

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **5**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012 年 5 月 第 17 卷 第 5 期(总第 193 期)

目 次

综述

中国图像工程:2011 章毓晋(603)

植物叶片表面质感建模与真实感绘制研究进展 田原,赵春江,陆声链,郭新宇(613)

图像处理和编码

高位有效位概率算术解码的图像密写改进算法 马丽红,吕先明,高小满(621)

滑动平均和改进权重函数的快速非局部平均图像去噪算法 熊波,尹周平(628)

提升小波的同态滤波在图像烟雾弱化中的应用 范有臣,李迎春,韩意,张来线(635)

图像分析和识别

基于彩色模型的重构标记分水岭分割算法 张桂梅,周明明,马珂(641)

均衡化特征匹配的非刚体细胞形态跟踪 陈莹,艾春璐(648)

图像斑状特征位置与尺寸的自动检测 王志衡,刘红敏(656)

改进极化白化滤波的边缘检测 邓少平,张继贤,李平湘,黄国满(665)

联合特征在行人检测中的应用 杨阳,杨静宇(671)

Gabor 相位特征的人脸光照不变量提取 范春年,张福炎(676)

图像理解和计算机视觉

保持几何特征的自适应弹性二次曲线模型 蒋建国,郝世杰,郭艳蓉,詹曙,李鸿(682)

局部颜色映射的彩色夜视算法 钱小燕,张天慈,王帮峰,黄圣国(689)

结合分支定界法和线性规划的摄像机位姿估计 马文娟(694)

金字塔评分改进主方向模板匹配的实时目标检索 洪朝群,朱建科,李娜,卜佳俊,陈纯(700)

计算机图形学

类曲率在曲线相似性判定中的应用 于昊,赵乃良,陈小雕(707)

虚拟现实与增强现实

人机系统中视域评估的可视化 李倩,吉晓民,林文周(715)

遥感图像处理

小波方向子带偏微分方程遥感图像去噪 王相海,李放,王爽(721)

遥感图像理想均衡化及图像质量定量评价 孟天佑,汪云甲(729)

地理信息技术

点要素扇形缓冲区的设计与应用 侯景伟,孔云峰,张迪,吕可文(740)

中国图象图形学报

刊名题字: 宋 健 月刊(1996 年创刊) 第 17 卷 第 5 期 2012 年 5 月 16 日出版

主管单位	中国科学院	Superintended by	Chinese Academy of Sciences
主 办	中国科学院遥感应用研究所 中国图象图形学学会 北京应用物理与计算数学研究所	Sponsored by	Institute of Remote Sensing Application, CAS China Society of Image and Graphics Institute of Applied Physics and Computational Mathematics
主 编	李小文	Chief editor	LI Xiaowen
编辑出版	《中国图象图形学报》编辑出版委员会 北京 9718 信箱 邮编 100101 电子信箱:jig@irsa.ac.cn 电话:010-68407995 010-82614429 网 址:www.cjig.cn	Editor, Publisher	Editorial and Publishing Board of Journal of Image and Graphics (P. O. Box 9718, Beijing 100101, China) E-mail:jig@irsa.ac.cn
印刷装订	北京北林印刷厂	Distributed by	Beijing Bureau for Distribution of Newspapers and Journals
广告经营许可证	京朝工商广字第 0346 号	Domestic	All Local Post Offices in China
总 发 行	北京报刊发行局	Foreign	China International Book Trading Corporation (P. O. Box 399, Beijing 100044, China)
订 购	全国各地邮局	Printed by	Beijing Beilin Printing House
国外发行	中国国际图书贸易总公司 (中国国际书店) (北京 399 信箱 邮编 100044)		

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 5 May 2012

Contents

Review

- Image engineering in China: 2011 Zhang Yujin(603)
- Advances in appearance modeling and photorealistic rendering of plant leaves
..... Tian Yuan, Zhao Chunjiang, Lu Shenglian, Guo Xinyu(613)

Image Processing and Coding

- Most significant bits probability arithmetic decoding for improved image steganography
..... Ma Lihong, Lv Xianming, Gao Xiaoman(621)
- Fast non-local means for image de-noising on moving average and modified weight function Xiong Bo, Yin Zhouping(628)
- Weakening of smoke for homomorphic filtering Fan Youchen, Li Yingchun, Han Yi, Zhang Laixian(635)

Image Analysis and Recognition

- Image segmentation algorithm for reconstruction labeling watershed in color space Zhang Guimei, Zhou Mingming, Ma Ke(641)
- Non-rigid cell contour tracking method for balanced feature matching Chen Ying, Ai Chunlu(648)
- Automatically detecting position and size of blob features in images Wang Zhiheng, Liu Hongmin(656)
- Improved polarimetric whitening filter for edge detection Deng Shaoping, Zhang Jixian, Li Pingxiang, Huang Guoman(665)
- Pedestrian detection based on compound feature Yang Yang, Yang Jingyu(671)
- Illumination invariant extraction on Gabor phase Fan Chunnian, Zhang Fuyan(676)

Image Understanding and Computer Vision

- Adaptive geometrical-feature-preserving elastic quadratic wire model
..... Jiang Jianguo, Hao Shijie, Guo Yanrong, Zhan Shu, Li Hong(682)
- Color night vision algorithm based on local color mapping ... Qian Xiaoyan, Zhang Tianci, Wang Bangfeng, Huang Shengguo(689)
- Camera pose estimation using branch and bound method with linear programming Ma Wenjuan(694)
- Real-time object retrieval with dominant orientation template matching improved by pyramid scoring
..... Hong Chaoqun, Zhu Jianke, Li Na, Bu Jiajun, Chen Chun(700)

Computer Graphics

- Quasi-curvature and its application in similarity measurement of curves
..... Yu Hao, Zhao Nailiang, Chen Xiaodiao(707)

Virtual Reality and Augmented Reality

- Visualization of the visual range assessment in man-machine system Li Qian, Ji Xiaomin, Lin Wenzhou(715)

Remote Sensing Image Processing

- Remote sensing image de-noising on partial differential equation in wavelet directional subband
..... Wang Xianghai, Li Fang, Wang Shuang(721)
- Ideal equalization of remote sensing images and quantitative assessment of image quality Meng Tianyou, Wang Yunjia(729)

Geoinformatics

- Design and applications of sector buffers for point feature
..... Hou Jingwei, Kong Yunfeng, Zhang Di, Lv Kewen(740)

中图法分类号: TP391.4 文献标志码: A 文章编号: 1006-8961(2012)-05-0694-06

论文引用格式: 马文娟. 结合分支定界法和线性规划的摄像机位姿估计[J]. 中国图象图形学报, 2012, 17(5): 694-699

结合分支定界法和线性规划的摄像机位姿估计

马文娟

浙江理工大学数字媒体技术系, 杭州 310018

摘要: 介绍了一种新的利用对应点估计摄像机位姿的算法。通常情况下, 摄像机位姿估计可以转化为一个最优化问题, 现有算法将问题转换成一个序列二阶锥规划问题, 通过对旋转矩阵所在空间进行分支定界搜索来求取全局最优解。对现有算法进行改进, 通过将二阶锥约束松弛为线性约束, 提出了一种结合分支定界法和线性规划方法的全局优化算法。该算法不仅能够求得全局最优解, 而且算法速度较现有算法提高了一倍以上。最后通过模拟数据和真实数据对该算法进行了验证, 结果表明了该算法的准确性和高效性。

关键词: 位姿估计; 全局最优化; 分支定界; 线性规划; 二阶锥规划

Camera pose estimation using branch and bound method with linear programming

Ma Wenjuan

Department of Digital Media Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China

Abstract: In this paper, we introduce a new algorithm for estimating camera pose from point correspondences. Generally, the camera pose problem could be formulated as an optimization problem. The current methods transform the problem into a set of second order cone programming (SOCP) feasibility problems which obtain the global optimal solution by searching the rotation space. In this paper, by relaxing the second-order cone constraints to linear constraints, we propose an improved method that combines branching and bounding with linear programming (LP). Our method cannot only get the global optimal pose but also runs two times faster than the current best method. Our approach has been tested on a number of synthetically generated and real data sets, and the results demonstrate the accuracy and the high speed of the proposed method.

Key words: pose estimation; global optimization; branch and bound; linear programming; second order cone programming

0 引言

摄像机位姿估计是计算机视觉领域和摄影测量学广泛研究的课题之一^[1], 其研究最早可追溯到1841年。这类问题不仅是几何视觉中的一个重要基本问题, 同时也是许多计算机视觉应用问题的核

心内容, 包括运动分割^[2], 物体识别^[3,4]和模型校准^[5]等。在过去几十年中, 已经针对摄像机位姿估计问题提出了许多算法^[6-10]。

摄像机的位姿是由摄像机相对于世界坐标系的方位决定的, 包括旋转矩阵 R 和平移矢量 C 。摄像机的位姿估计问题就是给定一系列3维点以及对应的图像点, 确定 R 和 C 。在理想情况下, 即不存在噪

收稿日期: 2011-06-15; 修回日期: 2011-10-24

基金项目: 浙江理工大学科研启动基金项目(1004838-Y)

第一作者简介: 马文娟(1982—), 女, 讲师, 2010年于上海交通大学获模式识别与智能系统专业博士学位, 主要研究方向为数字图像处理 and 计算机视觉。E-mail: mawj203@163.com

声时,可以采用直接线性变换(DLT)算法求得摄像机位姿^[1]。这种方法是利用线性算法求出初始解,然后进行非线性优化。虽然该方法比较有效,但是由于问题的非凸性而往往陷入局部最优。为了求得全局最优解,一些基于重投影误差的 L_∞ 范数的全局优化算法已经提出^[6-7],然而,这些算法是在假设摄像机的旋转矩阵 $\mathbf{R}$ 已知的情况下,仅仅确定摄像机的位置 $\mathbf{C}$ 。

文献[8-9]提出了摄像机位姿估计问题的第一种全局优化算法,这种方法是基于全局优化理论,在重投影误差的二范数下采用凸松弛结合分支定界法求解,但其缺点是收敛速度很慢。在文献[10-11]中,作者利用经典几何的一个简单结论推导 L_∞ 范数的必要条件,并将其应用到分支定界方法中进行寻优和检测外围点。该方法的缺点是分支定界方法的寻优区域很难定义。另外一种求取摄像机最优位姿的方法在文献[12-13]中提出,这种方法是在旋转空间寻优,通过解一系列的二阶锥规划问题得到最优解,其寻优空间是可以确定的,并且收敛速度比文献[8-9]的方法快得多。

本文将文献[12-13]中的方法进行了改进,通过将二阶锥约束松弛为线性约束,提出了一种结合分支定界法和线性规划法的全局优化方法。与文献[12-13]中的算法相比,本文方法在保证全局最优解的同时具有更快的求解速度。

1 基础知识

首先介绍一些基本概念和算法,包括凸优化、二阶锥规划和线性规划,以及分支定界算法。这些是本文算法的基础。

1.1 凸优化、二阶锥规划和线性规划

凸优化(CP)^[14]是近年来发展非常迅速的一类优化问题,其最大特点在于已有十分有效的求解方法对其求解,同时能够保证求得的是全局最优解。

如果函数 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 的定义域为一个凸集并且对于任意 $\mathbf{x}, \mathbf{y} \in \text{dom}f$,以及 $\theta(0 \leq \theta \leq 1)$,存在:

$$f(\theta \mathbf{x} + (1 - \theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1 - \theta)f(\mathbf{y}) \quad (1)$$

则将 f 称为凸函数。一个凸优化问题定义为

$$\begin{aligned} \min \quad & f_0(\mathbf{x}) \\ \text{s. t.} \quad & f_i(\mathbf{x}) \leq 0 \quad i = 1, \dots, m \end{aligned} \quad (2)$$

式中 $f_0: \mathbf{R}^n \rightarrow \mathbf{R}$ 。如果一个凸优化问题的形式为:

$$\begin{aligned} \min \quad & \mathbf{f}^T \mathbf{x} \\ \text{s. t.} \quad & \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\| \leq \mathbf{c}_i^T \mathbf{x} + d_i \quad i = 1, \dots, m \end{aligned} \quad (3)$$

$$\mathbf{F} \mathbf{x} = \mathbf{g}$$

式中, $\mathbf{A}_i \in \mathbf{R}^{n_i \times n}$,目标函数 $\mathbf{F} \in \mathbf{R}^{p \times n}$,则该问题被称为二阶锥规划问题(SOCP)。如果SOCP问题中没有目标函数,则称其为一个SOCP可行问题,形式为

$$\begin{aligned} \text{find} \quad & \mathbf{x} \\ \text{s. t.} \quad & \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\| \leq \mathbf{c}_i^T \mathbf{x} + d_i \quad i = 1, \dots, m \\ & \mathbf{F} \mathbf{x} = \mathbf{g} \end{aligned} \quad (4)$$

另一类特殊的凸优化问题称为线性规划(LP),形式为

$$\begin{aligned} \min \quad & \mathbf{f}^T \mathbf{x} \\ \text{s. t.} \quad & \mathbf{A} \mathbf{x} \leq \mathbf{b} \end{aligned} \quad (5)$$

1.2 分支定界算法

分支定界算法是非凸问题寻优的常用算法。其基本思想是对有约束条件的最优化问题的所有可行解(数目有限)空间进行搜索。该算法在具体执行时,把全部可行的解空间不断分割为越来越小的子集(称为分支),并为每个子集内的解计算一个下界或上界(称为定界)。在每次分支后,对凡是界限超出已知可行解的那些子集不再做进一步分支。这样,解的许多子集就可以不予考虑了,从而缩小了搜索范围。这一过程一直进行到找出可行解为止,该可行解不大于任何子集的界限。因此这种算法一般可以求得最优解。

2 位姿估计的线性松弛

2.1 问题描述

首先我们给出摄像机位姿估计问题的数学描述。设给定一系列的3维点 $\mathbf{X}_i(i = 1, \dots, n)$ 和对应的图像点 $\mathbf{v}_i$,要估计摄像机相对于世界坐标系的旋转矩阵 $\mathbf{R}$ 和摄像机的位置 $\mathbf{C}$ 。如文献[13]所述,在重投影误差的 L_∞ 范数度量下,位姿估计问题可以描述为minmax优化问题,即

$$\min_{\mathbf{R}, \mathbf{C}} \max_i \angle(\mathbf{v}_i, \mathbf{R}(\mathbf{X}_i - \mathbf{C})) \quad (6)$$

式中, $\angle(\cdot, \cdot)$ 代表两个矢量之间的夹角,摄像机内部参数被假设为单位矩阵。为了方便,图像点用球面坐标表示,这样式(6)中的图像坐标 $\mathbf{v}_i$ 表示摄像机中心到3维点的方向矢量,即

$$\mathbf{v}_i = \frac{\mathbf{R}(\mathbf{X}_i - \mathbf{C})}{\|\mathbf{R}(\mathbf{X}_i - \mathbf{C})\|} \quad (7)$$

估算的图像点用 $\hat{v}_i(\theta)$ 来表示, v_i 和 $\hat{v}_i(\theta)$ 之间的距离用球面坐标表示就是两矢量之间的夹角 $\angle(v_i, \hat{v}_i(\theta))$, 应该小于某个阈值 ε , 即

$$\frac{\|v_i \times \hat{v}_i(\theta)\|}{v_i^T \hat{v}_i(\theta)} \leq \tan(\varepsilon) \quad (8)$$

等价于

$$\|v \times \hat{v}(\theta)\| - \varepsilon' v^T \hat{v}(\theta) \leq 0 \quad (9)$$

式中, $\varepsilon' = \tan(\varepsilon)$, 式(9)符合二阶锥约束的形式, 但是由于矢量 $R(X_i - C)$ 不能表示成未知量 R 和 C 的线性形式, 所以该问题是非凸的, 不能用二阶锥规划求解。然而, 当旋转矩阵 R 已知, 该问题就是可解的。因此, 文献[13]通过在旋转空间搜索求得最优的 R 进行摄像机位姿估计。该方法是采用分支定界法搜索旋转空间。旋转量用轴角法表示, 那么所有的旋转量就可以用半径为 π 的球内的点表示。由于球不容易分割, 可以把球松弛为球的外切立方体, 立方体很容易被分割成小的立方体。对于包含球内任意点的小立方体, 对应于其中心的旋转量用 $\bar{R}$ 表示, 半径用 r_d 表示, r_d 为小立方体内点与中心 $\bar{R}$ 的最大距离, 可推导 r_d 满足不等式 $r_d \leq \sqrt{3}\sigma$, 其中 σ 等于小立方体边长的一半。在小立方体内判断是否存在 R 满足可行性问题

$$\begin{aligned} & \text{给定 } X_i, \bar{R}, \varepsilon_{\min} \\ & \text{寻找 } C, \delta r \in [-\sigma, \sigma]^3 \\ & \text{满足 } \angle(v_i, \bar{R}(X_i + \delta r \times X_i - C)) < \\ & \quad \varepsilon_{\min} + 3\sigma^2/2 \end{aligned} \quad (10)$$

式中, δr 是 $\delta R = \bar{R}^T R$ 的轴角表示形式, $\varepsilon_{\min}$ 是误差阈值。若可行性问题式(10)无解, 那么该小立方体被舍弃, 否则将小立方体继续分割, 直到不能再分割为止。文献[13]的方法将原优化问题转换成求解一系列二阶锥可行性问题。一般情况下二阶锥规划比线性规划速度慢得多, 那么用线性规划形式描述摄像机位姿估计问题, 算法速度会得到很大提高。基于这种思想, 本文提出了将分支定界法和线性规划方法相结合的新算法。

2.2 线性松弛方法

如2.1节所述, 文献[13]中的方法采用球面坐标表示图像点, 约束满足二阶锥规划的形式。本文用平面坐标表示图像点, 将约束表示成线性规划形式。图像球面和图像平面的关系如图1所示。

图1中, 由于摄像机内部参数矩阵已知, 因此图像点的坐标可以通过内参数矩阵归一化, 进而假设

摄像机焦距为1。给定图像球面上一点 $p_s = [u_x \ u_y \ u_z]^T$, 那么其在图像平面上的对应点为 $p = \frac{1}{u_z} [u_x \ u_y]^T$ 。令 $q_s = (x \ y \ z)^T$ 为图像球面上另外的任意一点, 它在图像平面上的对应点为 $q = \frac{1}{z} [x \ y]^T$ 。我们假定 q_s 和 p_s 的夹角为 α , 那么对于所有的点 $q_s = [x \ y \ z]^T$ 应满足

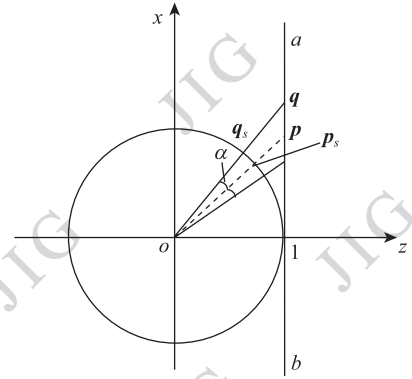


图1 图像球面(圆o)和图像平面(直线ab)的2维图

Fig. 1 The two-dimensional diagram of the image sphere (circle o) and the image plane (line ab)

$$\begin{cases} xu_x + yu_y + zu_z = \cos \alpha \\ x^2 + y^2 + z^2 = 1 \end{cases} \quad (11)$$

p 和 q 之间的距离可以表示为

$$\begin{aligned} d &= \left\| \frac{1}{z} \begin{bmatrix} x \\ y \end{bmatrix} - \frac{1}{u_z} \begin{bmatrix} u_x \\ u_y \end{bmatrix} \right\|^2 = \\ &= \frac{x^2}{z^2} - \frac{2xu_x}{zu_z} + \frac{u_x^2}{u_z^2} + \frac{y^2}{z^2} - \frac{2yu_y}{zu_z} + \frac{u_y^2}{u_z^2} = \\ &= \frac{x^2 + y^2}{z^2} - \frac{2(xu_x + yu_y)}{zu_z} + \frac{u_x^2 + u_y^2}{u_z^2} = \\ &= \frac{1 - z^2}{z^2} - \frac{2(\cos \alpha - zu_z)}{zu_z} + \frac{1 - u_z^2}{u_z^2} = \\ &= \frac{1}{z^2} - \frac{2\cos \alpha}{zu_z} + \frac{1}{u_z^2} \end{aligned} \quad (12)$$

根据 $d = \frac{1}{z^2} - \frac{2\cos \alpha}{zu_z} + \frac{1}{u_z^2}$ 的曲线图, 距离 d 的

最大值一定位于定义域(即 z 的范围)的边界上。 α 表示 q_s 和 p_s 之间的弧长, 那么定义域一定包含在 $[u_z - \alpha, u_z + \alpha]$ 内。将 z 的边界值代入式(12)就可以得到距离 d 的最大值, 记为 $r(\alpha)$ 。也就是说, 圆心位于点 p , 半径为 $r(\alpha)$ 的圆包含所有的对应图像球面点 q_s 的图像平面点 q , 其中 q_s 满足 $\angle(p_s, q_s) \leq$

α 。由式(10), $\alpha = \varepsilon_{\min} + 3\sigma^2/2$, 那么图像平面中圆的半径为

$$r(\alpha) = r(\varepsilon_{\min} + 3\sigma^2/2) \quad (13)$$

令 $\mathbf{p}_i = [u_i \ v_i \ 1]^T$ 是第 i 个图像平面点, 则其对应的 3 维点的球面投影与式(10)中的相同, 即 $\bar{\mathbf{R}}(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})$, 该球面投影的平面坐标可以表示为

$$\begin{bmatrix} \hat{u}_i \\ \hat{v}_i \end{bmatrix} = \begin{bmatrix} \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \\ \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \end{bmatrix} \quad (14)$$

式中, $\bar{\mathbf{R}}_j, j=1,2,3$ 代表 $\bar{\mathbf{R}}$ 的第 j 行。重投影误差可以写为

$$\left\| \begin{bmatrix} u_i - \hat{u}_i \\ v_i - \hat{v}_i \end{bmatrix} \right\| = \left\| \begin{bmatrix} u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \\ v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \end{bmatrix} \right\| \quad (15)$$

考虑式(13), 得到二阶锥约束

$$\left\| \begin{bmatrix} u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \\ v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \end{bmatrix} \right\| \leq r(\varepsilon_{\min} + 3\sigma^2/2) \quad (16)$$

式(16)中的二阶锥约束是对式(10)中二阶锥约束的松弛。换句话说, 如果存在一个 $(\mathbf{C}, \delta \mathbf{r})$ 满足式(10)中的约束, 它必满足式(16)中的约束。更明确地说, 如果问题(10)是可行的, 那么下面的问题一定可行。

给定 $\mathbf{X}_i, \bar{\mathbf{R}}$

寻找 $\mathbf{C}, \delta \mathbf{r} \in [-\sigma, \sigma]^3$

满足

$$\left\| \begin{bmatrix} u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \\ v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \end{bmatrix} \right\| \leq r(\varepsilon_{\min} + 3\sigma^2/2) \quad (17)$$

式(17)中的约束是凸约束, 可以利用二阶锥规划找到 $(\mathbf{C}, \delta \mathbf{r})$ 。下面考虑用 L_∞ 范数替换式(17)中的 L_2 范数, 于是有

$$\left\| \begin{bmatrix} u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \\ v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \end{bmatrix} \right\|_\infty \leq r(\varepsilon_{\min} + 3\sigma^2/2) \quad (18)$$

等价于

$$\left| u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \right| \leq r(\varepsilon_{\min} + 3\sigma^2/2)$$

且

$$\left| v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \right| \leq r(\varepsilon_{\min} + 3\sigma^2/2), i=1, \dots, n \quad (19)$$

式(19)给出 $2n$ 个线性不等式代表 n 个以摄像机为中心的方锥, 也就是说当用 L_∞ 范数代替 L_2 范数时, 式(17)中的 n 个圆锥被 n 个方锥替代。因为每一个方锥外切于对应的圆锥, 所以当问题式(17)可行时, 式(20)问题必定可行。

给定 $\mathbf{X}_i, \bar{\mathbf{R}}$

寻找 $\mathbf{C}, \delta \mathbf{r} \in [-\sigma, \sigma]^3$

满足 $\left| u_i - \frac{\bar{\mathbf{R}}_1(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \right| \leq r(\varepsilon_{\min} + 3\sigma^2/2)$

且 $\left| v_i - \frac{\bar{\mathbf{R}}_2(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})}{\bar{\mathbf{R}}_3(\mathbf{X}_i + \delta \mathbf{r} \times \mathbf{X}_i - \mathbf{C})} \right| \leq r(\varepsilon_{\min} + 3\sigma^2/2) \quad (20)$

可以得出结论, 如果问题式(10)可行, 那么问题式(17)必定可行, 从而推出问题式(20)可行。根据原命题和其逆否命题等价的原理, 如果问题式(20)不可行, 那么问题式(17)必不可行, 同理问题式(10)也不可行。

本文算法总结如下: 与文献[13]类似, 采用分支定界方法对旋转空间进行搜索, 旋转量用轴角表示为 $\mathbf{r} = \alpha \hat{\mathbf{r}}$, 其中 α 表示旋转量的角, $\hat{\mathbf{r}}$ 是表示旋转轴的单位矢量。这样所有的 3 维旋转量可以由半径为 π 的球体 $\mathbf{B}_\pi^3$ 内的点表示。为了简化算法, 将球松弛成它的外切立方体 $\mathbf{C}_\pi^3 = [-\pi, \pi]^3$, 把立方体 $\mathbf{C}_\pi^3$ 分割成小的立方体, 通常每次分割是将立方体平分成 8 块小立方体。随着立方体的尺寸逐渐变小, 解就会越来越接近最优解。因此, 分支定界算法可以描述为:

1) 将旋转空间(球 $\mathbf{B}_\pi^3$) 松弛成它的外切立方体 $\mathbf{C}_\pi^3 = [-\pi, \pi]^3$ 。将立方体 $\mathbf{C}_\pi^3$ 分割成小的立方体 $\mathbf{D}_k$, 对每一个小立方体, 计算其中心点 $\bar{\mathbf{R}}$, 判断是否有解 $(\mathbf{C}, \delta \mathbf{r})$ 满足问题式(20), 这是个可行性问题。如果问题式(20)不可行, 那么问题式(10)也不可行, 这个小立方体就被舍弃。

2) 否则将这个小立方体继续平分成 8 块。如果分得的小立方体中不包含原旋转空间(球 $\mathbf{B}_\pi^3$) 中

的任何元素,那么这个小立方体也被舍弃。

直至小立方体包含满足要求的解,算法终止。本文方法在分支定界过程中求解一系列线性规划的可行性问题,相对文献[13]中的二阶锥规划问题,速度要快得多。下面通过实验验证本文算法的有效性。

3 实验结果

3.1 模拟实验

本文首先采用计算机模拟数据进行了实验。令摄像机位姿参数 $R = I_{3 \times 3}$ (单位阵), $C = [0 \ 0 \ 0]^T$, 即摄像机坐标系与世界坐标系重合,选择 20 个空间点,使其图像投影均匀分布在一个 500×500 像素的模拟图像上,在图像坐标上添加均值为 0, 方差为 σ 的高斯白噪声, σ 从 0.2 到 2 均匀取 10 个值,即 σ 分别取 0.2, 0.4, ..., 2, 分别使用本文算法和文献[13]的算法求 R 和 C , 并与真值比较。对于 R , 用求得的 R 与真值间的角度 R 表示, 结果如图 2 所示; 对于 C , 用求得的 C 与真值 $[0 \ 0 \ 0]^T$ 差的模 C 来表示, 结果如图 3 所示。

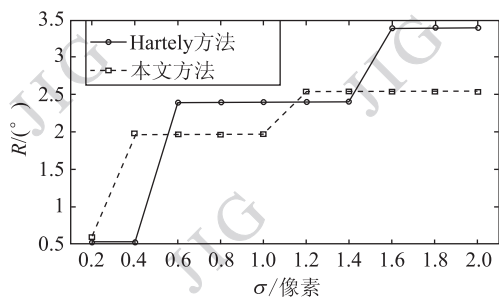


图 2 本文方法和 Hartely 方法求得的 R 与真值结果比较

Fig. 2 The result of our method and Hartely et al's method compared with the true value for R

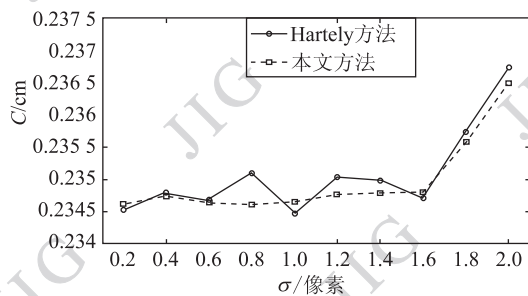


图 3 本文方法和 Hartely 方法求得的 C 与真值比较结果

Fig. 3 The result of our method and Hartely et al's method compared with the true value for C

本文算法运算时间平均为 143 s, 文献[13]算法平均为 351 s, 由此可见, 与文献[13]算法相比, 本文方法在保证精度的同时, 运算速度提高了一倍以上。

3.2 真实实验

采用位姿估计算法常用的恐龙图像序列^[15]进行真实实验。对于这组图像, 其特征点的 2 维信息和 3 维信息都是已知的, 包括 36 幅图像, 328 个图像点, 其中的 3 幅图像如图 4 所示。每一幅视图中有 22 至 154 个特征点可以被观测到, 并且没有外点。采用的计算机配置为 2.0 GHz CPU, 2 GB 内存。



图 4 恐龙数据中的 3 幅视图

Fig. 4 Three images of the dinosaur data set

实验中分别采用本文算法和文献[13]的算法对所有的 36 幅恐龙图像估计了摄像机的位姿参数 R 和 C 。两种方法都是采用分支定界法在旋转空间搜索来求取最优解, 搜索终止的判断依据分别为可行性问题式(20)和式(10)是否有解。试验结果如图 5 至图 7 所示, 图 5 列出了达到同样精度时本文算法和 Hartely 方法^[13]的运算时间, 对 36 幅图的实验结果表明本文方法比 Hartely 的方法速度提高一倍以上。图 6 是使用两种方法计算的 R 与真值比较的结果, 求得的 R 与真值 R' 的差用 $R'R^{-1}$ 的角度 $\theta (0 \leq \theta \leq \pi)$ 来表示, 结果显示两种算法与真值的误差都在 5° 以内。图 7 是使用两种方法计算的 C 与真

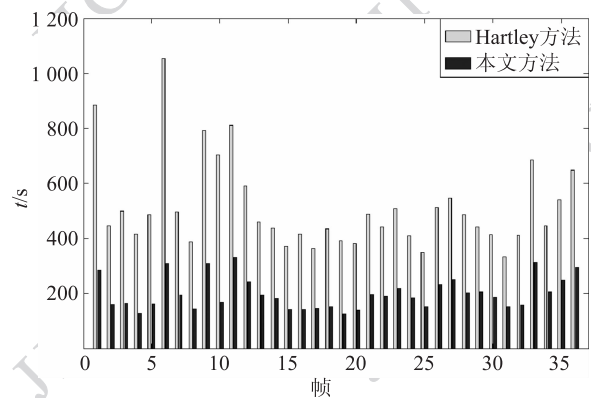


图 5 本文方法和 Hartely 方法的运算时间

Fig. 5 The computation time of our method and Hartely et al's method

值比较的结果,我们使用求得的 C 与真值 C' 的差的模来度量 C 的误差,结果显示两种方法求得的 C 与真值的差都较小。真实实验结果表明,本文算法在保证求解精度的同时,在速度上比文献[13]的算法提升了一倍以上。

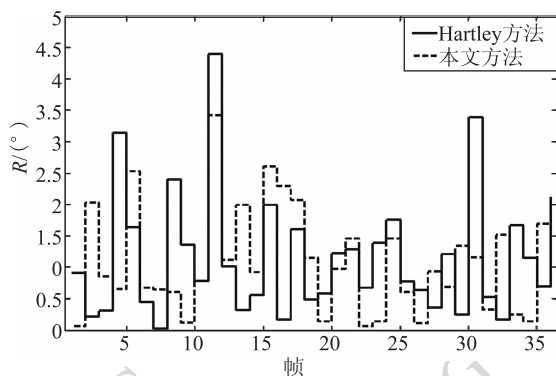


图6 采用两种方法计算出的 R 和真值之间的差
Fig. 6 The difference between the rotations R computed using two methods and the ground truth

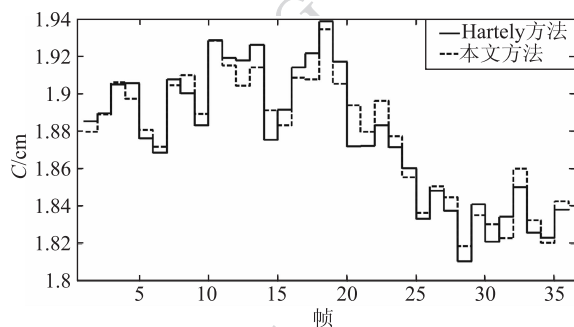


图7 采用两种方法计算出的 C 和真值之间的差
Fig. 7 The difference between the translations C computed using two methods and the ground truth

4 结 论

摄像机位姿估计问题可以转化为一个最小化重投影误差的优化问题。现有的方法^[12-13]将该问题转换成一个序列二阶锥规划问题,通过搜索旋转空间来求取全局最优解。但是一般情况下,二阶锥规划比线性规划速度慢得多,因此用线性规划形式描述摄像机位姿估计问题,算法速度会得到很大提高。基于这种思想,本文通过将二阶锥约束松弛为线性约束,提出一种结合分支定界法和线性规划的新算法,有效地估计了摄像机的最优位姿。最后通过模拟数据和真实数据进行实验,与文献[13]算法进行

了对比,结果表明本文方法在保证全局最优解的同时具有更快的求解速度。

参考文献 (References)

- [1] Hartley R I, Zisserman A. Multiple View Geometry in Computer Vision[M]. 2nd ed. Cambridge: Cambridge University Press, 2004.
- [2] Olson C. A general method for geometric feature matching and model extraction[J]. Journal Computer Vision, 2001, 45(1): 39-54.
- [3] Cass T. Polynomial-time geometric matching for object recognition [J]. Journal Computer Vision, 1999, 21(1/2): 37-61.
- [4] Jacobs D. Matching 3-d models to 2-d images [J]. Journal Computer Vision, 1999, 21(1/2): 123-153.
- [5] David P, Dementhon D, Duraiswami R, et al. Simultaneous pose and correspondence determination[J]. Journal Computer Vision, 2004, 59(3): 259-284.
- [6] Ke Q, Kanade T. Quasiconvex optimization for robust geometric reconstruction[C]//Proceedings of the International Conference on Computer Vision. Beijing: Microsoft Asia Research, 2005: 986-993.
- [7] Kahl F. Multiple view geometry and the L_∞ -norm [C]//Proceedings of the International Conference on Computer Vision. Beijing: Microsoft Asia Research, 2005: 1002-1009.
- [8] Olsson C, Kahl F, Oskarsson M. Optimal estimation of perspective camera pose [C]//Proceedings of the International Conference on Pattern Recognition. Hong Kong: Hong Kong University, 2006: 5-8.
- [9] Olsson C, Kahl F, Oskarsson M. Branch-and-Bound Methods for Euclidean Registration Problems[J]. IEEE Transactions. on Pattern Analysis and Machine Intelligence, 2009, 31(5): 783-794.
- [10] Enqvist O, Kahl F. Robust optimal pose estimation [C]//Proceedings of the European Conference on Computer Vision. Marseille, France: Palais des Congrès, 2008: 141-153.
- [11] Enqvist O, Josephson K, Kahl F. Optimal Correspondences from Pairwise Constraints [C]//Proceedings of the International Conference on Computer Vision. Kyoto, Japan: Kyoto University, 2009: 1295-1302.
- [12] Hartley R, Kahl F. Global optimization through searching rotation space and optimal estimation of the essential matrix [C]//Proceedings of the International Conference on Computer Vision. Brazil: Rio de Janeiro, 2007: 1-8.
- [13] Hartley R, Kahl F. Global optimization through rotation space search[J]. Journal Computer Vision, 2009, 82(1): 64-79.
- [14] Boyd S, Vandenberghe L. Convex optimization [M]. Cambridge: Cambridge University Press, 2004.
- [15] Visual Geometry Group. Multi-view and Oxford Colleges building reconstruction [DB/OL]. [2011-05-10] <http://www.robots.ox.ac.uk/~vgg/data/>.