

# 基于差分进化算法的射影重建

李 强, 余贞侠

(成都信息工程大学, 四川 成都 610225)

**摘要:**基于因式分解的射影重建可统一使用多视图图像而不需要指定参考图像,目前成为基于多视图三维重建的一个研究热点。而基于因式分解重建方法重要的一步就是定义射影深度矩阵约束条件,首先以基于射影深度矩阵的行列  $l_1$  范数作为约束条件,使其避免在迭代因式分解过程中收敛到错误结果,然后再使用差分进化算法在约束的解空间中得到射影深度矩阵,进而实现射影重建。模拟数据和真实数据的试验结果该方法是有有效的,且有鲁棒性好、重投影误差小等优点。

**关 键 词:**差分进化算法;射影重建;因式分解;多视图

## 0 引言

Tomasi-Kanade 提出的基于因式分解重建方法是从运动恢复结构(structure from motion, SFM)研究领域的一个热点<sup>[1]</sup>。其基本原理是通过奇异值分解(singular value decomposition, SVD),从多个视图的得到的测量矩阵  $M$  得到摄像机矩阵  $P$  和三维结构  $X$ 。有别于其他的重建方法,基于因式分解的重建方法能够统一的使用图像序列中的所有的图像,而且不需要指定参考图像。

Tomasi-Kanade 的因式分解方法是基于仿射分解的并不能直接应用于射影重建中,Strum 在其基础上提出了一种基于因式分解的射影重建方法<sup>[2]</sup>。基于因式分解的射影重建可以简单地表示为

$$A \otimes M = PX \quad (1)$$

其中,  $A$  是射影深度矩阵,  $\otimes$  是按元素的乘积,  $M$  是从多个是视图得到的像点组成的矩阵,  $P$  是多个视图组成的矩阵,  $X$  是场景的三维结构。对于一个图像序列来说,像点组成的矩阵  $M$  是已知的,而射影深度矩阵  $A$  是未知的,所以基于因式分解的射影重建的关键问题就是如何得到射影深度矩阵  $A$  的一个正确估计。另一个要解决的问题是,不是所有式(1)的解都是适用的,例如  $A$  为 0 矩阵,这就需要添加一些约束条件排除那些不适用的解。

起初,Strum 和 Triggs 给出了一种利用基础矩阵估计射影深度的非迭代方法,由于基础矩阵的计算对特征点的匹配误差非常敏感,该算法对真是的数据并不能给出很好的结果。Triggs<sup>[3]</sup> 又提出可以使用迭代的

方法来估计射影深度,不再计算基础矩阵,并且建议在每次迭代结束对得到的射影深度矩阵进行归一化处理,使矩阵的每行都有相同的欧氏距离,每列也有相同的欧氏距离。然而,使用迭代方法来估计射影深度也有一些不足之处,其迭代也不一定会收敛,Mahamud<sup>[4]</sup> 指出即使使用的 Triggs 提出的归一化算法也不能保证迭代一定收敛到正确结果。并提出一种使用列的  $l_2$  范数作为约束的迭代算法,但是只给出局部的收敛性证明,并没有给出全局收敛性的证明。与 Mahamud 类似,Heyden 使用行的  $l_2$  范数作为约束,利用线性子空间估计射影深度<sup>[5]</sup>。但是,如果只是每行都有相同的  $l_2$ ,显然 0 行也是符合这一条件。迭代算法也需要一个初始值,对于使用 1 作为初始射影深度的估计,Harley 和 Zisserman 认为,只有当真正的射影变换非常接近仿射的时候,这类迭代算法才能很好的工作<sup>[1]</sup>。Oliensis 对 Strum-Triggs 的迭代算法做了分析,指出没有进行归一化的 Strum-Triggs 迭代算法总是会收敛到一个错误的结果<sup>[6]</sup>。而 Mahamud<sup>[4]</sup> 的可证明的收敛算法会收敛到一个无用的结果,而对行或者列进行归一化可能会导致迭代过程不是很稳定,并给出一种规则化的迭代算法(CIESTA),该算法会收敛到一个稳定的结果,但是需要真正的射影深度接近于 1。

本文给出一种使用行和列的  $l_2$  范数作为约束的射影重建算法,并在每次迭代的过程中使用阶梯矩阵

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (2)$$

作为掩码保证不会收敛到错误的结果。在进行重建的过程仍然以矩阵的最小秩为目标函数,选择使用差分进化算法迭代估计摄影深度,进而使用 SVD 分解技术

得到射影重建结果。

## 1 基于因式分解的射影重建

### 1.1 基本原理

假设存在  $n$  个三维的点  $X_1, X_2, \dots, X_n \in R^4$  和  $m$  个射影矩阵  $P_1, P_2, \dots, P_m \in R^{3 \times 4}$ , 三维点  $X_j$  通过射影矩阵  $P_i$  得到像点  $x_{ij}$  且有:

$$\lambda_{ij} x_{ij} = P_i X_j \quad (3)$$

其中,  $\lambda_{ij}$  是射影深度是一个非 0 的缩放因子。射影深度  $\lambda_{ij} (i=1, 2, \dots, m, j=1, 2, \dots, n)$  可以组成一个  $m \times n$  的深度矩阵  $\Lambda = [\lambda_{ij}]$ , 同样图像的像点  $x_{ij}$  组成数据矩阵  $[x_{ij}]$ , 这样式(3)可以改写为

$$W = \Lambda \otimes [x_{ij}] = P X \quad (4)$$

其中,  $P = [P_1, P_2, \dots, P_m]^T, X = [X_1, X_2, \dots, X_n]$ ,  $\Lambda \otimes [x_{ij}] = [\lambda_{ij} x_{ij}]$ 。从式(4)可以看出测量矩阵  $W$  可以分解为  $3m \times 4$  的矩阵  $P$  和  $4 \times n$  的矩阵  $X$ , 也就是说测量矩阵  $W$  的秩最大为 4。

由于深度矩阵  $\Lambda$  是未知的, 从而测量矩阵  $W$  也是未知的, 故其分解式(4)也是未知的, 所以基于因式分解的射影重建就是通过某种方式找到射影深度  $\Lambda$  的一个正确估计, 然后对测量矩阵  $W$  进行奇异值分解得到摄像机矩阵以及场景的三维结构。

Sturm-Triggs 类型的迭代算法一般是将下面步骤交替进行取得一个射影深度矩阵  $\Lambda$  的估计:

步骤 1 固定射影深度矩阵  $\Lambda$ , 通过 SVD 分解得到  $P$  和  $X$ 。

步骤 2 固定  $P$  和  $X$ , 通过最小二乘法解得射影深度矩阵  $\Lambda$ 。

步骤 3 交替的执行上面两步, 直到结果收敛。

通常为了避免收敛到错误结果(所有的射影深度为 0, 或者有 4 个 0 列等), 会在步骤 1 和步骤 2 之间插入归一化的步骤(行列的欧氏长度相等, 行和列有相同的  $l_2$  范数等)。

### 1.2 重构的约束条件

通过以上的分析可知, 基于因式分解的射影重构首先需要找到深度矩阵  $\Lambda$  的一个合适估计, 然后应用式(4)对测量矩阵进行因式分解得到摄像机矩阵  $P$  和三维结构矩阵  $X$ 。在没有任何约束条件下, 求到一个合适  $\Lambda$  几乎是不可能的, 在文中应用了以下几个条件对深度矩阵  $\Lambda$  进行约束: 任何可以观测到的像点应该再摄像机的前面, 也就是说所有观测到的像点的射影深度都应该是正的; 测量矩阵  $W$  的秩不大于 4, 深度

矩阵  $\Lambda$  是  $W$  的子矩阵, 所以有  $\text{rand}(W) \leq 4 \Rightarrow \text{rand}(\Lambda) \leq 4$ ; 深度矩阵  $\Lambda$  每一行都有相同的  $l_1$  范数, 每一列也如此。

使用深度矩阵  $\Lambda$  行和列的  $l_2$  范数作为约束条件来源于  $\Lambda$  的一个特殊性质。任意的三维点  $X = [x, y, z, 1]^T$  通过射影矩阵  $P$  得到像点  $x = [x, y, 1]^T$  有:

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

其中,  $P_i$  是射影矩阵  $P$  的第  $i$  行, 从式(5)可以得到射影深度  $\lambda = P_3 \cdot X$ 。  $n$  个三维点  $m$  个视角的射影深度组成的  $m \times n$  射影深度矩阵可以表示为:

$$\Lambda = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & \vdots & \vdots \\ \lambda_{m1} & \dots & \lambda_{mn} \end{bmatrix} \quad (6)$$

对于射影深度矩阵  $\Lambda$ , 第  $i$  行是第  $i$  个视角下不同的三维点的射影深度; 第  $j$  列是第  $j$  个三维点在不同视角下的射影深度, 也就是说  $\Lambda$  中的每一行射影深度对应的摄像机矩阵是相同的, 每一列射影深度对应的三维点相同。从前面的分析有射影深度  $\lambda = P_3 \cdot X$  ( $P_3$  是摄像机矩阵的第 3 行), 那么可以得出射影深度矩阵  $\Lambda$  的第  $i$  行有一个公共因子  $\pi_i$ , 其第  $j$  列也有一个公共因子  $v_j$ , 具体公式如下:

$$\Lambda = \text{diag}(\pi) \Lambda' \text{diag}(v) \quad (7)$$

其中,  $\text{diag}(\pi)$  是向量  $[\pi_1, \pi_2, \dots, \pi_m]$  的对角矩阵,  $\text{diag}(v)$  是向量  $[v_1, v_2, \dots, v_n]$  的对角矩阵。

根据 Sinkhorn 著名的双随机矩阵<sup>[13]</sup>理论, 如果一个非负的方阵的每一行的都为 1, 每一列的和也为 1, 那么这个方阵就称为双随机矩阵 (doubly stochastic matrix)。Sinkhorn 证明了下面的理论:

定理 1: 任意的  $n$  阶正方形  $A$  都可以通过  $D_1 A D_2$  缩放为双随机矩阵,  $D_1$  和  $D_2$  是  $n$  阶正的对角矩阵。

该理论也可以很容易推广到  $m \times n$  的矩阵。

定理 2: 任意正的  $m \times n$  矩阵  $A$  都可以通过  $D_1 A D_2$  缩放为每一行的和为  $n$ , 每一列的和为  $m$ 。其中  $D_1$  是  $m \times m$  的对角矩阵,  $D_2$  是  $n \times n$  的对角矩阵<sup>[14]</sup>。

本文中对深度矩阵  $\Lambda$  使用的约束条件可以归结如下:

$$\text{rand}(\Lambda) \leq 4$$

$$\sum_i \lambda_{ij} = m, \quad j = 1, 2, \dots, n$$

$$\sum_j \lambda_{ij} = n, \quad i = 1, 2, \dots, m$$

$$\lambda_{ij} > 0$$

## 2 基于差分进化算法的射影深度估计

差分进化算法<sup>[6]</sup>由 Storn 和 Price 于 1996 年提出的一种新型进化算法,因其操作简单、易执行、性能卓越等特点因其众多研究者的关注。相比于梯度下降算法、牛顿法等需要问题是可微分的,差分进化算法在有优化问题时并不依赖于梯度,因此可以使用差分进化算法解决那些不连续,噪声随着时间发生变化的复杂问题的优化<sup>[11]</sup>。

如何的得到射影深度矩阵的一个正确估计是进行射影重构的关键,下面使用差分进化算法(differential evolution, DE)结合上文的约束条件得到射影深度矩阵的一个最优解。

### 2.1 解的编码和适应度函数

对于参数  $\lambda_{ij}(i=1, \dots, m, j=1, \dots, n)$  采用十进制编码。令  $\{\lambda_{ij}\}$  是射影深度矩阵的一个解。以  $\hat{\lambda} = \{\lambda_{11}, \dots, \lambda_{1n}, \lambda_{m1}, \dots, \lambda_{mn}\}$  作为一个个体进行编码。由于透视模型可以使用仿射模型来迭代逼近<sup>[3.3]</sup>,且仿射模型的仿射深度是一个常量(通常取 1),因此  $\lambda_{ij}$  可以在  $(0, 2]$  区间随机产生。

在理想无噪声情况下,测量矩阵  $W$  的秩  $\text{rand}(W) = 4$ ,故此定义差分进化算法的适应度函数为

$$f(\hat{\lambda}) = \frac{1}{1 + \sigma_5}$$

其中,  $\sigma_5$  为测量矩阵  $W$  的排序第 5 的奇异值,  $f(\hat{\lambda})$  是依赖于  $\hat{\lambda}$  的非负实数,为使  $\text{rand}(W) \leq 4$ ,则必有  $\sigma_5 = 0$ 。于是,求使  $f(\hat{\lambda})$  是最大值的  $\lambda$  既是射影深度的一个正确估计。由于  $f(\hat{\lambda})$  不是一个线性函数,故最大化  $\hat{\lambda}$  是一个非线性优化问题。

### 2.2 射影深度矩阵估计

基于差分进化算法的射影深度矩阵估计具体步骤如下:

步骤 1 从  $m$  幅图像中选取  $n$  个对应的特征点,并得到特征点的图像坐标  $m_{ij}(i=1, \dots, m, j=1, \dots, n)$ 。

步骤 2 对特征点的图像坐标进行归一化,即选择任一特征点  $m_i$  作为坐标原点,并计算  $\delta m_{ij} = m_{ij} - m_i$ ,然后将  $\delta m_{ij}$  乘以一个公共因子  $s_i(s_i = \sqrt{2}/\frac{1}{p} \sum_{j=1}^n \|\delta m_{ij}\|_2)$ ,使他们的平均范数为  $\sqrt{2}$ 。

步骤 3 在  $(0, 2]$  区间内,初始化种群  $\{A\}$ ,种群容量取 20。

步骤 4 变异操作,从当前种群中随机选择 3 个个

体作为父代进行变异操作,将选中的 2 个个体进行向量差缩放后和第 3 个个体进行向量和得到变异个体。

$$v_1(g+1) = v_{r1}(g) + F(v_{r2}(g) - v_{r3}(g)), i \neq r1 \neq r2$$

步骤 5 交叉操作,对第  $g$  代种群  $\{x_i(g)\}$  和其中间体  $\{v_{i(g+1)}\}$  之间做交叉操作

$$u_{j,i}(g+1) = \begin{cases} v_{j,i}(g+1), & \text{rand}(0,1) < CR \text{ or } j = j_{\text{rand}} \\ x_{j,i}(g), & \text{otherwise} \end{cases}$$

其中,  $CR$  为交叉概率,  $j_{\text{rand}}$  为  $[1, 2, \dots, m \times n]$  之间的整数。

步骤 6 选择操作,使用贪婪思想选择进入下一代的个体

$$x_i(g+1) = \begin{cases} u_i(g+1), & f(u_i(g+1)) \leq f(x_i(g)) \\ x_i(g), & \text{otherwise} \end{cases}$$

步骤 7 使用约束条件和式(2)对进入下一代的个体进行调整,约束条件如下:

$$\begin{aligned} \sum_i \lambda_{ij} &= m, \quad j = 1, 2, \dots, n \\ \sum_j \lambda_{ij} &= n, \quad i = 1, 2, \dots, m \\ \lambda_{ij} &> 0 \end{aligned}$$

步骤 8 进行上述操作至第  $t$  代后(这里取  $t = 20$ ),在第  $t$  代种群中选择适应度最大的个体即为所求的个体。

## 3 试验结果

### 3.1 模拟数据

为了验证本文提出重建方法的有效性,使用计算机模拟图像的产生过程,并利用模拟产生的图像进行射影重建。试验时,设定摄像机的内参为  $f_u = f_v = 1000, s = 0, u_0 = 320, v_0 = 240$ ,然后随机产生 100 个空间点,并通过改变摄像机的外参以得到 8 幅  $640 \times 480$  的图像。试验结果如图 1 所示,其中纵坐标是重投影误差,定义为

$$\sigma = \frac{1}{mn} \sum \sum \|P_i x_j - \lambda_{ij}\| \quad (8)$$

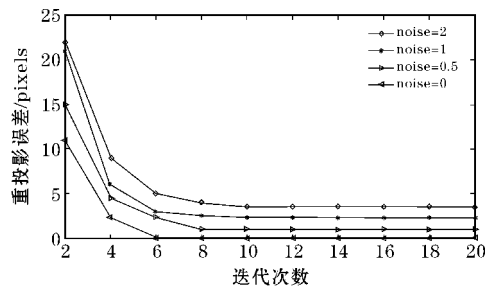


图1 重投影误差随迭代次数变化图

从图 1 可以看出,经过 10 次迭代后重投影误差趋



于稳定,这说明本文的重建方法具有良好的收敛性。在基于图像序列的重建算法中,图像的数量以及图像中的噪声也会对重建的结果产生重要的影响。为了验证本文的重建方法对图像数量以及噪声的敏感程度,使用计算机模拟了 20 幅图像,并添加了不同程度的噪声,其试验结果如图 2 所示。从图 2 可以看出不同程度的噪声在 10 次迭代后已趋于稳定,重投影误差随着图像数量的增加而增加,到达 8 次后开始趋于稳定。

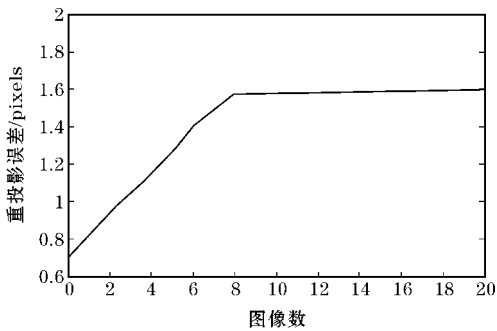


图2 重投影误差随图像噪声变化图

3.2 真实数据

为了验证本文方法的有效性及重建效果,使用牛津大学建筑数据集提供的图像序列,图像大小为 1024 × 768,其图像如图 3 所示。

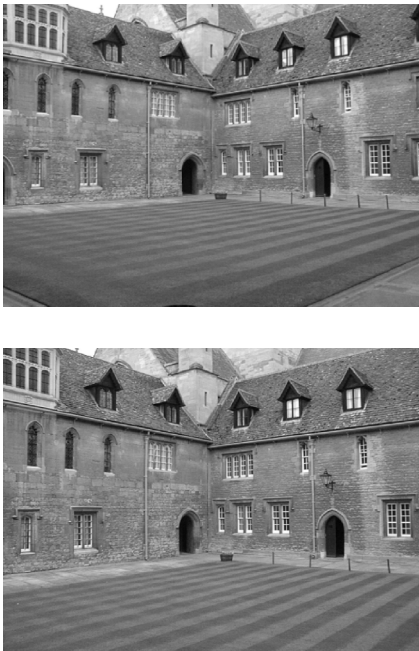


图3 牛津大学建筑数据集-Merton

为了直观地展现重建的结果并进行对比,将使用本文方法重建得到的三维点重新投影到原图的特征点上(图 4),从图 4 可以看出原图像中的特征点☆和重建得到的三维点○基本重合。同时,利用式(5)求得重投影误差为 0.54 个像素。从上述实验结果可以看

出本文的重建方法的重投影误差很小,具有很好的重建效果。

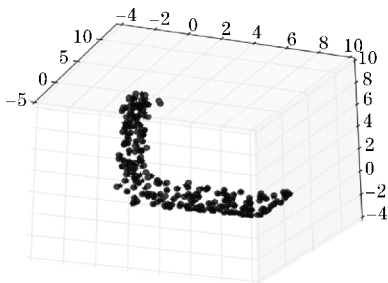


图4 重建后三维点和正确三维点对比

4 结束语

基于因式分解的射影重建,对射影深度矩阵的约束条件是能进行正确重建的关键。本文提出了一种基于射影深度矩阵行和列的范数的约束条件,并使用差分进化算法得到射影深度矩阵的一个估计,然后使用 SVD 分解得到场景的三维点和摄像机矩阵。经过模拟实验的对比表明,在不同的图像数量和噪声下,该方法经过 10 迭代就会收敛并且重投影误差小于图像噪声的 35%。结合真实图像的重建结果,可以得出本文的方法具有很好的收敛性,并且对噪声有一定抵抗性。

参考文献:

[1] Hartley R I,Zisserman A. Muitipie View Geometry in Computer Vision [ M ]. Cambridge University Press,2000.

[2] Sturm P,Triggs B. A Factorization-based Aigorithm for Multi-Image Projective Structure and Motion [ C]. Proc. of European Conference on Computer Vision, Cambridge, Engiand, 1996 :709-720.

[3] Triggs B. : Factorization methods for projective structure and motion. In: CVPR, IEEE Computer Society, 1996 :845-851.

[4] Mahamud S,Rebert M. Iterative Projective Recon- struction from Muitipie View[ C]. Proc. of IEEE Computer Vision and Pattern Recognition, 2000.

[5] A Heyden, R Berthilsson, G Sparr. An iterative factorization method for projective structure and motion from imagesequences [ J ]. Image Vision Comput. ,1999,17(13) :981-991.

[6] J Oliensis, R Hartley. Iterative extensions of the Sturm/Triggs algorithm: Convergence and noncon-

- vergence[J]. IEEE Trans. Pattern Anal. Mach. Intell. ,2007,29(12):2217–2233.
- [7] Sinkhorn,R. A relationship between arbitrary positive matrices and doubly stochastic matrices [J]. Ann. Math. Statist. 1964:876–879.
- [8] Sinkhorn,R. Diagonal equivalence to matrices with prescribed row and column sums[J]. The American Mathematical Monthly,1967:402–405.
- [9] 刘波,王凌,金以慧. 差分进化算法研究进展[J]. 控制与决策,2007,(7).
- [10] STORN R,PRICE K. Differential evolution;a simple and efficient adaptive scheme for globM optimization over continuous spaces[M]. Berkeley:International Computer Science Institute,1995.
- [11] Tomasi C,Kanade T. Shape and motion from image streams under orthography;a factorization method [J]. Int. J. Comput. ,1992,(9):137–154.
- [12] Triggs B,McLauchian P F,Hartley R I, et al. Bundie Adjustment a Modern Synthesis [J]. Vision Aigorithms:Theory and Practice,2000.