mathbf{q}, \quad (14)$$

从而得到闭式解

$$\mathbf{X} = (\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})^\dagger \mathbf{Q}^\top \mathbf{q}, \quad (15)$$

其中 $(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})^\dagger$ 表示矩阵 $(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})$ 的 Moore-Penrose 伪逆. 当矩阵 $(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})$ 可逆时, 线性系统 (14) 具有精确解. 当矩阵 $(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})$ 不可逆时, 以上的 Moore-Penrose 伪逆解给出线性系统 (14) 的极小范数最小二乘解. 实际中, 矩阵 $(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})$ 一般为可逆矩阵, 此时闭式解为 $\mathbf{X} = (\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})^{-1} \mathbf{Q}^\top \mathbf{q}$.

代价函数 (12) 的 Hessian 矩阵为 $2(\mathbf{Q}^\top \mathbf{Q} + \lambda \mathbf{H}^\top \mathbf{H})$, 由于 $\lambda \geq 0$, 这个矩阵为半正定矩阵, 而且代价函数 (12) 显然存在极小值, 因此闭式解 (15) 最小化代价函数 (12).

3.3 平衡参数的选择

提出的闭式解 (15) 包含相机标定信息、已知的平滑矩阵和图像观测, 唯一的参数是平衡参数 λ . λ 平衡了对于图像观测的拟合能力与运动目标三维轨迹的平滑性. 小的 λ 对于轨迹施加小的平滑约束, 通过对于模型 $QX = q$ 小的松弛寻求平滑的轨迹. 大的 λ 对于轨迹施加大的平滑约束, 从而较大程度的松弛模型约束 $QX = q$ 寻求平滑的轨迹.

当 $\lambda = 0$ 时, 对于轨迹不施加任何平滑约束, 模型 $QX = q$ 得到完全满足, 输出最小 L_2 范数解 $\hat{X} = Q^\dagger q$ 且观测重建误差为 0, 然而这一解具有大的平滑项, 往往不满足运动目标三维轨迹本身特性. 当 $\lambda \rightarrow +\infty$ 时, 重建轨迹 $X \rightarrow 0$, 输出的解为平凡解, 对于模型 $QX = q$ 没有任何约束, 观测重建误差极大. 当图像观测中存在较大噪声时, 需要减弱模型约束 $QX = q$ 而强化平滑约束, 即增大平衡参数 λ ; 当噪声较小时, 应强化模型约束 $QX = q$ 而减弱平滑约束, 即减小平衡参数 λ .

图 4 给出某一典型运动目标在不同平衡参数下的轨迹重建结果, 重建结果 (相对误差) 随平衡参数 λ 的变化关系曲线如图 5 所示. 从图中可以看出, 当 λ 趋于 0 时, 对于重建轨迹不施加或者施加很小的平滑性约束, 重建结果趋向于 $X = (Q^T Q)^\dagger Q^T q = Q^\dagger q$. 当 λ 取值同时满足模型约束与轨迹平滑约束时, 轨迹重建结果具有小的轨迹重建相对误差. 当 λ 不断增大时, 轨迹重建结果趋于常数, 轨迹重建相对误差不断增大. 当 λ 趋于 $+\infty$ 时, 重建轨迹 $X \rightarrow 0$, 重建结果为平凡解, 轨迹重建误差达到极大. 从图中可以看出, 当平衡参数 λ 位于一定的区间时, 运动目标三维轨迹的重建误差几乎保持不变且达到极小. 需要注意的是, 这一区间往往具有较大的范围, 从而保证 λ 在很大区间范围内都能较好重建运动目标三维轨迹.

3.4 部分观测缺失

实际系统中, 由于运动目标自身遮挡或者特征跟踪匹配的失效, 运动目标轨迹的图像观测中往往存在缺失数据. 当缺失数据在观测中呈现非强结构化 (成块缺失) 时, 平滑约束轨迹重建方法仍可应用. 令缺失观测模板为 Ω , 其定义为:

$$\Omega_{2i-1:2i} = \begin{cases} 1, & \text{第 } i \text{ 帧观测存在;} \\ 0, & \text{第 } i \text{ 帧观测缺失.} \end{cases} \quad (16)$$

则模型拟合残差项变为 $\|\Omega \odot (q - QX)\|_2^2$, $\odot$ 表示逐项相乘. 令 q' , Q' , X' 表示所有存在观测对应的矩阵. 采用如上表示, 部分观测缺失情形下的运动目标三维轨迹重建模型为

$$\min_{X'} f(X') = \|q' - Q'X'\|_2^2 + \lambda \|H'X'\|_2^2, \quad (17)$$

其中 H' 为对应可见观测的平滑矩阵. 其闭式解为 $X' = (Q'^T Q' + \lambda H'^T H')^\dagger Q'^T q'$. 因此本文提出的平滑约束轨迹重建方法适用于部分观测数据缺失的情形.

3.5 计算复杂度

平滑约束的单目视频运动目标轨迹三维重建方法具有以矩阵伪逆形式表达的闭式解. 矩阵的伪逆运算通过奇异值分解实现, 对于规模为 $m \times n$ 的矩阵, 奇异值分解的计算复杂度为 $O(4m^2n + 8mn^2 + 9n^3)$. 本文中矩阵的规模为 $3F \times 3F$, 计算复杂度为 $O(567F^3)$.

实际中, $Q^T Q$ 的秩为 $2F$, $H^T H$ 的秩为 $3(F-1)$, 一般情况下 $(Q^T Q + \lambda H^T H)$ 为满秩矩阵, 矩阵的伪逆运算可以简化为矩阵求逆运算, 其计算复杂度大大下降, 从而本文方法可以实现实时应用.

图 4 不同平衡参数下的轨迹重建结果

Figure 4 Trajectory reconstruction results under various trade-off parameter λ . (a) $\lambda = 0$; (b) $\lambda = 10^{-10}$; (c) $\lambda = 10^0$; (d) $\lambda = 10^3$; (e) $\lambda = 10^{10}$; (f) $\lambda = 10^{20}$

3.6 轨迹重建的唯一性

单目视频运动目标轨迹三维重建问题在不施加任何约束情况下是欠定的, 其解空间可以由相机观测运动目标视线 (由相机运动轨迹和运动目标轨迹确定) 的形式刻画. 预先定义基轨迹方法寻找解空间与预先定义基轨迹张成空间的交集. 本文提出的平滑约束方法则在解空间中寻求最平滑的轨迹. 然而, 在某些特定情形下, 即使给定以上约束仍无法准确重建运动目标轨迹.

Park 等^[1]对轨迹的可重建性进行了讨论, 将相机运动轨迹描述为 $\mathbf{C} = \Theta\beta_{\mathbf{C}} + \Theta^{\perp}\beta_{\mathbf{C}}^{\perp}$ 而目标运动轨迹为 $\mathbf{X} = \Theta\beta_{\mathbf{X}} + \Theta^{\perp}\beta_{\mathbf{X}}^{\perp}$, 其中 Θ 表示 DCT 基, $\perp$ 表示补空间. 可重建性定义为

$$\eta = \frac{\|\Theta^{\perp}\beta_{\mathbf{C}}^{\perp}\|}{\|\Theta^{\perp}\beta_{\mathbf{X}}^{\perp}\|}, \quad (18)$$

图 5 轨迹重建相对误差与平衡参数 λ 关系曲线Figure 5 Relationship between relative error in trajectory reconstruction and trade-off parameter λ

Park 等^[1]证明当 η 趋于无穷时, 基于 DCT 基的解趋向于问题的真实解.

于起峰等^[3]在多项式基描述下求解轨迹恢复问题, 给出了两种无法得到确定解的情形, 即 1) 相机运动轨迹也能用时间的多项式描述, 并且描述相机运动的时间多项式的次数等于或低于描述目标运动的时间多项式的次数; 2) 所有观察视线相交于同一点.

本文提出的平滑约束方法在轨迹重建解空间中的求解最平滑轨迹, 重建的轨迹总是唯一的. 然而, 当相机轨迹与目标运动轨迹满足特定条件时, 最平滑轨迹重建结果并不对应运动目标真实轨迹. 这里对两种典型退化情形进行分析.

退化情形 1: 视线相交. 首先我们讨论一种极端的退化情形, 即单静态相机运动目标轨迹重建, 显然给定单静态相机和运动目标的图像观测, 无法重建运动目标轨迹. 当单相机中心 C_i 保持不变时, 即使相机可以发生旋转, 仍然无法得到运动目标轨迹. 这是因为相机观测目标的视线只与相机位置有关而与相机自身的旋转无关. 对于以上两种情形, 由于相机中心是问题的解, 以最小化平滑约束为目标的方法恢复的运动目标轨迹为固定相机位置 $C_i = C_1$. 当相机位置沿观测视线移动并且自身旋转时, 所有观测视线仍相交于一点.

给定相机运动轨迹 C 和运动目标轨迹 X , 则时刻 i 的视线方向为 $X_i - C_i$. 所有视线相交的条件为: 对于所有的时刻 $i, i = 1, \dots, F$, 存在三维点 O , 满足如下约束:

$$X_i - O = c_i(C_i - O), \quad (19)$$

其中三维点 O 为视线相交点. 因此在视线相交情形下, 相交点 O 位于轨迹重建问题的解空间且平滑性为 0 达到最小, 因此本文提出的以平滑性最小为目标的重建结果为所有视线的交点 O , 显而易见这一平凡解不是运动目标真实轨迹. 同样, 多项式基和 DCT 基等预先定义基轨迹方法也无法获得确定解.

图 6 给出了轨迹重建问题的退化情形, 所有视线相交, 目标轨迹 1 和目标轨迹 2 具有相同的复杂度, 均位于解空间与预先定义基轨迹张成空间的交集, 因此预先定义基轨迹方法无法获得确定解. 由于平行视线相交于无穷远, 因此平行视线轨迹重建问题也无法得到确定解, 如图 7 所示.

如图 6 所示, 在观测视线相交情形下, 在视线上改变相机位置, 相机轨迹 (如相机轨迹 2) 的复杂度可以迅速增加而运动目标轨迹复杂度不变, 从而可以使得可重建性 η 趋于 $+\infty$, 但是显然在任何预先定义基轨迹下均无法获得确定唯一解, 这与 Park 等^[1]的结论不一致. 可见 Park 等^[1]提出的可重建性指标至少不适应所有观测视线相交于一点的情形.

图 6 动态轨迹重建中的视线相交退化情形

Figure 6 Degenerate case in dynamic trajectory reconstruction, where all the rays intersect at a point

图 7 动态轨迹重建中的视线平行退化情形

Figure 7 Degenerate case in dynamic trajectory reconstruction, where all the rays are parallel

图 8 动态轨迹重建中相机轨迹与运动目标轨迹均为低复杂度形式时的退化情形

Figure 8 Degenerate case in dynamic trajectory reconstruction, where both the camera trajectory and the target trajectory have low complexity

退化情形 2: 当相机运动轨迹的复杂度小于等于运动目标轨迹复杂度时, 两者线性组合的复杂度小于等于运动目标轨迹的复杂度. 这里的复杂度定义为轨迹在预先定义基轨迹下展开的阶数. 图 8 给出此情形的图示, 如图所示相机轨迹与运动目标轨迹均为低复杂度形式, 两者线性组合仍为低复杂度形式, 以平滑性度量也具有类似的结论.

令运动目标轨迹和相机轨迹在预先定义基轨迹下的表达为 $\mathbf{X} = \Theta\beta_{\mathbf{X}}$, $\mathbf{C} = \Theta\beta_{\mathbf{C}}$, 其中 Θ 为预先定义基轨迹, 如 DCT 基和多项式基等. 当 $\mu_i = \mu$ 时, 解空间可以表达为相机运动轨迹与目标运动轨迹的线性组合 $\mu\mathbf{X} + (1 - \mu)\mathbf{C}$, 即

$$\mu\mathbf{X} + (1 - \mu)\mathbf{C} = \Theta(\mu\beta_{\mathbf{X}} + (1 - \mu)\beta_{\mathbf{C}}), \quad (20)$$

则在预先定义基轨迹下, 轨迹的线性组合仍处于解空间与预先定义基轨迹张成空间的交集, 预先定义基轨迹方法无法区分这些轨迹, 从而无法得到唯一确定解. 本文提出的平滑约束方法在解空间中寻找最平滑轨迹, 由于平滑轨迹的线性组合仍具有平滑性, 因此最平滑解可能并不对应真实解而是相机轨迹与运动目标轨迹的某一线性组合.

令相机运动轨迹为 $\mathbf{C}$, 运动目标真实三维轨迹为 $\mathbf{X}$, 且两个轨迹均为平滑轨迹, 则两个轨迹的线性组合 $\hat{\mathbf{X}} = \mu\mathbf{X} + (1 - \mu)\mathbf{C}$ 的平滑程度为

$$\|\mathbf{H}\hat{\mathbf{X}}\|_2^2 = \mu^2\text{tr}(\mathbf{X}^T\mathbf{H}^T\mathbf{H}\mathbf{X}) + (1 - \mu)^2\text{tr}(\mathbf{C}^T\mathbf{H}^T\mathbf{H}\mathbf{C}) + 2\mu(1 - \mu)\text{tr}(\mathbf{X}^T\mathbf{H}^T\mathbf{H}\mathbf{C}), \quad (21)$$

以上目标函数为 μ 的二次函数形式, 平滑性函数最小点位置为

$$\hat{\mu} = \frac{\text{tr}(\mathbf{C}^T\mathbf{H}^T\mathbf{H}\mathbf{C}) - \text{tr}(\mathbf{X}^T\mathbf{H}^T\mathbf{H}\mathbf{C})}{\text{tr}(\mathbf{X}^T\mathbf{H}^T\mathbf{H}\mathbf{X}) + \text{tr}(\mathbf{C}^T\mathbf{H}^T\mathbf{H}\mathbf{C}) - 2\text{tr}(\mathbf{X}^T\mathbf{H}^T\mathbf{H}\mathbf{C})}, \quad (22)$$

一般情形下 $\hat{\mu} \neq 1$, 因此以最平滑为目标的重建得到的结果不同于运动目标真实轨迹. 随着 $\text{tr}(\mathbf{C}^T\mathbf{H}^T\mathbf{H}\mathbf{C})$ 的增大, $\hat{\mu}$ 不断增大, $\hat{\mu} \rightarrow 1$. 即相机运动轨迹与运动目标轨迹均平滑时, 相机运动轨迹越相对不平滑, 轨迹重建结果越趋向于运动目标真实轨迹.

重建结果的唯一性: 本文提出的平滑约束方法在由视线 (由相机运动轨迹和运动目标轨迹确定) 定义的解空间中求解最平滑的轨迹. 显然, 单目视频运动目标轨迹重建问题的解取决于相机运动轨迹、运动目标轨迹以及优化目标, 其唯一性和有效性同样取决于相机运动轨迹、运动目标轨迹以及选择的优化目标. 当所有观测视线相交时可以证明最平滑重建结果为交会点. 当视线不相交而相机运动轨迹与目标运动轨迹均平滑时, 两者的线性组合满足平滑性, 最平滑解不一定对应运动目标真实轨迹. 从相机运动轨迹规划角度讲, 增大相机运动轨迹的非平滑性和随机性, 可以提高相机运动轨迹与运动目标轨迹之间的不相关性, 并使得视线散乱分布, 视线方向不断变化, 从而在由视线定义的解空间中出现除真实运动目标轨迹外其他平滑轨迹的可能性降低, 提高运动目标轨迹的可重建性. 实际的手持、车载、机载和星载相机系统中, 相机运动轨迹并不平滑, 因而本文研究更加具有实用性.

4 实验

为了评测本文提出方法的性能, 我们在仿真与真实单目视频数据上进行了大量测试. 运动目标三维轨迹重建相对误差定义为

$$\epsilon = \frac{\|\hat{\mathbf{X}} - \mathbf{X}\|_2}{\|\mathbf{X}\|_2}, \quad (23)$$

其中 $\hat{\mathbf{X}}$ 为重建得到的运动目标三维轨迹, $\mathbf{X}$ 为真实的运动目标三维轨迹, $\|\cdot\|_2$ 表示向量的 L_2 范数.

4.1 仿真数据实验

首先在仿真数据上测试提出的方法, 衡量三维轨迹重建误差随图像观测中噪声强度的变化趋势, 并将本文方法与基于 DCT 基重建方法进行性能比较, 最后衡量提出方法在部分观测缺失情形下的性能.

图 9 轨迹重建结果对比

Figure 9 Comparison of trajectory reconstruction results. (a) X coordinate; (b) Y coordinate; (c) Z coordinate; (d) 3D trajectory

4.1.1 重建结果示例

为了获取真实平滑运动目标轨迹, 我们首先采用 CMU 运动捕获数据集 (<http://mocap.cs.cmu.edu/>) 中的人体运动数据. 给定二维图像观测矩阵和相机矩阵, 本文提出的方法以闭式解形式重建输出平滑约束的运动目标三维轨迹. 在 307 帧单目视频数据上的轨迹重建结果如图 9 所示. 图中给出了轨迹重建结果与真实轨迹在 X , Y , Z 坐标和三维轨迹的比较, 从图中可以看出提出方法很好地重建了运动目标的三维轨迹.

4.1.2 噪声性能

为了衡量提出方法在不同噪声强度下的统计性能, 我们在观测矩阵中添加零均值标准差为 σ 的高斯白噪声, 噪声强度定义为

$$\eta = \frac{\|\mathbf{n}\|_2}{\|\mathbf{W}\|_2}, \quad (24)$$

其中 $\|\mathbf{n}\|_2$ 表示噪声能量, η 为噪声能量与图像观测能量的比例. 不同噪声水平下轨迹重建实验结果如图 10 所示, 对于每一种噪声水平, 我们进行 50 次随机实验以给出统计结果. 从图可以看出, 提出的平滑约束轨迹重建方法具有比 DCT 基轨迹重建方法更小的轨迹重建误差, 当噪声强度较大时, 提出的平滑约束轨迹重建方法的优势更加明显.

图 10 平滑约束轨迹重建与 DCT 基轨迹重建性能比较
Figure 10 Performance evaluation of smoothness constrained trajectory reconstruction and DCT basis based trajectory reconstruction

图 11 不同观测缺失比例下轨迹重建相对误差
Figure 11 Relative error in trajectory reconstruction under various ratio of missing measurements

4.1.3 部分观测缺失时的性能

提出的平滑约束轨迹重建方法可以适用于部分观测数据缺失情形的轨迹重建. 为了衡量平滑约束重建方法在部分观测缺失情形下的性能, 我们在仿真数据上进行测试. 首先给定高斯白噪声强度, 随机构造缺失模板, 产生缺失观测, 在每一观测缺失比例下进行 50 次随机实验以获取统计结果. 高斯白噪声强度为 0.05 和 0.25 时不同观测缺失比例下本文提出的方法性能如图 11 所示. 从图中可以看出不同噪声强度下轨迹重建相对误差均随观测缺失比例的增加而增加, 同时大的噪声强度下具有更大的相对误差. 当观测缺失比例超过 50% 后轨迹重建相对误差迅速增加. 这是因为随着观测缺失比例的增加, 重建轨迹的平滑性不断下降, 不再满足平滑约束.

4.1.4 典型场景重建实验

为了测试提出的平滑约束轨迹重建方法的性能, 我们在若干典型场景进行轨迹重建实验. 实验情形包括空间直线运动, 平面二次曲线运动与空间任意平滑三维曲线运动, 真实的轨迹与重建轨迹如图 12 所示. 可见提出的方法在这些情形下均重建得到运动目标的三维轨迹. 特别的, 对于存在一定程度运动突变情形如图 12(c), 本文方法也得到满意结果.

4.2 单目视频复杂运动轨迹重建

为了测试平滑约束轨迹重建方法在长序列复杂轨迹重建中的性能, 我们采用 CMU 人体运动跟踪数据中的长序列复杂运动. 实验中产生 4 个透视相机位姿, 每一个图像帧时刻随机选择其中的一个相机观测运动目标, 从而得到单相机视频观测数据. 获取的图像观测添加高斯白噪声以模拟真实图像获取与特征点跟踪中的噪声. 利用本文提出的方法和 DCT 基方法重建运动目标的三维轨迹, 结果如图 13 所示. 各个重建情形中的轨迹重建相对误差也标在图中.

图 12 典型运动轨迹重建结果

Figure 12 Demonstrative trajectory reconstruction results. (a) Moving along 3D line; (b) curve on plane; (c) 3D curve

图 13 平滑约束运动目标复杂运动轨迹重建结果

Figure 13 Trajectory reconstruction results of our proposed method on complex trajectory scenarios. Our proposed method achieves relative error in trajectory reconstruction as 0.068%, 0.063% and 0.0648% compared with 8.6%, 13.42% and 1.83% by DCT basis based method in (a), (b) and (c) respectively

图 14 平滑约束运动目标复杂轨迹重建结果 (平滑约束轨迹重建方法误差为 0.064%, DCT 基方法轨迹重建误差为 15.39%)

Figure 14 Trajectory reconstruction results of smoothness constrained trajectory reconstruction method on complex movement, where the relative error in trajectory reconstruction of smoothness constrained method is 0.064% compared with 15.39% by DCT basis based method. (a) 3D trajectory; (b) X coordinate; (c) Y coordinate; (d) Z coordinate

图 14 给出了复杂运动轨迹重建结果及对应的 X, Y, Z 坐标轨迹重建结果, 从图中可以看出提出的平滑约束轨迹重建方法精确地重建了运动目标的复杂三维轨迹.

4.3 真实单目视频实验

为了衡量平滑约束运动目标轨迹三维重建方法在真实单目视频数据上的性能, 我们在单目手势跟踪与轨迹重建和单目视频航模运动跟踪与重建中进行测试. 需要注意的是, 我们获取到的是目标的三维轨迹而不是二维轨迹.

单目视频手势跟踪与轨迹重建实验结果如图 15 所示. 图 15(a) 为手势跟踪结果图像, 图中采用三维标定模板获取相机信息. 图 15(b) 为手势运动轨迹重建结果. 从重建的三维轨迹测量得到手势运动的距离为 103.80 cm, 而真实的距离约为 105 cm. 对比重建的轨迹和单目视频以及运动距离的度量, 可见提出的方法较精确重建了手势运动轨迹.

单目视频中航模运动跟踪结果如图 16 所示, 其中跟踪得到的运动目标以矩形框出, 重建的轨迹为目标质心的运动轨迹. 航模目标局部特写如图 16(e) 所示. 根据场景中的标定信息可以获取每帧图像对应的透视相机投影矩阵, 本文方法重建得到的航模的三维轨迹如图 17(a) 所示. 重建轨迹的波动源自航模实验中气流等原因引起的扰动以及目标跟踪与相机标定中的波动. 采用 DCT 基轨迹重建方法^[1]重建的航模运动轨迹如图 17(b) 所示. 为了进一步比较重建结果的不同, 我们给出了两种轨迹重

图 15 单目视频手势跟踪轨迹重建

Figure 15 Gesture tracking and trajectory reconstruction from monocular video. (a) Monocular video sequence; (b) trajectory reconstruction result

图 16 单目视频航模运动目标跟踪结果 (跟踪得到的运动目标在图像中用矩形标识)

Figure 16 Tracking result of the model airplane from monocular video sequence, where the moving target is labeled in rectangle

建结果在 XY 平面上的投影如图 17(c) 和 (d) 所示. 从图中可以看出, 提出的平滑约束轨迹重建方法精确的重建了运动目标的三维轨迹. 而 DCT 基轨迹重建方法得到的轨迹与真实轨迹差距较大, 在度量意义下具有更大的偏差 (单目视频中航模目标进行的是绕场飞行并包含俯冲和拉升动作, 这与平滑约束轨迹重建结果一致).

5 结论

本文引入运动目标三维轨迹的平滑约束作为求解单目视频运动目标轨迹三维重建的先验信息, 与基于 DCT 基和多项式基等预先定义基轨迹方法相比, 平滑性约束更具有一般性与通用性, 在约束问题求解的同时没有对于轨迹的过约束, 可以获取更加准确的轨迹重建结果. 引入平滑约束后, 得到轨

图 17 单目视频航模运动目标轨迹三维重建结果

Figure 17 Trajectory reconstruction result of the model airplane from monocular video sequence. (a) Trajectory reconstruction result of smoothness constrained method; (b) trajectory reconstruction result of DCT basis based method; (c) X - Y plane of trajectory reconstruction result of smoothness constrained method; (d) X - Y plane of trajectory reconstruction result of DCT basis based method

迹重建的闭式解, 可以显著减少计算时间. 提出的方法不需要基轨迹数量信息而且适用于部分观测缺失的情形. 本文进一步讨论了轨迹重建中的唯一性与可重建性问题, 对于不能获得确定轨迹的情形进行了分析. 通过仿真实验和真实单目视频序列实验进行了验证. 实验结果表明, 本文方法可以重建得到较高精度的运动目标三维轨迹. 同时, 本文提出的方法适用于移动手持、车载、机载和星载等运动平台情形, 具有更大的实用性.

进一步的研究包括引入时间因素, 在考虑轨迹重建同时衡量运动速度、加速度和角速度等的精度, 并将单目标运动跟踪与轨迹重建方法拓展至多目标情形. 同时本文方法也可以扩展应用于月球车巡游、机动平台多视角测绘、自然人机交互、目标跟踪、非刚性结构恢复^[14]和单目测量等.

参考文献

- 1 Park H S, Shiratori T, Matthews I, et al. 3D reconstruction of a moving point from a series of 2D projections. In: Proc. European Conf. on Computer Vision, Heraklion, 2010. 158–171
- 2 Akhter I, Sheikh Y, Khan S, et al. Trajectory space: A dual representation for nonrigid structure from motion. IEEE Trans Pattern Anal Mach Intell, 2011, 33: 1442–1456
- 3 Yu Q F, Shang Y, Zhou J, et al. Monocular trajectory intersection method for 3D motion measurement of a point target. Sci China Ser E-Tech Sci, 2009, 52: 3454–3463 [于起峰, 尚洋, 周剑, 等. 测量点目标运动参数的单目运动轨迹交会法. 中国科学 E 辑: 技术科学, 2009, 39: 1919–1927]
- 4 Xu Y H, Li S Q, Jia Y D. A vision-based method for finger-screen interaction. Acta Electron Sin, 2007, 35: 2236–2240

- [徐一华, 李善青, 贾云得. 一种基于视觉的手指屏幕交互方法. 电子学报. 2007, 35: 2236–2240]
- 5 Liu W, Yu H F, Yang H, et al. A new method for generalized obstacle detection based on monocular vision. *Acta Electron Sin*, 2011, 39: 1793–1799 [刘威, 于红绯, 杨恒, 等. 一种新的基于单目视觉的广义障碍物检测方法. 电子学报, 2011, 39: 1793–1799]
 - 6 Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision*. 2nd Edition. Cambridge: Cambridge University Press, 2004
 - 7 Avidan S, Shashua A. Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence. *IEEE Trans Pattern Anal Mach Intell*, 2000, 22: 348–357
 - 8 Shashua A, Wolf L. Homography tensors: On algebraic entities that represent three views of static or moving planar points. In: *Proc. European Conf. on Computer Vision*, Dublin, 2000. 507–521
 - 9 Kaminski J Y, Teicher M. A general framework for trajectory triangulation. *J Math Imaging Vis*, 2004, 21: 27–41
 - 10 Wolf L, Shashua A. On projection matrices $\mathcal{P}^k \rightarrow \mathcal{P}^2, k = 3, \dots, 6$ and their applications in computer vision. *Int J Comput Vis*, 2002, 48: 53–67
 - 11 Park H S, Sheikh Y A. 3D reconstruction of a smooth articulated trajectory from a monocular image sequence. In: *Proc. International Conf. on Computer Vision*, Barcelona, 2011. 201–208
 - 12 Chen M Y, AlRegib G, Juang B H. Trajectory triangulation: 3D motion reconstruction with l_1 optimization. In: *Proc. IEEE Conf. on Acoustics, Speech and Signal Processing*, Prague, 2011. 4020–4023
 - 13 Zhu Y Y, Cox M, Lucey S. 3D motion reconstruction for real-world camera motion. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Colorado Springs, 2011. 1–8
 - 14 Dai Y C, Li H D, He M Y. A simple prior-free method for non-rigid structure-from-motion factorization. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Providence, 2012. 2018–2025

Generic smoothness constrained 3D trajectory reconstruction of moving object from monocular video

DAI YuChao^{1*}, HE MingYi^{1*}, LI HongDong² & LI Bo¹

1 *School of Electronics and Information, Northwestern Polytechnical University, Shaanxi Key Laboratory of Information Acquisition and Processing, Xi'an, 710129, China;*

2 *Research School of Engineering, Australian National University, Canberra 2600, Australia*

*E-mail: daiyuchao@gmail.com, myhe@nwpu.edu.cn

Abstract In this paper, generic smoothness constrained 3D trajectory reconstruction of moving object from monocular video is proposed. By introducing the generic smoothness constraint on the 3D trajectory, unconstrained optimization model for 3D trajectory reconstruction is achieved and a closed-form solution is derived. Compared with the predefined basis methods such as Discrete Cosine Transform based method and polynomial basis based method, the proposed method is more generic and can be applied to incomplete measurement case, thus having broad applicability. Geometric explanation and uniqueness of 3D trajectory reconstruction are given. Experimental results on both synthetic data and real monocular video sequences have shown the validity and advancement of the proposed method.

Keywords smoothness constraint, 3D trajectory reconstruction, multi-view, moving object, monocular video, closed-form solution, line of sight

DAI YuChao was born in 1982. He received the B.E., M.E. and Ph.D. degree all in signal and information processing from Northwestern Polytechnical University, Xi'an in 2005, 2008 and 2012, respectively. Currently, he is a Post-doctoral Fellow with Research School of Computer Science, Australian National University. His research interests include computer vision, human-computer interaction and optimization. He won the best paper award in CVPR 2012.

HE MingYi received the B.Sc. and M.Sc. degrees in electronic engineering from Northwestern Polytechnical University, Xi'an in 1982 and 1985, respectively, and the Ph.D. degree in signal and information processing from Xidian University, Xi'an in 1994. Professor He is the Founder and Director of the Shaanxi Key Laboratory of Information Acquisition and Processing, a Professor in School of Electronics and Information, and the Director and Chief Scientist of the Center for Earth Observation Research, Northwestern Polytechnical University. He won 9 scientific prizes from China and the best paper award in CVPR 2012. He has made contributions to hyperspectral remote-sensing data processing, image processing and pattern recognition, neural networks, and 3D technique.

LI HongDong is currently a Fellow with the Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia. He is also seconded to National ICT Australia as a Research Scientist with the Canberra Research Laboratory. His current research interests include geometric computer vision, structure from motion, image restoration, medical image and bionic eyes, and also discrete combinatorial optimization. He constantly serves as a Programme Committee Member and/or a Reviewer on the top forums of computer vision research, i.e., ICCV, CVPR, and ECCV. He won the best paper award in CVPR 2012.

LI Bo was born in 1988. He received the B.S. degree in electronics and information engineering from Northwestern Polytechnical University, Xi'an in 2011. Currently, he is a Ph.D. student at Northwestern Polytechnical University. His research interests include pattern recognition, sparse representation, remote sensing images, etc.