射影定理在运动推断结构中的应用

于 XX

(北京航空航天大学 自动化科学与电气工程学院, 北京 100191)

摘　　要：运动推断结构是基于多视角几何的三维重建的第一个步骤，其在即时定位与地图构建中也有重要应用。经典的运动推断结构算法总会涉及到最小二乘问题的求解。运动推断结构算法中的三角测量算法是计算场景结构的必要步骤，而即使是相机运动求解也需要三角测量算法的辅助，因此三角测量算法是经典的运动推断结构算法中的非常重要的一步。三角测量的非线性优化问题可转化为线性最小二乘问题和最佳逼近问题。三角测量问题最终可以使用射影定理的思想进行求解。

关　键　词：射影定理；运动推断结构；三角测量；最小二乘问题

1　引　言

1.1运动推断结构问题描述

基于多视角几何的三维重建主要采用传统几何方法从输入的序列图像中重建出三维模型。从多张图像中重建三维模型，其首先需要通过运动推断结构（Structure from Motion, SfM）计算相机运动和场景结构[1]。

从输入输出的角度来看，运动推断结构可以被描述为如图 1 所示的问题。输入为多张摄像机在不同位置不同角度拍摄的照片（相邻两张之间有重叠区域），输出为被拍摄场景的三维结构（场景三维点坐标）及每张照片拍摄时对应的相机运动（平移和旋转）。图 1 中描述的 SfM 问题的输入为五张图片，输出为场景三维结构（黑色点）及相机位置姿态（红色点），图片与红色点之间的红色连线为相机位置和照片的对应关系。

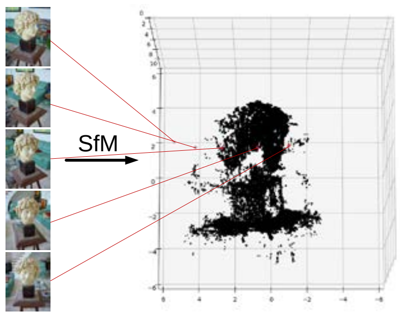

图 1 运动推断结构问题描

基于几何方法的 SfM 问题往往可以被分解为多个子问题，而每个子问题最终可被转化为求解线性方程组的最小二乘解的问题。即使是非线性的迭代估计方法，最后也会被转化为求线性方程组的最小二乘解问题。其中，可以利用射影定理来求解三角测量问题。

2　正　文

2.1简单的运动推断结构流程

简单的运动推断结构流程包括：1.相机标定；2.特征点提取；3.特征点匹配；4.对相邻的每对相机求解相机运动；5.三角测量解算场景结构点坐标；6.捆集调整进行全局的相机运动和场景结构优化。需要注意的是，流程并非线性的，尤其是 4，5，6 步骤需要穿插进行。其中步骤 5 可以利用泛函分析中射影定理相关知识来求解。

步骤 1-步骤 4 结合步骤 5 可以解算出两个摄像机拍摄位置精确的相机相对运动，解算得到的相机相对运动如图 2 所示。其中，$\mathrm{R_n}, T_n, n=1,2,3,4$分别为相机的旋转矩阵和平移向量。

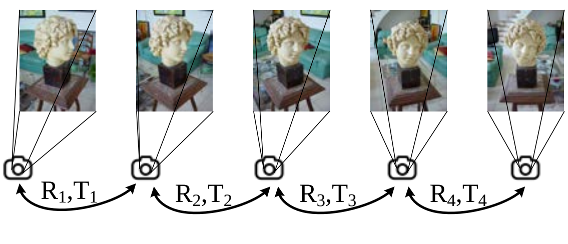

图 2 求解得到的相机相对运动示意图

2.2三角测量与射影定理

重点介绍与射影定理相关的三角测量问题。三角测量问题如图 3 所示，三角测量问题可以描述为，已知两个相机的摄像机矩阵$\mathbf{M}=\mathbf{K}[I|0], \mathbf{M}'=\mathbf{K}'[\mathbf{R}|\mathbf{T}]$，和测量点对$p, p'$，求解相应的三维点$P$的坐标。显然三角测量问题可以转化为求解射线$l, l'$的交点的问题。可以使用非线性三角测量法近似求解射线l, l'的交点。

首先由线性三角测量法求解出场景三维点坐标的值。该值可以作为非线性三角测量法[3]的初值输入。

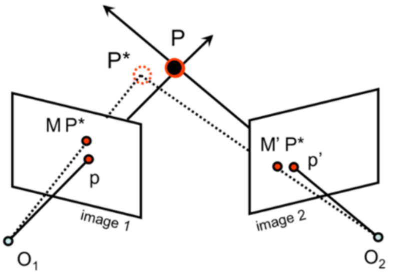

图 3 三角测量问题示意图[2]

非线性测量问题可以写为：

$$\arg\min_{P}\sum_{i}\|\mathbf{M}_i P - p_i\| \tag{1}$$

式中，p_i为图像平面上的点；$\mathbf{M}_i$为对应点p_i的摄像机矩阵；P为场景三维点坐标。摄像机矩阵可将三维点P映射至二维图像平面上。式(1)的含义为，在最小化重投影误差的前提下估计出最优三维点坐标。

设残差：

$$e_i(P) = \mathbf{M}_i P - p_i \tag{2}$$

将残差写为矩阵形式：

$$e = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix} \tag{3}$$

线性化：

$$e(\hat{P} + \delta_P) \approx e(\hat{P}) + J\delta_P \tag{4}$$

式中，$\hat{P}$ 为当前估计的场景三维点坐标；δ_P 为迭代步长，雅可比矩阵：

$$J = \frac{\partial e}{\partial \hat{P}} = \begin{bmatrix} \frac{\partial e_1}{\partial \hat{P}_1} & \frac{\partial e_1}{\partial \hat{P}_2} & \frac{\partial e_1}{\partial \hat{P}_3} \\ \vdots & \vdots & \vdots \\ \frac{\partial e_N}{\partial \hat{P}_1} & \frac{\partial e_N}{\partial \hat{P}_2} & \frac{\partial e_N}{\partial \hat{P}_3} \end{bmatrix} \tag{5}$$

则非线性优化问题可以线性化为：

$$\arg\min_{\delta_P}\left\|J\delta_P - (-e(\hat{P}))\right\|^2 \tag{6}$$

在每一次迭代过程中，需要按照式(6)进行迭代步长的求解。由射影定理[4]可知，求最小化范数 $\left\|J\delta_P - (-e(\hat{P}))\right\|^2$ 的 δ_P，就是求解在 J 的列空间中最接近 $-e(\hat{P})$ 的一个解，即 $-e(\hat{P})$ 在 J 的列空间中的最佳逼近元。图 4 所示为最佳逼近问题的示意图，射影定理告诉我们，最佳逼近问题的解就是 $-e(\hat{P})$ 在 J 的列空间中的投影。

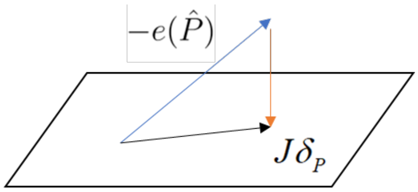

图 4 最佳逼近问题示意图

有 $J\delta_P - (-e(\hat{P}))$ 正交于 J 的列空间，即 $J\delta_P - (-e(\hat{P}))$ 正交于 J 的每一列：

$$J^{\mathrm{T}}(J\delta_P - (-e(\hat{P}))) = 0 \tag{7}$$

式(7)可以转化为：

$$J^{\mathrm{T}} J\delta_P = J^{\mathrm{T}}(-e(\hat{P})) \tag{8}$$

可知该正规方程的解为[5]：

$$\delta_P = -(J^T J)^{-1} J^T e \tag{9}$$

至此便利用射影定理解决了运动推断结构中的三角测量问题。

3 总结

射影定理在工程中有非常重要的作用。在三维重建中，射影定理使得运动推断结构中场景结构的求解变得非常简洁直观，是一种非常有用的工具。

参考文献

[1] Furukawa, Yasutaka & Hernandez, Carlos. (2015). Multi-View Stereo: A Tutorial. Foundations and Trends® in Computer Graphics and Vision. 9. 1-148. 10.1561/0600000052.

[2] Web.stanford.edu. (2019). CS231A: Computer Vision, From 3D Reconstruction to Recognition. [online] Available at: http://web.stanford.edu/class/cs231a/ [Accessed 21 Dec. 2019].

[3] Forsyth D A, Ponce J. Computer vision: a modern approach[M]. Prentice Hall Professional Technical Reference, 2002.

[4] 王永革. 应用泛函分析[M]. 北京航空航天大学出版社，2012.

[5] Hartley, R.I. & Zisserman, A. (2003). Multi-View Geometry in Computer Vision.