

针对摄像机矩阵估计的增强连续禁忌搜索方法

周果清, 王 庆

(西北工业大学计算机学院, 陕西西安 710072)

摘 要: 摄像机矩阵估计是机器视觉的一个重要问题. 在 2 范数误差代价函数模型下, 最小二乘法简单而有效, 但因误差代价函数非凸, 容易陷入局部最优. 在无穷范数误差代价函数模型下, 凸优化方法理论上可以获得全局最优, 但计算效率较低, 其计算耗时随着问题规模的增大而急剧增加. 现代优化论中的增强连续禁忌搜索(Enhanced continuous taboo search, ECTS)方法具有逃离局部最优的优良性质, 因此本文在 2 范数误差代价函数模型下提出一种针对摄像机矩阵估计的 ECTS 算法. 在 ECTS 置信区间序列构造及最大置信区间选择环节, 本文提出了一种非迭代的方法获取包含全局最优解的凸包. 在增强禁忌搜索环节, 本文提出了一种基于伪凸函数的候选解邻域构造方法. 同时, 给出了本文算法以概率 1 收敛于全局最优的理论证明. 对虚拟场景和真实场景的实验结果表明本文算法可以快速获取摄像机矩阵估计的全局最优解.

关键词: 多视几何; 摄像机矩阵; 全局最优; 禁忌搜索

中图分类号: TP319.7 **文献标识码:** A **文章编号:** 0372-2112 (2014)12-2422-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2014.12.013

A Resection Method Based on Enhanced Continuous Taboo Search

ZHOU Guo-qing, WANG Qing

(School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China)

Abstract: Resection is one of important issues in machine vision. Although L2 norm based least square method is reasonably fast, the globally optimal solution cannot be obtained theoretically due to its non-convexity of the objective function. Optimization using the L_∞ norm has been becoming an effective way to solve parameter estimation problems in multiview geometry. But the computational cost increases rapidly with the size of measurement data. In the paper, we propose a novel approach under the framework of enhanced continuous taboo search (ECTS) for resection in multiview geometry. ECTS is an optimization method in the domain of artificial intelligence, which has an interesting ability of covering a wide solution space by promoting the search far away from current solution and consecutively decreasing the possibility of trapping in the local minima. We propose the corresponding ways in the key steps of ECTS, diversification and intensification. We also present theoretical proof to guarantee the global convergence of search with probability one. Experimental results validate that the ECTS can obtain the global optimum effectively and efficiently. Potentially, the novel ECTS framework can be employed in many applications of multi-view geometry.

Key words: multiview geometry; resection; global optimization; taboo search

1 引言

摄像机矩阵估计是机器视觉中重要而基础的问题. 针对摄像机矩阵估计, 文献[1]给出了不同误差度量形式的代价函数, 包括: 代数距离、几何距离、重投影误差和 Sampson 误差. 常用的代价函数是基于 2 范数的重投影误差函数. 目前已经有求解该代价函数的方法, 比如 Newton 法和 Leven-berg-Marquardt 迭代法^[1], 光束法平差(Bundle Adjustment)算法^[2]. 该类方法能够获取较好

的摄像机矩阵估计结果. 但基于 2 范数的代价函数理论上是非凸的, 在图像点存在噪声的情况下该类方法容易陷入局部最优.

为了克服基于 2 范数误差代价函数方法的缺点, 文献[3]提出用无穷范数代替 2 范数重新构造误差代价函数. 在无穷范数框架下, 摄像机矩阵估计的误差代价函数在理论上是拟凸函数. 因此将误差代价函数的求解归结为一个最小化最大误差优化问题, 可采用分支限界策略结合二阶锥规划(Second-Order Cone Programing, SOCP)

进行求解.该方法在理论上可以确保获得全局最优解,但缺点是计算速度很慢而且对外点极其敏感^[4].文献[5]利用无穷范数代价函数是伪凸函数的性质对摄像机矩阵进行全局最优估计.文献[6]指出基于无穷范数的摄像机矩阵估计问题本质上是一个广义分式规划问题,提出基于无穷范数的牛顿迭代法实现摄像机矩阵的估计.文献[7]发现 SOCP 迭代过程中,相邻凸优化解之间存在数值相关性,提出一种类牛顿法实现了摄像机矩阵的估计.

另外,文献[8]提出首先在 2 范数误差代价函数模型下用非线性方法获取摄像机矩阵的估计结果,再对该结果进行全局最优判别,如判别成功则将其作为全局最优解,若不成功就调用基于无穷范数的方法进行求解.该方法只是关注了 2 范数结果的全局最优判别,对判别失败的点没有提出更好的解决方法.文献[9]也对类似的问题进行了研究和探讨.

现有获取摄像机矩阵估计的优化方法都属于经典优化论范畴,迄今为止利用现代优化论进行摄像机矩阵估计的方法还极少见报道.现代优化论中的禁忌搜索(Taboo search, TS)思想最早由文献[10]提出,它是对局部邻域搜索的一种扩展,是一种全局逐步寻优算法,是对人类智力过程的一种模拟. TS 算法通过引入一个灵活的存储结构和相应的禁忌准则来避免迂回搜索,并通过藐视准则来赦免一些被禁忌的优良状态,进而保证多样化的有效探索以最终实现全局优化.文献[11]提出了增强连续禁忌搜索(Enhanced Continuous Taboo Search, ECTS)方法用于连续最优化问题的求解.文献[12,13]进一步提出了量子禁忌搜索等改进算法.

本文在 ECTS 框架下,提出一种新的摄像机矩阵估计方法.该方法以简单线性方法的解作为初值,利用 ECTS 方法来获取摄像机矩阵的全局最优估计.在置信区间序列构造及最大置信区间选择环节,本文提出一种非迭代的方法获取包含全局最优解的凸包.在增强禁忌搜索环节,本文提出一种基于伪凸函数的候选解邻域构造方法.最后,本文证明了基于 ECTS 方法的摄像机矩阵估计方法的全局最优收敛性.相比现有的全局最优方法,本文提出的方法不但能获取摄像机矩阵的全局最优估计,而且能明显降低计算代价.

2 问题描述

根据多视几何理论,在齐次坐标系下,空间点和图像点之间存在以下关系:

$$\mathbf{u}_i = \mathbf{P} \mathbf{U}_i, \quad i = 1, 2, \dots, N \quad (1)$$

其中 \mathbf{P} 表示摄像机矩阵, $\mathbf{U}_i = (U_{i1}, U_{i2}, U_{i3}, 1)^T$ 表示空间点坐标, $\mathbf{u}_i = (u_{i1}, u_{i2}, 1)^T$ 表示空间点 \mathbf{U}_i 对应的图像

点坐标, i 是空间点编号.如果已知空间点坐标 \mathbf{U}_i 和对应的图像点坐标 \mathbf{u}_i ,则摄像机矩阵估计就是求解式(1)中的 \mathbf{P} .

因为 \mathbf{P} 是一个拥有 12 个元素的 3×4 矩阵,在忽略缩放因子的情况下它有 11 个自由度.理论上只要 6 组对应点,就可以通过求解线性方程完成摄像机矩阵估计.但由于图像点噪声不可避免,因此常用基于 2 范数的空间点重投影误差代价函数来估计摄像机矩阵.代价函数的形式为:

$$f(\mathbf{P}) = \sum_{i=1}^N \|\mathbf{u}_i - \mathbf{P} \mathbf{U}_i\|^2 \quad (2)$$

为了在后面更好的描述参数估计问题,我们在此将 \mathbf{P} 按行拉直成 $\mathbf{x} \in \mathbf{R}^n$, $n = 11$,即

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & 1 \end{bmatrix} \xrightarrow{\text{reshape}} \mathbf{x} = \hat{\mathbf{x}} = (x_1, x_2, \dots, x_{11})^T \quad (3)$$

简单变换后可以得到基于 2 范数的摄像机矩阵估计优化目标函数及约束为:

$$\begin{aligned} \min \sum_{i=1}^N f_i(\mathbf{x}) \\ \text{s.t. } f_i(\mathbf{x}) &= \frac{\sum_{j=1}^2 (\mathbf{a}_{ij}^T \mathbf{x} + \tilde{a}_{ij})^2}{(\mathbf{b}_i^T \mathbf{x} + \tilde{b}_i)^2} \\ \mathbf{b}_i^T \mathbf{x} + \tilde{b}_i &> 0 \\ \mathbf{x}, \mathbf{a}_{ij}, \mathbf{b}_i &\in \mathbf{R}^n \quad \text{and} \quad \tilde{a}_{ij}, \tilde{b}_i \in \mathbf{R} \end{aligned} \quad (4)$$

其中

$$\mathbf{a}_{i1} = [U_{i1}, U_{i2}, U_{i3}, 1, 0, 0, 0, 0, -u_{i1} U_{i1}, -u_{i1} U_{i2}, -u_{i1} U_{i3}]^T \quad (5)$$

$$\mathbf{a}_{i2} = [0, 0, 0, 0, U_{i1}, U_{i2}, U_{i3}, 1, -u_{i2} U_{i1}, -u_{i2} U_{i2}, -u_{i2} U_{i3}]^T \quad (6)$$

$$\mathbf{b}_i = [0, 0, 0, 0, 0, 0, 0, 0, U_{i1}, U_{i2}, U_{i3}]^T \quad (7)$$

$$\tilde{a}_{i2} = -u_{i2}, \quad \tilde{a}_{i1} = -u_{i1}, \quad \tilde{b}_i = 1 \quad (8)$$

$\mathbf{b}_i^T \mathbf{x} + \tilde{b}_i > 0$ 表示出现在图像上的空间点一般都应该在摄像机前方.

对应的,基于无穷范数的摄像机矩阵估计误差代价函数为:

$$\min_x \max_i f_i(\mathbf{x}) = \frac{\sum_{j=1}^2 (\mathbf{a}_{ij}^T \mathbf{x} + \tilde{a}_{ij})^2}{(\mathbf{b}_i^T \mathbf{x} + \tilde{b}_i)^2} \quad (9)$$

$$\text{s.t. } \mathbf{b}_i^T \mathbf{x} + \tilde{b}_i > 0, \quad i = 1, 2, \dots, N$$

3 摄像机矩阵估计增强连续禁忌算法

经典的 ECTS 方法包含 5 个步骤:①参数设置;②构造置信区间序列;③寻找最大置信区间;④最大置信区

间内的增强禁忌搜索;⑤最优解输出.其中,中间三个是 ECTS 算法的关键步骤.在构造置信区间序列时,ECTS 方法首先扫描整个解空间并检测出可能包含全局最优的置信空间,并将置信空间的中心存放在置信队列中.然后,根据给定的判别准则从置信队列中选出最大置信区间.当最大置信区间确定后,增强禁忌搜索开始执行.在增强禁忌搜索中,ECTS 方法通过不断压缩最大置信空间和不断简化近邻解集的结构实现全局最优解的获取.

3.1 构造置信区间序列及寻找最大置信区间

本小节讨论在摄像机矩阵估计中如何实现经典 ECTS 方法中最大置信空间的构造.换句话说,如何构造一个包含全局最优解的凸包.与经典的 ECTS 方法先通过不断迭代构造置信区间序列然后再选择最大置信区间不同,本文利用给定的初值直接构造最大置信区间.

首先用简单线性方法获取代价函数式(4)的初始解 \mathbf{x}_{init} .如果 \mathbf{x}_{opt} 是基于 2 范数方法的全局最优解,那么:

$$E_2(\mathbf{x}_{\text{opt}}) = \min_{\mathbf{x}} \sum_{i=1}^n f_i(\mathbf{x}_{\text{opt}}) \leq E_2(\mathbf{x}_{\text{init}}) = \delta^2 \quad (10)$$

其中 δ 是给定的误差限.从式(10)可以看出 $f_i(\mathbf{x}_{\text{opt}})$ 总是小于等于 δ^2 .因此,式(10)可以写成如下的形式:

$$\sqrt{\frac{(\mathbf{a}_{i1}^T \mathbf{x}_{\text{opt}} + \tilde{a}_{i1})^2 + (\mathbf{a}_{i2}^T \mathbf{x}_{\text{opt}} + \tilde{a}_{i2})^2}{(\mathbf{b}_i^T \mathbf{x}_{\text{opt}} + \tilde{b}_i)^2}} \leq \delta \quad (11)$$

进一步的,我们有

$$\left| \frac{(\mathbf{a}_{i1}^T \mathbf{x}_{\text{opt}} + \tilde{a}_{i1})}{\mathbf{b}_i^T \mathbf{x}_{\text{opt}} + \tilde{b}_i} \right| \leq \delta \quad \text{和} \quad \left| \frac{(\mathbf{a}_{i2}^T \mathbf{x}_{\text{opt}} + \tilde{a}_{i2})}{\mathbf{b}_i^T \mathbf{x}_{\text{opt}} + \tilde{b}_i} \right| \leq \delta \quad (12)$$

如果在不等式的两边同时乘以深度项 $\mathbf{b}_i^T \mathbf{x}_{\text{opt}} + \tilde{b}_i$,则对于有 N 个空间点与图像点对应的摄像机矩阵估计问题,我们可以得到总共 $4N$ 个关于 \mathbf{x}_{opt} 的线性约束.我们的目标是构造一个包含 \mathbf{x}_{opt} 的凸包,也就是寻找 $\mathbf{x}_{1 \min}, \mathbf{x}_{1 \max}, \dots, \mathbf{x}_{11 \min}, \mathbf{x}_{11 \max}$, 使它们满足:

$$\mathbf{x}_{1 \min} \leq \mathbf{x}_{1 \text{opt}} \leq \mathbf{x}_{1 \max}, \dots, \mathbf{x}_{11 \min} \leq \mathbf{x}_{11 \text{opt}} \leq \mathbf{x}_{11 \max} \quad (13)$$

因此,本文采用线性规划算法,通过求解下列不等式获得包含 \mathbf{x}_{opt} 的凸包.

$$\begin{aligned} -(\mathbf{a}_{i1}^T + \mathbf{b}_i^T \delta) \mathbf{x}_{\text{opt}} - \tilde{a}_{i1} - \tilde{b}_i \delta &\leq 0 \\ (\mathbf{a}_{i1}^T + \mathbf{b}_i^T \delta) \mathbf{x}_{\text{opt}} + \tilde{a}_{i1} - \tilde{b}_i \delta &\leq 0 \\ -(\mathbf{a}_{i2}^T + \mathbf{b}_i^T \delta) \mathbf{x}_{\text{opt}} - \tilde{a}_{i2} - \tilde{b}_i \delta &\leq 0 \\ (\mathbf{a}_{i2}^T + \mathbf{b}_i^T \delta) \mathbf{x}_{\text{opt}} - \tilde{a}_{i2} - \tilde{b}_i \delta &\leq 0 \end{aligned} \quad (14)$$

相比经典 ECTS 方法通过不断迭代来寻找最大置信空间,本文在给定初值的情况下通过线性规划直接获取了包含摄像机矩阵最优估计的凸包.因此,本文的最大置信区间的构造方法效率更高.

3.2 增强禁忌搜索

在经典 ECTS 算法中,增强禁忌搜索的流程是:构

造当前解的近邻解集,选择最佳近邻解,更新禁忌表和调整相关参数,迭代以上步骤直到满足终止条件.关键在于,如果当前解是局部最优解,算法必须找到一个可行的方向使得逃离局部最优成为可能.经典 ECTS 方法构造近邻解集的方法是以当前解为中心构造一个高维的超矩形,然后在该矩阵空间内随机选择一定数量的点,如果这些随机点不在禁忌表中,则用它们构建近邻解集.接下来,通过一定的判别准则挑选出近邻解集中的最优解作为算法的新最优解,即使它比当前的算法最优解要差.本文针对摄像机矩阵估计问题提出一种新的近邻解集构造方法,该方法不但能有效减少近邻解集元素的个数,而且能够确保 ECTS 算法以概率 1 收敛于全局最优解.

根据伪凸函数的定义^[14],式(4)中的 $f_i(\mathbf{x})$ 是伪凸函数.一般来说伪凸函数的驻点就是它的全局最优点.虽然 $f_i(\mathbf{x})$ 是伪凸函数,但是 $\max_i f_i(\mathbf{x})$ 却不是伪凸函数.即便如此, $\max_i f_i(\mathbf{x})$ 也具有一些良好的性质,其中最重要的性质是 minmaxKKT 条件.对于该结论的证明细节可以参考文献[15].

推论 1(minmaxKKT 条件) 令 $f(\mathbf{x}) = \max_i f_i(\mathbf{x})$, \mathbf{x}^* 是 $u^* = \min_{\mathbf{x} \in S} f(\mathbf{x})$, $S = \{\mathbf{x} | \mathbf{b}_i^T \mathbf{x} + \tilde{b}_i > 0, \forall i\}$ 的全局最优解,当且仅当存在 λ_i^* 满足:

$$\sum_{i=1}^m \lambda_i^* \nabla f_i(\mathbf{x}^*) = 0 \quad (15)$$

如果 $f_i(\mathbf{x}^*) = u^*$, 则 $\lambda_i^* \geq 0, i = 1, 2, \dots, m$; 如果 $f_i(\mathbf{x}^*) < u^*$, 那么 $\lambda_i^* = 0$, 且 $\sum_i \lambda_i^* = 1$.

推论 1 的几何含义是:对于任意一个 \mathbf{x} , 如果目标函数不是所有分量 $f_i(\mathbf{x})$ 的梯度都消失,那么在任一方向 d 上至少有一个 i 使得 $\nabla f_i(\mathbf{x})^T d \geq 0$, 也就是说在某一个方向至少有一个 $f_i(\mathbf{x})$ 是增大的.换句话说,只有在全局最优点梯度才会完全消失.根据推论 1, 本文在经典 ECTS 算法增强禁忌搜索迭代过程中仅在 $f_i(\mathbf{x}_k)$ (k 表示迭代的次数) 的梯度方向上选取随机点构造近邻解集.经典的 ECTS 算法中,近邻解集的选取方法是在以 \mathbf{x} 为中心的超矩形的每个维度方向上选取 p 个点 (p 是给定的参数,一般情况 $p \leq 10$). 对于摄像机矩阵估计问题,每次迭代产生的近邻解集元素个数是 $11 * p$. 本文方法产生的近邻解集元素个数是 p . 因此,本文方法的计算复杂度优于经典的 ECTS 算法.

本文近邻解集生成的具体方法是:用 \mathbf{z} 表示近邻解, \mathbf{z}^j 表示 \mathbf{z} 的第 j 维, $j = 1, 2, \dots, n$. \mathbf{z}^j 满足以 \mathbf{x}_k 为均值, σ 为标准方差的高斯分布. σ 的初始值为 1, 当迭代次数 k 增加时, $\sigma = d \times \sigma$ (d 是一个介于 0.997 与 0.999 之间的因子). 当 $\sigma < 10^{-4}$ 时, 将 σ 的值固定为 10^{-4} . 在

图像点存在噪声或者少量外点的情况下,推论 1 保证了近邻解集中包含着使 2 范数代价函数值变小的候选解,因此 ECTS 方法可不断的逼近全局最优解。

3.3 全局最优收敛证明

本小节我们讨论本文算法的全局最优收敛性。回顾本文的代价函数表达式(4),本质上是一个连续函数的全局优化问题,可以简写为:

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}) \quad (16)$$

其中 $\Omega = \{\mathbf{x} \in \mathbf{R}^n \mid x_{i_{\min}} \leq x_i \leq x_{i_{\max}}, x_{i_{\min}}, x_{i_{\max}} \in \mathbf{R}\}, i = 1, 2, \dots, n$.

本文提出的基于 ECTS 框架的摄像机矩阵估计方法满足记忆禁忌搜索(Memory Taboo Search, MTS)的约束条件. MTS 方法的简单流程如下:

步骤 1 产生代价函数的一个初始解 $\mathbf{x}_0 \in \Omega$, 并设置 $\mathbf{x}_0^* = \mathbf{x}_0$ 和 $k = 0$.

步骤 2 如果给定的终止条件满足,算法终止. 否则,用某种概率分布函数产生一个随机向量 \mathbf{y} .

步骤 3 如果 $f(\mathbf{y}) \leq f(\mathbf{x}_k^*)$, 则 $\mathbf{x}_{k+1}^* = \mathbf{y}$ 和 $\mathbf{x}_{k+1} = \mathbf{y}$; 否则,如果 $f(\mathbf{y}) \leq f(\mathbf{x}_k)$ 则 $\mathbf{x}_{k+1} = \mathbf{y}$, 否则如果 \mathbf{y} 不在禁忌表中, 则 $\mathbf{x}_{k+1} = \mathbf{y}$ 否则 $\mathbf{x}_{k+1} = \mathbf{x}_k$. 将 \mathbf{x}_{k+1} 投入禁忌表, 转到步骤 2.

为了讨论本文算法的收敛性,我们引入下面的定义^[15]:

定义 1 令 $\{\xi_m\}, (m = 0, 1, \dots)$ 是定义在概率空间的随机数序列. 如果 $\{\xi_m\}$ 以概率收敛于随机数 ξ , 当存在任意 $\epsilon > 0$, 有:

$$\lim_{m \rightarrow \infty} P\{|\xi_m - \xi| < \epsilon\} = 1 \quad (17)$$

定义 2 令 $\{\xi_m\}, (m = 0, 1, \dots)$ 是定义在概率空间的随机数序列. 如果 $\{\xi_m\}$ 以概率 1 (几乎必然) 收敛于随机数 ξ , 有:

$$P\{\lim_{m \rightarrow \infty} \xi_m = \xi\} = 1 \quad (18)$$

或者对于任意 $\epsilon > 0$, 有

$$P\{\bigcap_{m=1}^{\infty} \bigcup_{k \geq m} \{|\xi_k - \xi| \geq \epsilon\}\} = 1 \quad (19)$$

显而易见的,以概率 1 收敛要强于以概率收敛.

定理 1 (Borel-Cantelli 定理) 令 A_1, A_2, \dots 表示概

率空间的事件序列,集合 $P_k = Pr\{A_k\}$. 如果 $\sum_{n=1}^{\infty} P_k <$

∞ , 则 $Pr\{\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k\} = 0$. 如果 $\sum_{n=1}^{\infty} P_k = \infty$ 并且 A_k 是独

立的,那么 $Pr\{\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k\} = 1$.

假设基于 2 范数的摄像机矩阵估计代价函数 f 有全局最优解: $f^* = \min_{\mathbf{x} \in \Omega} f(\mathbf{x})$, 对于任意的 $\epsilon > 0$, 令 $D_0 = \{\mathbf{x} \in \Omega \mid f(\mathbf{x}) - f^* < \epsilon\}$, $D_1 = \Omega \setminus D_0$. 下列的引理 2 和推论 2 给出了可以确保 MTS 方法以概率 1 收敛于全

局最优.

引理 2^[16] 利用 MTS 方法求解连续全局优化问题(即式(15)), 当 $\mathbf{x}_k^* \in D_1$, 设 $\mathbf{x}_{k+1}^* \in D_1$ 的概率为 q_{k+1} , $\mathbf{x}_{k+1}^* \in D_0$ 的概率为 p_{k+1} . 如果 $\mathbf{y}^j, j = 1, 2, \dots, n$ 满足高斯分布, 那么 $q_{k+1} \leq c, c \in (0, 1)$.

推论 2^[16] 利用 MTS 方法求解连续全局优化问题式(15), 当 $\mathbf{x}_k^* \in D_1$, 如果 $\mathbf{y}^j, j = 1, 2, \dots, n$ 满足高斯分布, 则 $Pr\{\lim_{k \rightarrow \infty} f(\mathbf{x}_k^*) = f^*\} = 1$. 即 \mathbf{x}_k^* 以概率 1 收敛于全局最优. (相关证明见参考文献[16])

因此,如果给定摄像机矩阵估计的初值 \mathbf{x}_0 和任意小的误差限,本文提出的 ECTS 算法可以以概率 1 收敛于全局最优.

4 本文算法

本节给出在 ECTS 方法框架下,实现摄像机矩阵的全局最优估计的具体算法.

算法 1 基于 ECTS 的摄像机矩阵估计算法

输入:空间图像点坐标 U_i , 对应的图像点坐标 u_i , 预设的增强连续禁忌搜索的迭代次数 $K = 13$, 禁忌搜索的迭代步长 $t = 10^{-4}$, 算法终止的误差限 $\epsilon > 0$.

输出:摄像机矩阵的全局最优估计 \mathbf{x}_{opt} .

步骤 1:将简单线性性方法的解作为算法初始值 $\mathbf{x}_0^* = \mathbf{x}_0$. 设定 $k = 0$, 将禁忌表置为空.

步骤 2:调用 2.1 给出的线性规划算法构建包含全局最优点 \mathbf{x}_{opt} 的凸包 Ω .

步骤 3:如果 $|\sum_i f_i(\mathbf{x}_{k+1}^*) - \sum_i f_i(\mathbf{x}_k^*)| < \epsilon$ 或者 $k > K$, 算法终止.

步骤 4:构造近邻解集. 如果存在 $\nabla f_i(\mathbf{x}_k) > 0$, 则沿着某一个梯度 $\nabla f_i(\mathbf{x}_k)$ 的方向构造近邻候选解集 S 的元素 \mathbf{z} . 其中, \mathbf{z}^j 是 \mathbf{z} 的第 j 维, \mathbf{z}^j 满足以 \mathbf{x}_k 为均值, σ 为标准方差的高斯分布. σ 的初始值为 1, 当迭代次数 k 增加时, $\sigma = d \times \sigma$ (d 是一个介于 0.997 与 0.999 之间的因子). 当 $\sigma < 10^{-4}$ 时, 我们将 σ 的值固定为 10^{-4} . 如果 $\mathbf{z} \in \Omega$ 并且 \mathbf{z} 不在禁忌表中, 我们将其投入近邻候选解集 S 中.

步骤 5:对于每一个 $\mathbf{z}_s, s = 1, \dots, |S|$, 计算 $\mathbf{y} = \arg(\min_s (\max_i f_i(\mathbf{z}_s)))$.

步骤 6:如果 $\max_i f_i(\mathbf{y}) \leq \max_i f_i(\mathbf{x}_k^*)$ 则 $\mathbf{x}_{k+1}^* = \mathbf{y}, \mathbf{x}_{k+1} = \mathbf{y}$; 否则, 如果 $\max_i f_i(\mathbf{y}) \leq \max_i f_i(\mathbf{x}_k)$ 否则, 如果 \mathbf{y} 不在禁忌表中则 $\mathbf{x}_{k+1} = \mathbf{y}$ 否则 $\mathbf{x}_{k+1} = \mathbf{x}_k$. 将 \mathbf{x}_{k+1} 投入禁忌表.

步骤 7: $k = k + 1$. 转到步骤 3.

值得注意的是,在步骤 1 中,如果初始值是一个较差的线性方法的结果或者是一个任意值,则会设置置信区间的尺度变大,相应的会增加后续搜索迭代的次数,在一定程度上降低算法的效率. 但根据收敛性的理论分析,它并不影响算法的收敛.

5 实验与讨论

为了验证本文算法的精度和速度,我们采用虚拟场景和真实场景进行摄像机矩阵估计. 虚拟场景来自

Lund 大学的 Linfinity-1.0 数据集^[17], 真实场景来自于 Oxford 大学^[18]的 VGG 数据集. 本文对比算法是文献[4]的基于无穷范数的全局最优算法(Bisect-I), 它被认为是当前最好的全局最优算法. 本文的实验环境是标准 PC 机(Intel P8600 CPU, 6GB 内存, Matlab2010a, Win7 64bits). 本文用 Matlab 的 profile 命令来记录时间和进行性能对比. 基于无穷范数的对比算法 Bisect-I 中的 SOCP 问题求解使用的是 SUDUMI 工具包, 本文算法中使用的线性规划工具包是 MOSEK.

5.1 虚拟场景实验结果

首先, 我们用 Linfinity-1.0 生成 N 个空间点及对应的图像点坐标. 空间点位于一个给定的立方体内, 图像点坐标归一化到了区间 $[-1, 1]$ 并加入标准方差为 0.01 的高斯噪声. 我们比较了两种算法在不同空间点数量下的速度差异(表 1)、不同空间点数量下的平均重投影误差(图 1). 然后, 我们生成了 1000 个空间点以及对应的图像坐标, 在图像坐标中加入了不同水平的高斯噪声, 比较了相同空间点数量不同噪声水平下两种算法的平均重投影误差(图 2).

表 1 不同数量空间点下算法运行速度比较

空间点数量 /个	耗时/秒		速度提升 /倍
	Bisect-I 算法	本文算法	
10	4.2865	0.0583	71.8
100	5.4075	0.2102	25.7
300	9.3340	0.8087	11.5
500	9.6668	0.8856	10.9
700	14.4366	1.3857	10.4
1000	17.7806	1.7039	10.4

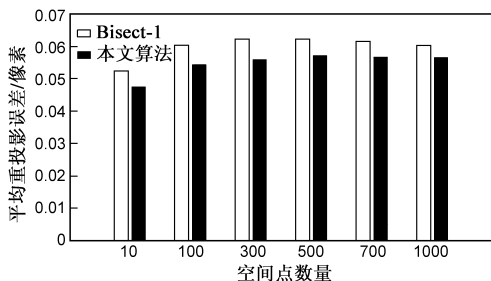


图 1 相同噪声水平不同空间点数量下平均重投影误差比较

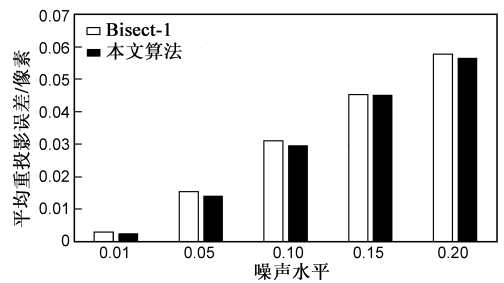


图 2 相同空间点数量不同噪声水平下平均重投影误差比较

5.2 真实场景实验结果

我们选取了 VGG 数据集的 4 组数据作为实验数据. 表 2 给了 2 种算法的计算耗时比较和平均重投影误差比较. 由于实验数据集中给出了摄像机的内参数, 因此本文在获得摄像机矩阵后通过矩阵分解可以进一步获得摄像机在空间的位置和朝向. 因为 University library 数据和 Oxford 数据的摄像机位置估计结果比较难展示, 本文通过图 3 和图 4 分别比较了 Model house 和 Wadham 两组数据的摄像机位置估计结果. 图中, 黑色矩形框表示真实的摄像机位置, 绿色矩形框表示 Bisect-I 算法获得的摄像机位置, 红色矩形框表示本文算法获得的摄像机位置.

5.3 分析与讨论

下面从摄像机矩阵估计的计算速度、精度和抗噪声能力三个方面对实验结果进行分析和讨论.

在计算速度方面, 虚拟场景实验结果表明在空点数量少于 100 个点时, 本文算法的速度提升在 20 倍以上, 空间点数量少于 1000 点时本文速度提升在 10 倍以上(表 1). 真实场景实验结果中, 除了 Wadham 数据外, 其他三组数据的速度提升都在 20 倍以上(表 2). Wadham 数据提升速率较低的原因是图像点坐标中存在的少量外点导致了禁忌搜索迭代次数的增加, 从而增加了计算时间. 综合虚拟场景和真实场景的实验结果可以看出本文算法相比 Bisect-I 算法在摄像机矩阵估计的速度上有明显提升.

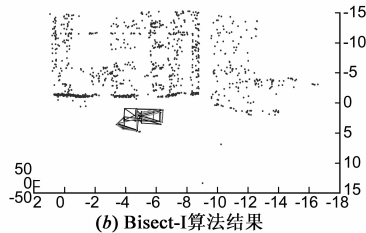
在计算精度方面, 针对虚拟场景, 从相同噪声水平不同空间点数量的平均重投影误差比较可以看出本文

表 2 真实数据计算精度和速度比较

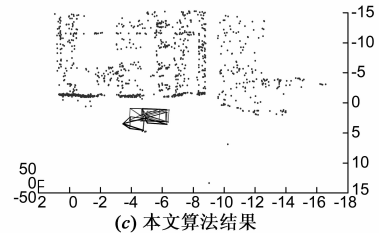
数据集	图像数量	空间点数量	平均重投影误差/像素		计算耗时/秒		速度提升 /倍
			Bisect-I	本文算法	Bisect-I	本文算法	
Model house	10	672	0.1000	0.0368	18.0754	0.5111	35.4
Wadham	5	1331	0.0091	0.0084	13.6696	1.9669	6.9
Library	3	667	0.0173	0.0124	7.8832	0.3587	22.0
Oxford	11	737	0.0615	0.0305	21.4552	0.8945	24.0



(a) 2D图像



(b) Bisect-I算法结果

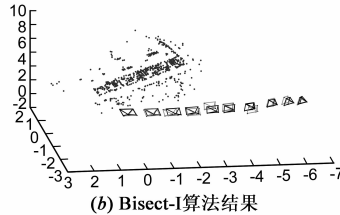


(c) 本文算法结果

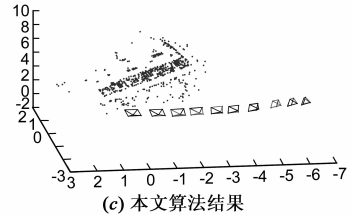
图3 Wadham数据摄像机矩阵估计结果对比



(a) 2D图像



(b) Bisect-I算法结果



(c) 本文算法结果

图4 Model house数据摄像机矩阵估计结果对比

算法与对比算法的精度相当,甚至略好一些(图2)。真实实验数据中,本文算法的计算精度都高于对比算法的计算精度(表2)。因此,本文算法的计算精度高于对比算法,这验证了本文算法可以以概率1收敛于全局最优。

在相同空间点数量,不同噪声水平的虚拟场景实验中,本文算法的平均重投影误差与对比算法基本相当,个别组次略好于对比算法。这表明本文算法的抗噪声能力和对比算法相当。在不同的噪声水平下,本文算法依然可以以概率1收敛于全局最优。

6 总结与展望

本文提出了一种针对摄像机矩阵估计的增强连续禁忌搜索算法。该算法以基于2范数的简单算法的估计结果为初值,通过构造一个包含全局最优的凸包,然后利用增强连续禁忌搜索策略实现对以概率1收敛全局最优解的获取。实验结果表明,同基于无穷范数的全局最优算法相比本文算法可以获得同样的计算精度,但在计算效率上有明显提高。对于多视几何中具有类似误差代价函数的其他问题,如单应估计、多视图三维重建等,本文提出的算法也能获取以概率1收敛的全局最优解。

参考文献

- [1] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press, 2004. 86 - 98.
- [2] Triggs B, McLauchlan P, et al. Bundle Adjustment a Modern Synthesis[M]. Berlin Heidelberg: Springer, 2000. 298 - 372.
- [3] Ke Q, Kanade T. Quasiconvex optimization for robust geometric reconstruction[A]. IEEE International Conference on Com-

puter Vision[C]. Beijing: IEEE, 2005. 986 - 993.

- [4] Kahl F. Multiple view geometry and the L_∞ -norm[A]. IEEE International Conference on Computer Vision[C]. Beijing: IEEE, 2005. 1002 - 1009.
- [5] Olsson C, Eriksson A, et al. Efficient optimization of L_∞ -problems using pseudoconvexity[A]. IEEE International Conference on Computer Vision[C]. Rio de Janeiro: IEEE, 2007. 1 - 8.
- [6] Agarwal S, Snavely N, et al. Fast algorithms for L_∞ -Norm problems in multiview geometry[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Alaska: IEEE, 2008. 1 - 8.
- [7] Dai Z, Wu Y, Zhang F, et al. A novel fast method for L_∞ problems in multiview geometry[A]. Computer Vision-ECCV 2012[C]. Florence: Springer Berlin Heidelberg, 2012. 116 - 129.
- [8] Hartley R, Kahl F, et al. Verifying global minima for L_2 minimization problems in multiple view geometry[J]. International Journal of Computer Vision, 2013, 101(2): 288 - 304.
- [9] 张闯, 迟健男, 张朝晖, 王志良. 一种基于立体视觉的视线估计方法[J]. 电子学报, 2010, 38(05): 1008 - 1014. Zhang C, Chi J N, Zhang Z H, Wang Z L. A novel gaze estimation method based on stereo vision[J]. Acta Electronica Sinica, 2010, 38(05): 1008 - 1014. (in Chinese)
- [10] Glover F. Tabu search methods in artificial intelligence and operations research[J]. ORSA Artificial Intelligence, 1987, 1(2): 6.
- [11] Chelouah R, Siarry P. Tabu search applied to global optimization[J]. European Journal of Operational Research, 2000, 123(2): 256 - 270.
- [12] 陈晓峰, 姜慧研. 量子禁忌搜索算法的研究[J]. 电子学报, 2013, 41(11): 2161 - 2166. Chen X F, Jiang H Y. Research of quantum tabu search algorithm[J]. Acta Electronica Sinica, 2013, 41(11): 2161 -

2166. (in Chinese)

- [13] 许鹏飞, 苗启广, 李伟生, 张军英. 基于函数复杂度的自适应模拟退火和禁忌搜索新算法[J]. 电子学报, 2012, 40(6): 1218 – 1222.
 Xu P F, Miao Q G, Li W S, Zhang J Y. Adaptive simulated annealing algorithm and tabu search algorithm based on the function complexity [J]. Acta Electronica Sinica, 2012, 40(06): 1218 – 1222. (in Chinese)
- [14] Boyd S, Vandenberghe L. Convex Optimization [M]. Cambridge Univ Press, 2004. 95 – 104.
- [15] Olsson C, Kahl F. Generalized convexity in multiple view geometry [J]. Journal of Mathematical Imaging and Vision, 2010, 38(1): 35 – 51.
- [16] Ji M, Tang H. Global optimizations and tabu search based on memory [J]. Applied Mathematics and Computation, 2004, 159(2): 449 – 457.
- [17] Kahl F. Multiple View Geometry and the L-infinity Norm [DB/OL]. <http://www.maths.lth.se/matematiklth/personal/fredrik/download.html>, 2013.
- [18] Werner T, Zisserman A. New techniques for automated architecture reconstruction from photographs proc [A]. 7th European Conference on Computer Vision [C]. Copenhagen, Denmark; ECCV, 2002. 541 – 555.

作者简介



周果清 男, 1980 年生, 四川资中人. 2003 年和 2013 年于西北工业大学获工学学士和博士学位. 现为西北工业大学计算机学院讲师. 研究兴趣包括: 计算机视觉、计算摄影术和全局优化方法.

E-mail: zhouguoqing@nwpu.edu.cn



王庆(通讯作者) 男, 1969 年生, 陕西西安人. 博士、博士生导师, 中国计算机学会高级会员、IEEE 会员、ACM 会员. 1991 年和 2000 年分别在北京大学、西北工业大学获理学学士和工学博士学位. 现为西北工业大学计算机学院教授、陕西省语音与图像信息处理重点实验室副主任, 主要从事计算机视觉和模式识别等方面的研究工作. E-mail: qwang@nwpu.edu.cn