

线性迭代子空间射影重建法

刘侍刚^{①②} 吴成柯^② 赵录刚^① 宁纪锋^②

^①(西安电子科技大学综合业务网国家重点实验室 西安 710071)

^②(陕西师范大学计算机科学学院 西安 710065)

摘要: 该文提出了一种基于子空间线性迭代的射影重建方法, 该方法利用所有的图像序列构成的行向量生成的线性子空间之和与射影重建结构点构成的行向量生成的子空间是同一线性子空间及在该子空间中任何一个基底都可以作为射影重建的特性, 线性迭代地求取射影重建及图像深度因子。模拟实验和真实实验表明, 该射影重建方法具有鲁棒性好、收敛性好及重投影误差小等优点。

关键词: 图像处理; 射影重建; 分解算法; 线性迭代

中图分类号: TP391, P232

文献标识码: A

文章编号: 1009-5896(2007)02-0451-04

An Iterative Method Based on Subspace for Projective Reconstruction

Liu Shi-gang^{①②} Wu Cheng-ke^② Zhao Lu-gang^① Ning Ji-feng^②

^①(National Key Laboratory of ISN, Xidian University, Xi'an 710071, China)

^②(School of Computer Science, Shanxi Normal University, Xi'an 710065, China)

Abstract: An iterative factorization method based on linear subspace for projective reconstruction is presented in the paper. It relies on the facts that the rows in the matrix including all the image points span the same linear subspace as the rows in the matrix including space points and the fact that any basis of the subspace can be regarded as projective reconstruction. The projective reconstruction and the depth factors are obtained based on linear iteration. The experiments with both simulate and real data show that the method is efficient, robust and has good property of convergence and small reprojection errors.

Key words: Image processing; Projective reconstruction; Factorization methods; Linear iteration

1 引言

从图像序列中重建出三维场景结构是计算机视觉的主要目标之一。目前, 它仍然是计算机视觉领域中的研究热点。Faugeras指出^[1], 如果没有任何先验知识, 从图像测量中只能得到射影重建。

现在所提出的射影重建算法大部分都是基于多线性约束, 当只有两幅图像时, 存在一个双线性约束关系^[2], 一般也称为对极几何约束; 当只有 3 幅或 4 幅图像时, 存在三线性约束^[3]或四线性约束关系^[4]; 当图像数大于 4 时, 则不存在这样的几何约束关系^[5]。利用多线性约束关系进行射影重建的缺点是它仅能够利用 2 幅至 4 幅图像, 而不能利用更多的图像, 同时, 这种方法并没有把所有的图像统一来看待, 而是偏向某几幅图像。

为了克服上述缺点, Peter 基于 Tomasi-Kanade^[6] 正交投影分解算法的思想, 提出了利用分解算法来进行射影重建^[7]。该方法的缺点是, 在进行分解之前, 各个图像点的深度因子必须要事先求得, 而深度因子的求解又要利用对极几何约束关系。这样, 又遇到了前面所述的问题。

针对上述缺点, 本文利用所有的图像序列构成的行向量生成的线性子空间之和与射影重建结构点构成的行向量生成的子空间是同一线性子空间的特性, 提出了在该线性子空间中, 任何一个基底都可以作为射影重建的性质, 线性迭代地得到射影重建及图像深度因子。该方法可以避免求取多线性约束关系, 而且各个图像都是同等的看待, 并没有倚重某些图像。模拟实验和真实实验数据结果表明, 该射影重建方法具有鲁棒性好、收敛性好及重投影误差小等优点。

2 相机模型及投影矩阵

本文假定摄像机模型为经典的针孔模型^[8], 即成像过程可以用下列方程表示:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ z \\ 1 \end{bmatrix} \Leftrightarrow \lambda \mathbf{m} = \mathbf{K} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \mathbf{X} \quad (1)$$

式中 $\mathbf{X} = [x \ y \ z \ 1]^T$ 表示三维空间点的齐次坐标, $\mathbf{m} = [u \ v \ 1]^T$ 为对应的图像平面点的齐次坐标; \mathbf{K} 为摄像机的内参矩阵, 其中, $(u_0 \ v_0)$ 为主点坐标, f_u 为图像 u 轴的尺度因子, f_v 为图像 v 轴的尺度因子, s 为畸变因子; \mathbf{R}, \mathbf{t} 分别为摄像机在拍摄位置对应的旋转矩阵和平移矩阵, 即摄

像机的外参矩阵。

若令

$$P = K[R \ t] \quad (2)$$

那么 P 称摄像机的投影矩阵, 则式(1)可以简写成

$$\lambda m = PX \quad (3)$$

3 子空间描述^[9]

假设有 m 幅图像, 每幅图像上有 n 个对应的图像点, 则式(3)为

$$\lambda_{i,j} m_{i,j} = P_i X_j, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n \quad (4)$$

式中 i 和 j 表示第 i 幅图像和第 j 个图像点。

将所有空间点的齐次坐标表示成一个矩阵:

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ y_1 & y_2 & y_3 & \cdots & y_n \\ z_1 & z_2 & z_3 & \cdots & z_n \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \\ \mathbf{1} \end{bmatrix} \quad (5)$$

则对于第 i 幅图像上所有点及深度因子也有下式

$$m_i = \begin{bmatrix} u_{i,1} & u_{i,2} & u_{i,3} & \cdots & u_{i,n} \\ v_{i,1} & v_{i,2} & v_{i,3} & \cdots & v_{i,n} \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \\ \mathbf{1} \end{bmatrix} \quad (6)$$

$$A_i = \text{diag}(\lambda_{i,1} \quad \lambda_{i,2} \quad \cdots \quad \lambda_{i,n}) \quad (7)$$

根据式(4), 则有

$$m_i A_i = P_i X \quad (8)$$

若将所有图像及投影矩阵放在一起, 则有

$$M = PX \quad (9)$$

式中 $M = [m_1 A_1 \quad m_2 A_2 \quad \cdots \quad m_m A_m]^T$, $P = [P_1 \quad P_2 \quad \cdots \quad P_m]^T$ 。

用 Ω_i 表示由 m_i 的行向量生成的三维线性子空间, 即

$$\Omega_i = \text{span}\{\mathbf{u}_i, \mathbf{v}_i, \mathbf{1}\} \quad (10)$$

式中 \mathbf{u}_i 和 \mathbf{v}_i 分别表示由第 i 幅图像的所有图像点的 u 和 v 分量组成的行向量, $\mathbf{1}$ 表示由 n 个 1 组成的行向量。

同样, 用 Ω 表示由 X 的行向量生成的四维线性子空间, 即

$$\Omega = \text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{1}\} \quad (11)$$

式中 \mathbf{x} , \mathbf{y} 和 \mathbf{z} 分别表示由空间点 x 轴, y 轴和 z 轴的分量组成的行向量。

对于 $m_i A_i$, 同样可以生成一个三维的线性子空间, 即

$$\Omega_i A_i = \text{span}\{\mathbf{u}_i A_i, \mathbf{v}_i A_i, \mathbf{1} A_i\} \quad (12)$$

由式(8)可得

$$\Omega_i A_i \subset \Omega \quad (13)$$

$$\Omega_1 A_1 + \Omega_2 A_2 + \cdots + \Omega_m A_m \subseteq \Omega \quad (14)$$

若式(13)中取 \subset 号, 表示所有的子空间 $\Omega_i A_i$ 重合, 也就是表示所有的图像一样, 这是一种特殊情况^[9]。对于一般情况下取等号, 在本文中, 并不考虑这种特殊情况, 因此本文取等号。

从式(9) M 矩阵的构成中可以看出, 由 M 矩阵的行向量所生成的线性子空间为 $\Omega_1 A_1 + \Omega_2 A_2 + \cdots + \Omega_m A_m$, 即为 Ω 。

同时, 从式(8)中可以看出, 对于任何一个非奇异矩阵 T , 都有

$$m_i A_i = P_i X = P_i T^{-1} T X = P'_i X' \quad (15)$$

式中 $P'_i = P_i T^{-1}$, $X' = T X$ 。即对于一组图像点序列 m , 如果存在一组重建 P 和 X , 则同样存在另一组重建 P' 和 X' 。对于无穷多的非奇异矩阵 T , 存在无穷多组重建满足要求。如果没有任何先验知识, 不能知道所求得的重建对应的是哪一个变换 T 。从而这个重建落在射影空间而不是欧氏空间。但从式(15)可以看出, 由于 T 为非奇异矩阵, 因此, 重建 X 和 X' 的行向量生成的线性子空间为同一子空间 Ω , 只是它们的基底不一样, 而且 X 和 X' 的行向量都可以作为这个子空间的一个基底。

4 线性迭代射影重建

定理 在由空间点 X 生成的线性子空间 Ω 中, 任何一个基底都可以作为射影重建 X' 的结果。

证明 设 $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ 是由 X 生成的子空间中的任意一个基底, 则存在一个非奇异矩阵 T , 使

$$(\sigma_1 \quad \sigma_2 \quad \sigma_3 \quad \sigma_4)^T = T X \quad (16)$$

由式(15)可得, 若令 $P'_i = P_i T^{-1}$, $X' = T X = (\sigma_1 \quad \sigma_2 \quad \sigma_3 \quad \sigma_4)^T$, 则 P' 和 X' 可以认为是一组射影重建。

因此, 线性子空间 Ω 中的任何一组基底都可以当作射影重建 X 的结果。证毕

从式(10)可以看出, 矩阵 M 的秩满足 $\text{Rank}(M) = 4$ 。将 M 矩阵进行奇异值分解(SVD)

$$M_{3m \times n} = S_{3m \times 3m} V_{3m \times n} D_{n \times n}^T \quad (17)$$

式中 S 和 D 为正交阵, $V = \text{diag}(v_1 \quad v_2 \quad v_3 \quad v_4 \quad 0 \quad 0 \quad \cdots \quad 0)$ 。

由于图像中含有噪声, 因此不可能保证在对矩阵 M 进行 SVD 分解后的 V 中刚好对角线中前 4 个元素为非零, 而其它元素都为零, 可以令对角线元素 $v_i = 0 (i \geq 5)$ 。

式(17)可以化简, 得

$$M_{3m \times n} = S'_{3m \times 4} V'_{4 \times 4} D'^T_{n \times 4} \quad (18)$$

式中 $S'_{3m \times 4}$ 和 $D'_{n \times 4}$ 分别为 $S_{3m \times 3m}$ 和 $D_{n \times n}$ 的前 4 列, $V'_{4 \times 4} = \text{diag}(v_1 \quad v_2 \quad v_3 \quad v_4)$ 。

可以证明, 由 D' 的列向量所生成的线性子空间也为 Ω , 因此, 可以令投影矩阵及射影空间点为

$$P = S'_{3m \times 4} V'_{4 \times 4}, \quad X = D'^T_{n \times 4} \quad (19)$$

但是, 为了计算更加稳定, 令

$$P = S'_{3m \times 4} W_{4 \times 4}, \quad X = W_{4 \times 4} D'^T_{n \times 4} \quad (20)$$

式中 $W_{4 \times 4} = \text{diag}(\sqrt{v_1} \quad \sqrt{v_2} \quad \sqrt{v_3} \quad \sqrt{v_4})$ 。

通过以上分析, 可知若深度因子 A_i 已知, 可以通过 SVD 分解来完成射影重建。下面来求解深度因子 A_i 。

将空间一个向量投影到由矩阵 X 的行向量所生成的线性子空间的正投影矩阵 T 为^[10]

$$T = X^T (XX^T)^{-1} X \quad (21)$$

投影到其正交补空间中的投影矩阵 T_{\perp} 为

$$T_{\perp} = I - X^T (XX^T)^{-1} X \quad (22)$$

由式(13)及投影矩阵 T_{\perp} 的性质可知

$$m_i A_i T_{\perp} = \mathbf{0}_{3 \times n} \quad (23)$$

令 $T_{\perp} = [t_1 \ t_2 \ \dots \ t_n]^T$, 则式(23)为

$$\lambda_{i,1} \begin{bmatrix} u_{i,1} \\ v_{i,1} \\ 1 \end{bmatrix} t_1 + \lambda_{i,2} \begin{bmatrix} u_{i,2} \\ v_{i,2} \\ 1 \end{bmatrix} t_2 + \dots + \lambda_{i,n} \begin{bmatrix} u_{i,n} \\ v_{i,n} \\ 1 \end{bmatrix} t_n = \mathbf{0}_{3 \times n} \quad (24)$$

整理得

$$\begin{bmatrix} u_{i,1} t_1^T & u_{i,2} t_2^T & \dots & u_{i,n} t_n^T \\ v_{i,1} t_1^T & v_{i,2} t_2^T & \dots & v_{i,n} t_n^T \\ t_1^T & t_2^T & \dots & t_n^T \end{bmatrix}_{3n \times n} \begin{bmatrix} \lambda_{i,1} \\ \lambda_{i,2} \\ \vdots \\ \lambda_{i,n} \end{bmatrix} = \mathbf{0}_{3n \times 1} \quad (25)$$

从式(25)可以看, 若投影空间点已知, 可以线性地求解深度因子 A_i 。

从以上分析可知, 我们可以构造一个线性迭代算法, 可以线性迭代地求取射影重建及深度因子。

5 算法描述

- (1) 假设所有的深度因子都为1, 令 ε 为任意小的一个正数。
- (2) 对 M 矩阵进行SVD分解, 利用式(18)和式(20)求得射影重建 P 和 X 。若 $v_5 < \varepsilon$, 则停止; 否则到第(3)步。
- (3) 利用求得的射影空间点 X , 构造投影矩阵 T_{\perp} , 根据式(25)求得深度因子 A_i , 返回第(2)步。

6 模拟与真实图像实验

6.1 模拟实验

为了检验本文所提出的算法的收敛性, 用计算机模拟在单位球内随机产生 100 个空间点, 模拟摄像机的内参数 $f_u = 1000$, $f_v = 1000$, $s=0$, $u_0 = 320$, $v_0 = 240$, 变换摄像机的外参数以产生 8 幅大小为 640×480 的图像, 并利用这些模拟图像点进行射影重建, 模拟结果如图 1 所示。

从图 1 可以看出, 本文所提出的算法迭代 10 次以内就能够达到收敛, 因此该算法具有良好的收敛性。

为了比较本文所提出的算法和 Peter 算法的性能, 在图像像素中加入高斯噪声, 在每种噪声水平下各进行 100 次实验, 利用式(26)求得重投影误差, 然后取均值, 实验结果如图 2 所示。

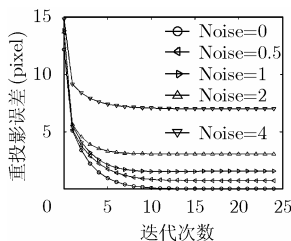


图 1 重投影误差随

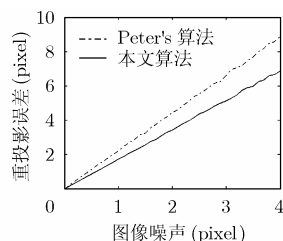


图 2 重投影误差随

迭代次数变化图

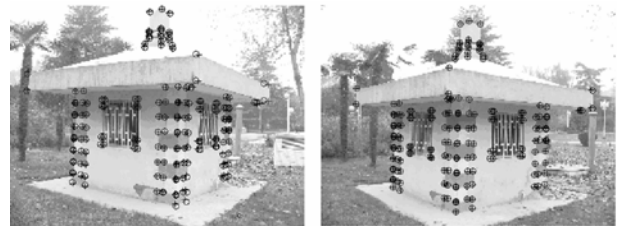
图像噪声变化图

$$\sigma = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n \|P_i x_j - u_{ij}\| \quad (26)$$

从图 2 中可以看, 本文提出的射影重建方法的重投影误差比Peter的分解算法^[7]重投影误差要小, 原因主要有两方面: (1)本文的方法不需要计算基础矩阵, 基础矩阵的计算对图像噪声非常敏感^[5], 而Peter的方法需要计算基础矩阵。(2)本文的方法把所有的图像统一看待, 并没有倚重哪一幅图像, 而Peter的方法在计算深度因子时, 却并没有把所有的图像统一看待。因此本文的方法具有更小的重投影误差。从图中还可以看出, 在 4 个像素的噪声之内, 本文算法求得的重投影误差只比原来的图像噪声约大 70%, 具有较小的重投影误差。

6.2 真实图像实验

为了验证本算法的正确性, 本文获得一个由 10 幅图像组成的房子序列, 图像的大小为 512×480 , 其中两幅图像如图 3 所示。用 Harris 算子提取了 120 个特征点(如图 3(a)所示), 并用这些特征点进行射影重建, 重建后再重投影的点如图 3(b)所示。从图中可以看出, 本文提出的算法具有良好的重建结果。



(a) 第 3 幅图像 (b) 第 7 幅图像

图 3 图像的特征点(+)及重投影点(O)

7 结束语

本文提出了一种利用子空间线性迭代求取射影重建的方法, 模拟实验结果表明, 该方法只需要 10 次以内迭代就可以达到收敛, 同时, 重投影误差约大于图像噪声的 70%, 因此, 该方法具有良好的收敛性、鲁棒性及重投影误差小等优点。同时, 真实图像实验结果表明, 本文提出的射影重建方法是非常有效的。

参考文献

- [1] Faugeras O. What can be seen in three dimensions with an uncalibrated stereo rig? ECCV'92, Lecture notes in Computer Science, Springer-Verlag, 1992, 588: 563-578.
- [2] Kuthirummal S, Jawahar C, and Narayanan P. Planar shape recognition across multiple views. International Conference on Pattern Recognition, Quebec City, Canada, Aug. 2002: 482-488.
- [3] Hartley R. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 1997,

- 22(2): 125–140.
- [4] Hartley R. Ambiguous configurations for 3-view projective reconstruction. Proceedings of European Conference on Computer Vision, Ireland, 2000: 922–935.
- [5] Liu S, Wu C, Tang L, and Jia J. An iterative factorization method based on rank 1 for projective structure and motion. *The IEICE Transactions on Information and Systems*, 2005, E88-D(9): 2183–2188.
- [6] Tomasi C and Kanade T. Shape and motion from image streams under orthography: A factorization method. *International Journal of Computer Vision*, 1992, 9(2): 137–154.
- [7] Sturm P and Triggs B. A factorization based algorithm for multi-image projective structure and motion. Proceedings of European Conference on Computer Vision, Cambridge, 1996: 709–720.
- [8] 刘侍刚, 吴成柯, 唐丽, 严尧平. 一种基于线性迭代自定标方法. 电子学报, 2004, 32(10): 1716–1719.
- Liu Shi-gang, Wu Cheng-ke, Tang Li, and Yan Rao-ping. A new self-calibration algorithm based on linear iteration. *Acta Electronica Sinica*, 2004, 32(10): 1716–1719.
- [9] Heyden A, Berthilsson R, and Sparr G. An iterative factorization method for projective structure and motion from image sequences. *Image and Vision Computing*, 1999, 17(13): 981–991.
- [10] 程云鹏, 张凯院, 徐仲. 矩阵论. 西安: 西北工业大学出版社, 2001: 289–293.
- 刘侍刚: 男, 1973 年生, 博士. 主要研究方向为计算机视觉和图像处理等方面的研究工作.
- 吴成柯: 男, 1938 年生, 教授, 博士生导师. 主要研究方向为计算机视觉和图像通信等方面的研究工作.
- 赵录刚: 男, 1975 年生, 博士. 主要研究方向为计算机视觉和图像处理等方面的研究工作.
- 宁纪锋: 男, 1975 年生, 博士. 主要研究方向为计算机视觉和图像处理等方面的研究工作.