



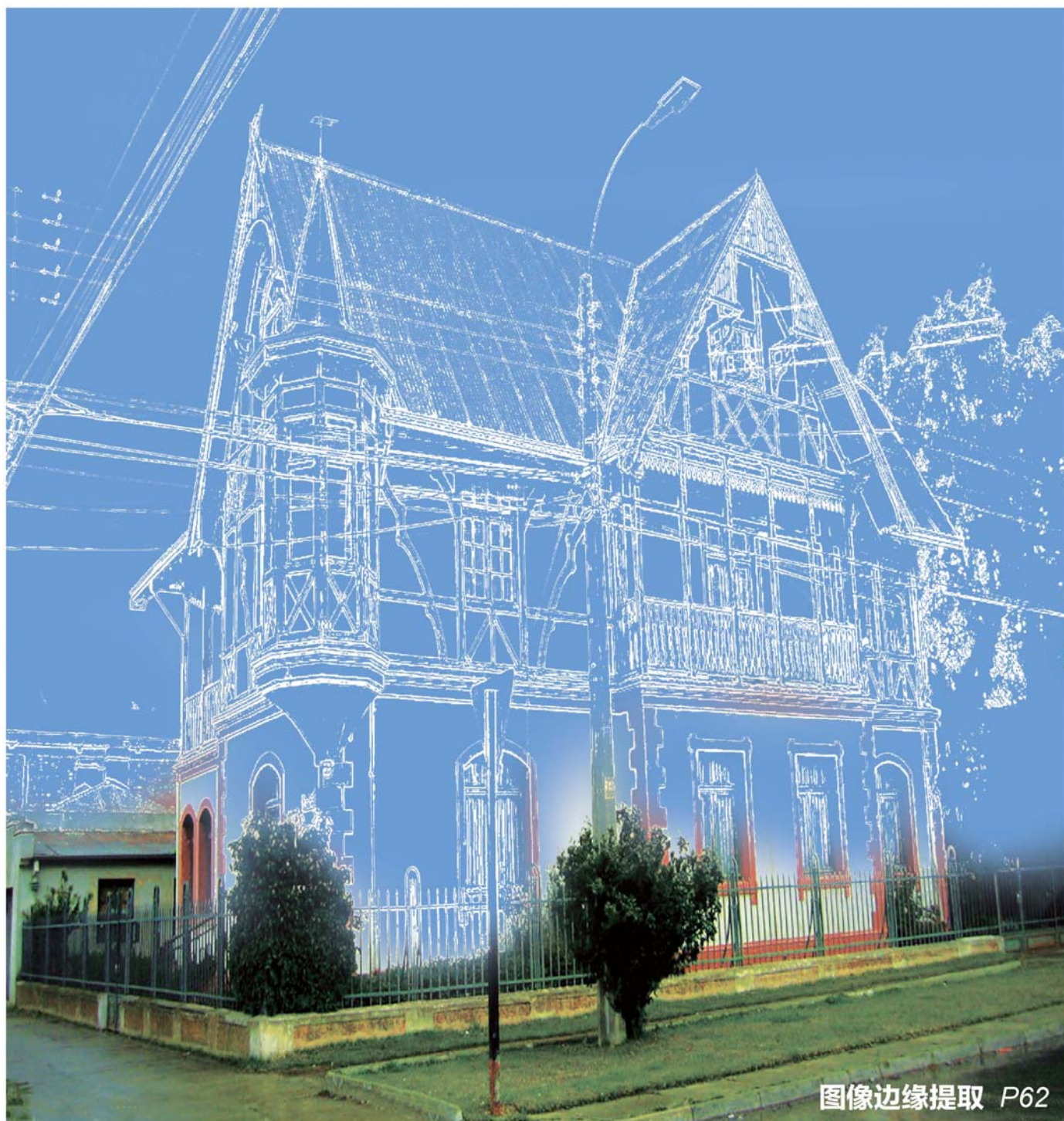
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2014
01
VOL.19

ISSN1006-8961
CN11-3758/TB



图像边缘提取 P62



第19卷第1期（总第213期）

2014年1月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

2013 all rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@irsa.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司
(邮政信箱：北京399信箱 邮编：100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@irsa.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

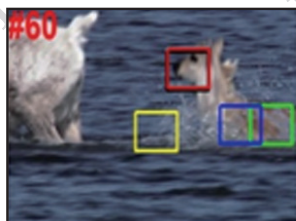
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

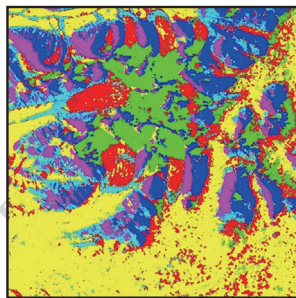
国内定价 50.00元



基于 L_2 范数最小化的实时目标跟踪(第36页)



数字地球渲染中单精度浮点数误差改正(第119页)



全极化合成孔径雷达影像地形纠正及其在雪冰制图中的应用(第150页)

图像处理和编码

图像阈值化的自适应粗糙熵方法

吴涛 0001

利用局部极值点特性的多聚焦图像快速融合

杨达, 王孝通, 徐冠雷, 吴俊波 0011

摄像机位姿的高精度快速求解

李书杰, 刘晓平 0020

整数变换实现大容量图像无损信息隐藏

邱应强 0028

图像分析和识别

基于 L_2 范数最小化的实时目标跟踪

齐美彬, 杨勋, 杨艳芳, 陆磊, 蒋建国 0036

利用七宫格的遮挡车辆凹性检测与分割

刘万军, 陈虹宇, 姚雪, 曲海成 0045

基于体素的人手结构自动估测

姚争为, 潘志庚 0054

结合高斯加权距离图的图像边缘提取

贾迪, 孟祥福, 孟琰, 董娜, 方金凤 0062

基于标定的任意半径柱面上2维条形码畸变校正

王伟, 何卫平, 郭改放, 牛晋波, 曹西征 0069

面向惯性约束聚变实验靶图像的快速椭圆检测

吴倩, 邹伟, 徐德, 张峰 0076

融合MBP和EPMOD的人脸识别

闫海婷, 王玲, 李昆明, 刘机福 0085

遮挡目标的分片跟踪处理

张彦超, 许宏丽 0092

图像理解和计算机视觉

多尺度分析的运动注意力计算

刘龙, 樊波阳 0101

多尺度空间判别性概率潜在语义分析的场景分类

季海峰, 高隽, 郑鹏, 王婧 0109

计算机图形学

数字地球渲染中单精度浮点数误差改正

李尚林, 郑利平, 张静, 杨柳青 0119

医学图像处理

小波与双边滤波的医学超声图像去噪

张聚, 王陈, 程芸 0126

Contourlet变换系数加权的医学图像融合

张鑫, 陈伟斌 0133

遥感图像处理

滑坡高分辨率遥感多维解译方法及其应用

鲁学军, 史振春, 尚伟涛, 周和颐 0141

全极化合成孔径雷达影像地形纠正及其在雪冰制图中的应用

符思涛, 李震, 田帮森 0150

结合邻域相关影像与最大相关性最小冗余性特征选择的面向对象变化检测

邹利东, 潘耀忠, 朱文泉, 周公器, 李宜展 0158

CONTENTS

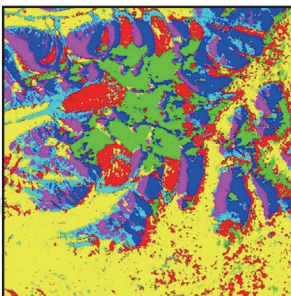
JOURNAL OF IMAGE AND GRAPHICS



Real-time object tracking based on L_2 -norm minimization(P36)



Correction method of single-precision floating-point error for digital Earth rendering(P119)



Topographic correction of polarimetric synthetic aperture radar and its application in snow and ice mapping(P150)

Image Processing and Coding

- Adaptive rough entropy method for image thresholding 0001
Wu Tao
- Multi-focus image fusion based on properties of local extrema 0011
Yang Da, Wang Xiaotong, Xu Guanlei, Wu Junbo
- An accurate and fast algorithm for camera pose estimation 0020
Li Shujie, Liu Xiaoping
- Large capacity lossless information hiding in images using integer transform 0028
Qiu Yingqiang

Image Analysis and Recognition

- Real-time object tracking based on L_2 -norm minimization 0036
Qi Meibin, Yang Xun, Yang Yanfang, Lu Lei, Jiang Jianguo
- Detection and segmentation of occluded vehicles based on seven grids 0045
Liu Wanjun, Chen Hongyu, Yao Xue, Qu Haicheng
- Automatic calibration of hand structures based on voxel data 0054
Yao Zhengwei, Pan Zhigeng
- Image edge extraction combined with a Gaussian weighted distance graph 0062
Jia Di, Meng Xiangfu, Meng Lu, Dong Na, Fang Jinfeng
- Restoration of distorted 2D barcode marked on arbitrary radius cylinder based on calibrated camera 0069
Wang Wei, He Weiping, Guo Gaifang, Niu Jinbo, Cao Xizheng
- Fast ellipse detection for target images in ICF experiment 0076
Wu Qian, Zou Wei, Xu De, Zhang Feng
- Face recognition by fusing MBP and EPMOD 0085
Yan Haiting, Wang Ling, Li Kunming, Liu Jifu
- Fragments tracking under occluded target 0092
Zhang Yanchao, Xu Hongli

Image Understanding and Computer Vision

- Multi-scale analysis based motion attention computation 0101
Liu Long, Fan Boyang
- Scene classification based on multi-scale spatial discriminative probabilistic latent semantic analysis 0109
Ji Haifeng, Gao Jun, Zheng Peng, Wang Jing

Computer Graphics

- Correction method of single-precision floating-point error for digital Earth rendering 0119
Li Shanglin, Zheng Liping, Zhang Jing, Yang Liuqing

Medical Image Processing

- Despeckling for medical ultrasound images based on wavelet and bilateral filter 0126
Zhang Ju, Wang Chen, Cheng Yun
- Medical image fusion based on weighted Contourlet transformation coefficients 0133
Zhang Xin, Chen Weibin

Remote Sensing Image Processing

- The method and application of multi-dimension interpretation for landslides using high resolution remote sensing image 0141
Lu Xuejun, Shi Zhenchun, Shang Weitao, Zhou Heyi
- Topographic correction of polarimetric synthetic aperture radar and its application in snow and ice mapping 0150
Fu Sitao, Li Zhen, Tian Bangsen
- The object-oriented change detection based on neighborhood correlation images and the minimum-redundancy-maximum-relevance feature selection 0158
Zou Lidong, Pan Yaozhong, Zhu Wenquan, Zhou Gongqi, Li Yizhan

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2014)01-0020-08

论文引用格式: 李书杰, 刘晓平. 摄像机位姿的高精度快速求解[J]. 中国图象图形学报, 2014, 19(1): 20-27. [DOI: 10.11834/jig.20140103]

摄像机位姿的高精度快速求解

李书杰, 刘晓平

合肥工业大学计算机与信息学院, 合肥 230009

摘要: 目的 针对对应点个数大于等于6的摄像机位姿估计问题, 提出一种既适用于已标定也适用于未标定摄像机的时间复杂度为 $O(n)$ 的高精度快速算法。方法 首先选取4个非共面虚拟控制点, 并根据空间点和虚拟控制点的空间关系以及空间点的图像建立线性方程组, 以此求解虚拟控制点的图像坐标及摄像机内参, 再由 POSIT (pose from orthography and scaling with iterations) 算法根据虚拟控制点及其图像坐标求解旋转矩阵和平移向量。结果 模拟数据实验和真实图像实验表明该算法时间复杂度和计算精度均优于现有的已标定摄像机位姿的高精度快速求解算法 EPnP (efficient perspective-n-point)。结论 该算法能够同时估计摄像机内外参数, 而且比现有算法具有更好的速度和精度。

关键词: 已标定摄像机; 未标定摄像机; PnP 问题; EPnP 算法; POSIT 算法

An accurate and fast algorithm for camera pose estimation

Li Shujie, Liu Xiaoping

School of Computer and Information, Hefei University of Technology, Hefei 230009, China

Abstract: **Objective** Aiming at the camera pose estimation problem, with the corresponding points number greater than six, we propose an accurate and fast algorithm which has a time complexity of $O(n)$ and is applicable to both: calibrated cameras and uncalibrated cameras. **Method** First, four non-coplanar virtual control points are selected and linear equations according to the relationship between space points and virtual control points and the image of space points are being setup. Then the image coordinates of the virtual control points and the camera internal parameters are solved by these equations. Second, the rotation matrix and the translation vector are solved by the POSIT (pose from orthography and scaling with iterations) algorithm according to virtual control points and their image points. **Result** Experimental results with simulative data and real images indicate that the time complexity and calculation precision of the proposed algorithm are better than EPnP (efficient perspective-n-point) algorithm which is the existing accurate and fast pose estimation algorithm for calibrated camera. **Conclusion** The proposed algorithm can estimate both camera internal and external parameters, and time complexity and calculation precision are better than existing methods.

Key words: calibrated camera; uncalibrated camera; PnP problem; EPnP algorithm; POSIT algorithm

收稿日期: 2013-03-15; 修回日期: 2013-06-27

基金项目: 安徽省教育厅自然科学研究重大项目 (KJ2013ZD02); 高等学校博士学科点专项科研基金项目 (20120111110003); 中央高校基本科研业务费专项资金 (2013HGBH0001, 2013HGBZ0170)

第一作者简介: 李书杰 (1983—), 女, 讲师, 2012 年于中国科学技术大学获模式识别与智能系统专业博士学位, 主要研究方向为基于视频和图像的快速建模。E-mail: lisjhut@hfut.edu.cn

0 引言

摄像机参数估计分为内参数估计和外参数估计,内参数估计被称为摄像机标定(camera calibration)^[1-2],而外参数估计被称为摄像机位姿估计(pose estimation)^[3-8]。摄像机位姿估计问题通常是指在摄像机已标定,即内参数已知的情况下,用 n 个空间点与它们对应的图像点估计摄像机的外参数,即旋转矩阵 $\mathbf{R}$ 与平移向量 $\mathbf{t}$,也称为PnP(perspective-n-point)问题。PnP问题可以分为两类,一类是 $3 \leq n \leq 5$ 的情况,另一类是 $n \geq 6$ 的情况^[8]。第一类PnP问题的研究焦点为确定问题的实解最多可达多少个,结论包括:P3P问题最多有4个解^[3];当4个控制点共面时,P4P问题有唯一解^[4-5],而当4个控制点不共面时,P4P问题最多有4个解^[6];P5P问题最多可以有两个解^[7]。但是由于此类PnP问题用较少的点估计 $\mathbf{R}$ 和 $\mathbf{t}$,通常存在多解而且对图像噪声极为敏感,因此第二类PnP问题也一直受到研究者的关注^[9-14],研究焦点为求解的精度、鲁棒性及运算复杂度等。DLT(direct linear transformation)算法一直是解决 $n \geq 6$ 的PnP问题的经典方法,但是由于其未考虑旋转矩阵的正交约束,使得位姿估计精度较低,时间复杂度为 $O(n^3)$ 。Quan和Lan通过构造四次方程组的方法求解摄像机位姿,精度较高,但时间复杂度为 $O(n^5)$ ^[9]。Fiore提出一种时间复杂度为 $O(n^2)$ 的方法^[10],但是鲁棒性较低。Li和Xu提出一种复杂度为 $O(n)$ 的鲁棒算法^[11],但是存在多个解。2009年Lepetit和Moreno提出的高精度快速位姿估计算法EPnP(efficient perspective-n-point)^[12],其核心思想是将空间点都通过4个虚拟控制点来表示,再通过空间点和图像点之间的对应信息求解虚拟控制点的摄像机坐标,继而求得所有空间点的摄像机坐标,最后用Horn绝对定位算法求得旋转矩阵和平移向量。EPnP算法被认为是目前最高效的摄像机位姿估计算法之一,并有研究者在此基础上做进一步工作。Yang和Wu在EPnP基础上,提出加权EPnP算法,使得在时间复杂度略有增加的基础上,位姿估计的精度提高^[13];Chen借鉴EPnP算法的思想,提出一种基于虚拟控制点的迭代算法,其位姿估计的精度和鲁棒性较以往迭代算法均有所提高^[14]。

但是上述位姿估计算法中除DLT算法外均需以摄像机内参已知为前提,然而在一些情况下摄像

机内参数需要经常调节变化(如主动视觉系统中),因此对未知内参数,即未标定摄像机的位姿估计问题同样具有重要意义。由于旋转矩阵和平移向量共有6个未知参数,所以未标定摄像机的位姿估计至少需要4个点。关于未标定摄像机的P4P和P5P问题已有广泛研究,如文献[15-18]讨论了如何用4个控制点求解摄像机焦距及位姿;文献[5,19]用4个控制点同时求解摄像机焦距、像素比例因子和位姿;文献[18,20]用5个控制点求解摄像机焦距、像素比例因子、主点和位姿。但此类方法因为用较少的点估计内外参数,同样面临对噪声敏感的问题,因此有必要用较多对应点估计未标定摄像机的位姿,但是 $n \geq 6$ 的未标定摄像机PnP问题相关文献较少。

考虑到摄像机的畸变参数和主点相对稳定,未标定摄像机的位姿估计可以集中在未知焦距摄像机的位姿估计问题上,针对 $n \geq 6$ 的PnP问题,在EPnP算法和POSIT(pose from orthography and scaling with iterations)算法^[21]的基础上提出一种既适用于已标定摄像机也适用于未知焦距摄像机的高精度快速算法:首先选取4个非共面虚拟控制点,根据空间点和虚拟控制点的空间关系以及空间点的图像坐标建立线性方程组,获得最小化代数误差意义下的虚拟控制点图像坐标,若为未知焦距摄像机则再由虚拟控制点空间距离计算摄像机焦距;再将焦距作为已知量由POSIT算法根据虚拟控制点及其图像坐标求解旋转矩阵和平移向量。由于采用少量虚拟控制点及其最小化代数误差意义下的图像坐标计算摄像机内外参数,使得本文提出的算法在计算精度和时间复杂度上均优于EPnP算法。

1 问题描述

给定 n 个空间点在世界坐标系及图像坐标系下的对应点对 $\{\mathbf{P}_i^w, \mathbf{u}_i \mid i=1, 2, \dots, n\}$,式中 $\mathbf{P}_i^w = (x_i, y_i, z_i, 1)^T$ 是空间点坐标,右上标 w 用来标注在世界坐标系下的坐标,记 $\tilde{\mathbf{P}}_i^w = (x_i, y_i, z_i)^T$ 是该空间点的非齐次世界坐标, $\mathbf{u}_i = (u_i, v_i, 1)^T$ 表示 $\mathbf{P}_i^w$ 对应的图像点齐次坐标,记 $\tilde{\mathbf{u}}_i = (u_i, v_i)^T$ 是该图像点的非齐次坐标。空间点 $\mathbf{P}_i$ 和图像点 $\mathbf{u}_i$ 的投影成像关系为

$$\lambda_i \mathbf{u}_i = \mathbf{K}[\mathbf{R} \quad \mathbf{t}] \mathbf{P}_i \quad (1)$$

式中, λ_i 是空间点 $\mathbf{P}_i$ 的深度值。摄像机具有等效焦距,其内参数矩阵 $\mathbf{K}$ 形如

$$K = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

式中,焦距 f 为已知或未知。本文所讨论的问题是如何根据空间点和图像点之间的对应关系求解旋转矩阵 K 和平移向量 t ,以及在摄像机焦距未知的情况下求解 K 。

2 算法描述

2.1 求解虚拟控制点图像坐标

在世界坐标系中设定4个非共面虚拟控制点,在世界坐标系下的齐次坐标分别记为 $\{C_i^w | i=1,2,3,4\}$,如图1所示。这4个控制点通常选为

$$C_1^w = [0,0,0,1]^T$$

$$C_2^w = [1,0,0,1]^T$$

$$C_3^w = [0,1,0,1]^T$$

$$C_4^w = [0,0,1,1]^T$$

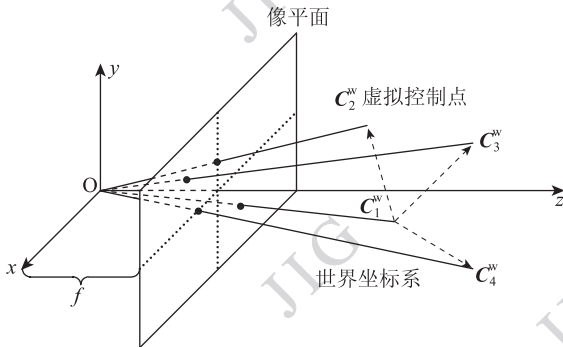


图1 虚拟控制点及其图像点

Fig. 1 Virtual control points and their image points

为与世界坐标系下的坐标加以区别,用右上标 c 表示摄像机坐标系下的坐标。令 $\{C_j^c | j=1,2,3,4\}$ 为虚拟控制点在摄像机坐标系下的齐次坐标, $\{\tilde{C}_j^c | j=1,2,3,4\}$ 为相应的非齐次坐标。 $\{c_j | j=1,2,3,4\}$ 为虚拟控制点对应图像点的齐次坐标, $\{\tilde{c}_j | j=1,2,3,4\}$ 为相应的非齐次坐标。 $\{P_i^w | i=1,2,\dots,n\}$ 为空间点在摄像机坐标系下的齐次坐标, $\{\tilde{P}_i^c | i=1,2,\dots,n\}$ 为相应的非齐次坐标。空间点的世界坐标和控制点的世界坐标存在关系

$$P_i^w = \sum_{j=1}^4 \alpha_{ij} C_j^w, i=1,2,\dots,n \quad (3)$$

式中,向量 $[\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4}]^T$ 为空间点 P_i^w 在以控制

点 $\{C_i^c | i=1,2,3,4\}$ 为基的欧氏空间的坐标,即

$$(\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4})^T = [C_1^w, C_2^w, C_3^w, C_4^w]^{-1} P_i^w \quad (4)$$

由线性关系在欧氏变换下的不变性得

$$P_i^c = \sum_{j=1}^4 \alpha_{ij} C_j^c, i=1,2,\dots,n \quad (5)$$

所以有

$$\lambda_i u_i = K \tilde{P}_i^c = K \sum_{j=1}^4 \alpha_{ij} \tilde{C}_j^c \quad (6)$$

$$i=1,2,\dots,n$$

设 $\tilde{C}_j^c = [x_j^c, y_j^c, z_j^c]^T$,则

$$\lambda_i = \sum_{j=1}^4 \alpha_{ij} z_j^c \quad (7)$$

把式(7)代入式(6),可得到两个独立的方程,即

$$\sum_{j=1}^4 \alpha_{ij} f x_j^c - \alpha_{ij} u_i z_j^c = 0 \quad (8)$$

$$\sum_{j=1}^4 \alpha_{ij} f y_j^c - \alpha_{ij} v_i z_j^c = 0 \quad (9)$$

令 $Z = [Z_1^T, Z_2^T, Z_3^T, Z_4^T]^T$, $Z_j^c = [f x_j^c, f y_j^c, z_j^c]^T$, $j=1,2,3,4$,则由 n 个空间点和图像点之间的对应可以得到含 $2n$ 个方程的线性方程组,记为

$$MZ = 0 \quad (10)$$

则 Z 的解是矩阵 M 的核空间,即

$$Z = \sum_{i=1}^N \beta_i W_i \quad (11)$$

式中, W_i 是对应于 $M^T M$ 零特征值的特征向量, N 是 $M^T M$ 核空间的维数, β_i 是待定系数。对于透视投影模型, N 的值为1^[12-13],所以有

$$Z = \beta W \quad (12)$$

式中, $W = [w_1^T, w_2^T, w_3^T, w_4^T]^T$, $w_j = [w_{j1}, w_{j2}, w_{j3}]^T$,则可以得到4个虚拟控制点的图像坐标,即

$$c_j = \left\{ \frac{w_{j1}}{w_{j3}}, \frac{w_{j2}}{w_{j3}}, 1 \right\}, j=1,2,3,4 \quad (13)$$

2.2 求解焦距 f

若焦距 f 未知,则还可以根据式(12)及欧氏变换保距性计算 f 。由式(12)可得

$$\begin{pmatrix} f x_j^c \\ f y_j^c \\ z_j^c \end{pmatrix} = \beta \begin{pmatrix} w_{j1} \\ w_{j2} \\ w_{j3} \end{pmatrix}, j=1,2,3,4 \quad (14)$$

即

$$\begin{pmatrix} x_j^c \\ y_j^c \\ z_j^c \end{pmatrix} = \begin{pmatrix} \beta \\ f \\ \beta w_{j3} \end{pmatrix}, j=1,2,3,4 \quad (15)$$

由欧氏变换的保距性, 可知

$$\|\tilde{\mathbf{C}}_i^c - \tilde{\mathbf{C}}_j^c\| = \|\tilde{\mathbf{C}}_i^w - \tilde{\mathbf{C}}_j^w\| \quad (16)$$

$$i, j = 1, 2, 3, 4$$

将式(15)代入式(16)可得到以 $\frac{\beta^2}{f^2}$ 和 β^2 为未知量的 6 个线性方程, 即

$$((w_{i1} - w_{j1})^2 + (w_{i2} - w_{j2})^2) \frac{\beta^2}{f^2} +$$

$$(w_{i3} - w_{j3})^2 \beta^2 = \|\tilde{\mathbf{C}}_i^w - \tilde{\mathbf{C}}_j^w\|^2 \quad (17)$$

$$i, j = 1, 2, 3, 4$$

由上述 6 个方程构成的线性方程组仅含两个未知数, 因此可以用最小二乘法求解。令 $x = \frac{\beta^2}{f^2}, y = \beta^2$,

则用最小二乘法求解式(17)可得 x 和 y , 所以有

$$f = \sqrt{y/x} \quad (18)$$

2.3 求解旋转矩阵及平移向量

POSIT 算法是针对已标定摄像机的精度较高的位姿求解算法, 通常仅需 45 步迭代即可收敛至精确解, 而且每步迭代仅需求解线性代数方程, 时间复杂度较低^[21]。因此, 由较少的虚拟控制点用 POSIT 算法可以快速求解摄像机位姿。

用 2.1 节中求解得到的 4 个虚拟控制点图像坐标, 以及已知的焦距 f 或由 2.2 节中求解得到的焦距 f , 带入 POSIT 算法中求解旋转矩阵 $\mathbf{R}$ 及平移向量 $\mathbf{t}$ 。

2.4 算法的两点讨论

1) 算法的精度

由式(8)和式(9)构成的方程组求出的最小二乘解相当于求解如下最小二乘问题, 即

$$\mathbf{Z} = \arg \min \sum_{i=1}^n \left\| \sum_{j=1}^4 \alpha_{ij} (z_j^c \mathbf{u}_i - \tilde{\mathbf{f}}_j^c) \right\|^2 \quad (19)$$

所以, 由式(8)和式(9)构成的方程组可以获得最小化代数误差意义下的虚拟控制点图像坐标。

获得最小化代数误差意义下的虚拟控制点图像坐标后, 再由精度较高的 POSIT 算法迭代计算得到旋转矩阵和平移向量, 是本文算法获得较高精度的关键所在。

2) 算法的时间复杂度

本文算法共有 3 步: 1) 求解式(8)和式(9)构成的方程组获得虚拟控制点图像坐标; 2) 根据式(12)及欧氏变换保距性求解焦距 f ; 3) 由 POSIT 算法根据虚拟控制点及其图像坐标求解旋转矩阵 $\mathbf{R}$ 与平移向量 $\mathbf{t}$ 。最耗时的部分为步骤 1) 中求解对应于

$\mathbf{M}^T \mathbf{M}$ 零特征值的特征向量, 时间复杂度为 $O(n)$ ^[9], n 为对应点的个数。步骤 2) 用固定点数求解焦距 f , 耗费时间固定。由于仅用固定的 4 个虚拟控制点求解, 而且 POSIT 算法仅需少量迭代次数即可获得较精确的摄像机位姿。步骤 3) 耗费的时间较少, 且不随对应点个数的增长而增长。综上, 本文算法的时间复杂度为 $O(n)$ 。

另外, 与 EPnP 算法相比, 本文算法不需根据全部点对应由 Horn 绝对定位算法求解旋转矩阵和平移向量, 仅采用 4 个虚拟控制点求解, 这正是本文算法计算时间少于 EPnP 算法的关键所在。

3 实验及分析

3.1 模拟数据实验

通过计算机仿真实验测试本文算法性能。设置等效焦距 $f_u = f_v = 800$, 主点 $[u_0, v_0] = [0, 0]$ 。空间点按照均匀分布产生于 $[-2, 2] \times [-2, 2] \times [4, 9]$ 的立方体区域内。旋转矩阵的 3 个分别沿 X 、 Y 、 Z 轴的坐标均按均匀分布产生于 $0^\circ \sim 45^\circ$ 之间。算法求得的 $\hat{\mathbf{K}}$ 、 $\hat{\mathbf{R}}$ 及 $\hat{\mathbf{t}}$ 与真实 $\mathbf{K}$ 、 $\mathbf{R}$ 及 $\mathbf{t}$ 的误差分别表示为

$$E_K = \|\hat{\mathbf{K}} - \mathbf{K}\|_F / \|\mathbf{K}\|_F \quad (20)$$

$$E_R = \|\hat{\mathbf{R}} - \mathbf{R}\|_F / \|\mathbf{R}\|_F \quad (21)$$

$$E_t = \|\hat{\mathbf{t}} - \mathbf{t}\|_F / \|\mathbf{t}\|_F \quad (22)$$

为数据对比说明方便起见, 称本文中用于已标定摄像机的位姿估计算法为 AFPnP (accurate and fast algorithm for PnP problem) 算法, 称用于未标定摄像机的位姿估计算法为 AFUFnP 算法 (accurate and fast algorithm for unknown focal length PnP problem)。

考虑到 EPnP 算法目前被认为是最好的高精度快速位姿估计算法以及 DLT 算法是解决对应点个数 $n \geq 6$ 未标定 PnP 问题的经典算法, 对比 DLT 算法、EPnP 算法以及提出的 AFPnP、AFUFnP 算法进行实验。

1) 对应点数目实验

实验目的是比较在不同对应点数目下各种算法的计算时间及计算精度。图像点添加均值为 0, 标准差为 2 的高斯噪声。对应点数目的变化范围为 $[20, 120]$, 变化步长为 10。对每个对应点数目进行 10 000 次独立实验, 再对计算时间和误差求取平均值。图 2 为实验结果。

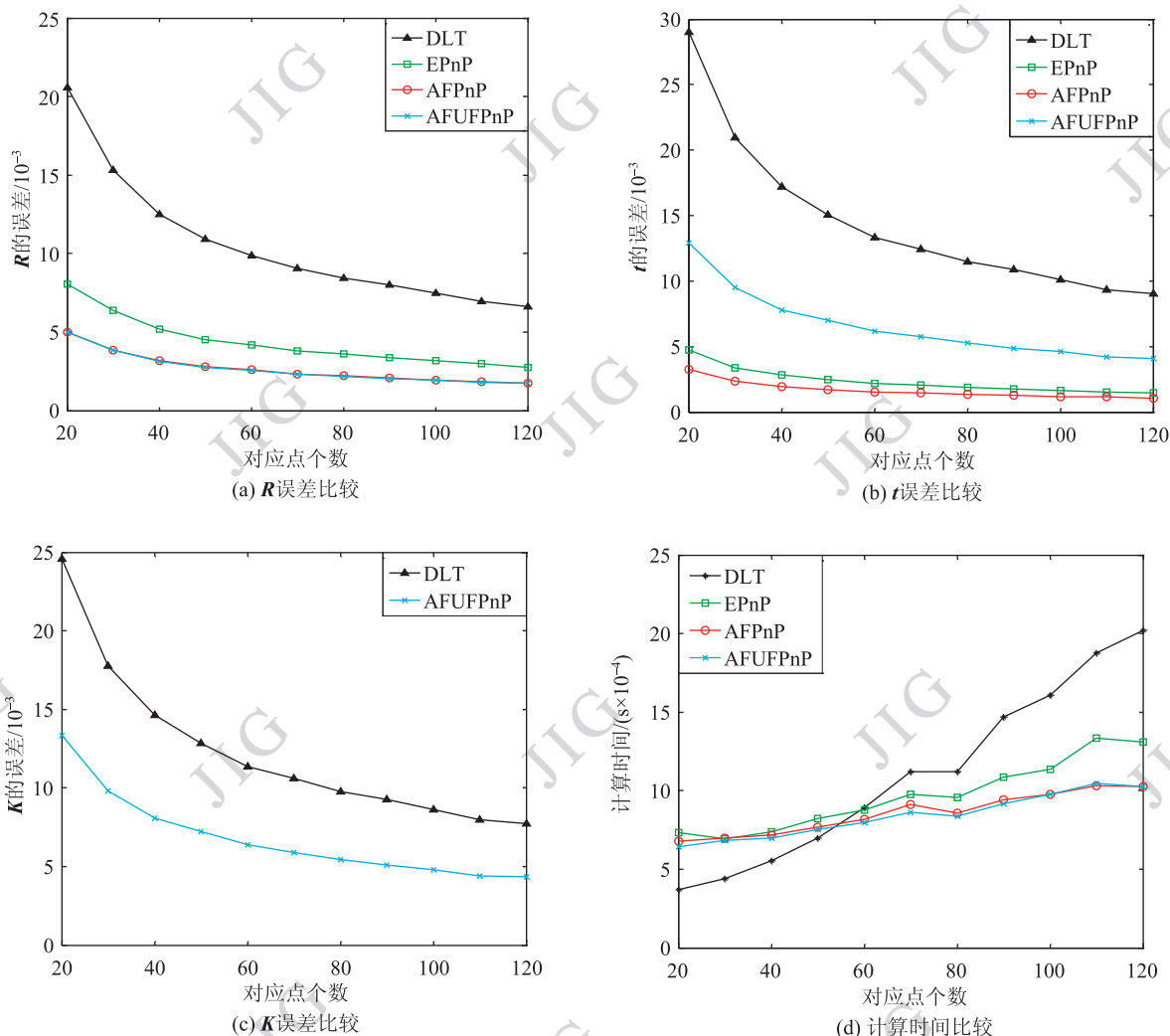


图2 对应点数目实验结果

Fig. 2 Number of corresponding points experiment results

实验结果说明:

(1) 对于各对应点数目, AFPnP 算法获得的 $\hat{R}$ 及 $\hat{t}$ 的精度均优于 DLT 算法和 EPnP 算法。

(2) 对于各对应点数目, AFUFnP 算法获得的 $\hat{R}$ 的精度优于 EPnP 算法和 DLT 算法。由于 AFUFnP 算法是在摄像机焦距未知的情况下求解摄像机位姿, AFUFnP 算法获得的 $\hat{t}$ 的精度略低于已标定摄像机的 EPnP 算法, 但仍高于 DLT 算法。对于各对应点数目, AFUFnP 算法获得的 $\hat{K}$ 的精度高于 DLT 算法。

(3) 对于各对应点数目, AFPnP 和 AFUFnP 算法的计算时间均少于 EPnP 算法。而且 AFPnP 和 AFUFnP 算法的计算时间随对应点数增加的速度

缓于 EPnP 算法。

2) 图像噪声实验

实验目的是为了比较各种算法在不同噪声水平下的计算精度。实验中设置对应点数目为 100, 图像点添加均值为 0, 标准差为 σ 的高斯噪声, 其中 σ 在 $[0, 5]$ 内变化, 变化步长为 0.5。对每个噪声做 10 000 次独立实验, 再对各种算法的计算误差求平均值。图 3 是实验结果。

实验结果说明:

(1) 当噪声为 0 时, AFPnP 算法和 AFUFnP 算法与 EPnP 算法和 DLT 算法一样可以获得精确值, 即误差为零。

(2) 在各噪声水平下, AFPnP 算法获得的 $\hat{R}$ 及 $\hat{t}$ 的精度均优于 DLT 算法和 EPnP 算法。

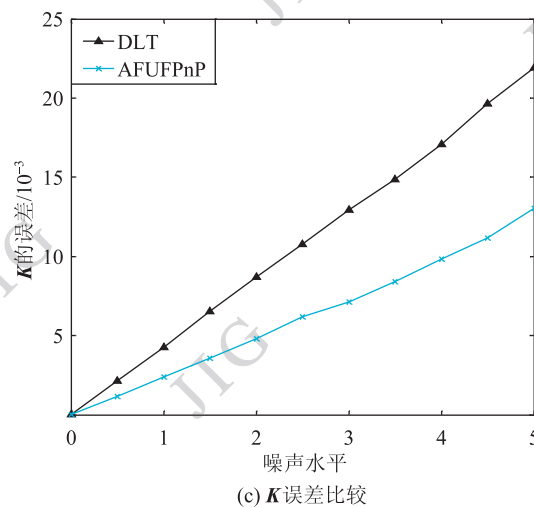
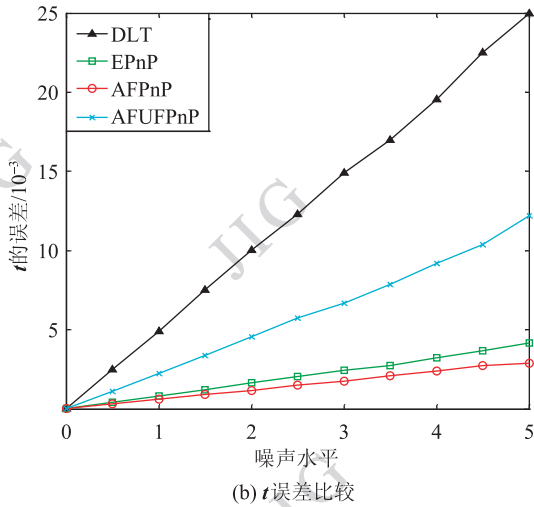
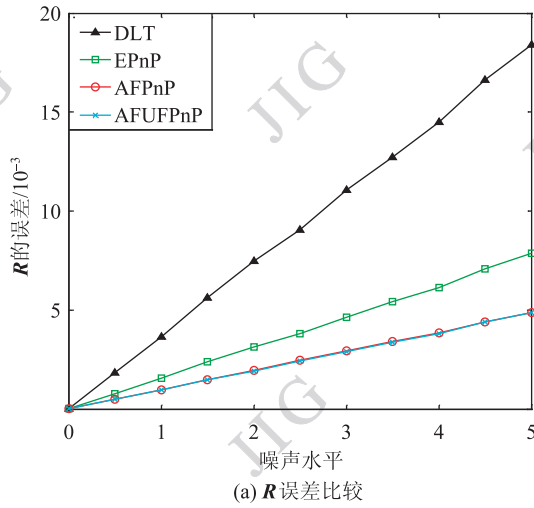


图3 图像噪声实验结果

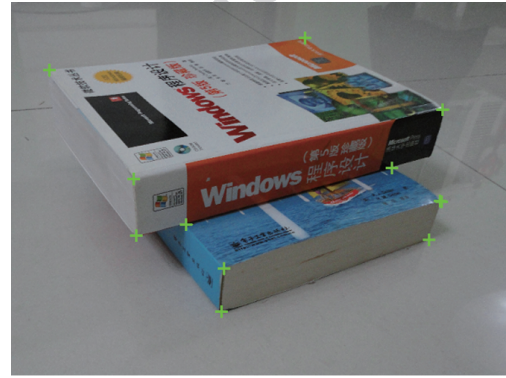
Fig. 3 Results of image noise experiment

(3) 在各噪声水平下, AFUFnP 算法获得的 $\hat{R}$ 的精度优于 EPnP 算法和 DLT 算法。AFUFnP 算

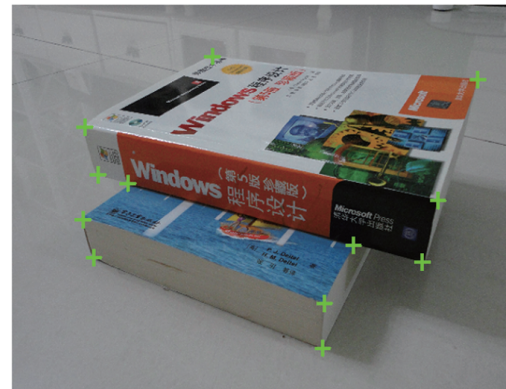
法获得的 $\hat{t}$ 的精度略低于已标定摄像机的 EPnP 算法, 但仍高于 DLT 算法。在各噪声水平下, AFUFnP 算法获得的 $\hat{K}$ 的精度均高于 DLT 算法。

3.2 真实图像实验

在真实图像实验中, 用 SONY CyberShot W3500 拍摄两幅图像, 如图 4 所示。由于特征点识别非本文涉及内容, 所有图像特征点均由手工标定。相机主点和畸变参数相对稳定, 在实验中预先由 Heikkilä 算法^[22] 预先标定并作为已知值。经过后期处理, 每幅图像获得 11 个点对应, 如图 4 所示。对图像点添加均值为 0, 标准差为 σ 的高斯噪声, 其中 σ 在 $[0.5, 3.5]$ 内变化, 变化步长为 0.5。每个噪声水平进行 10 000 次独立实验。在真实图像实验中, 由于无法获得真实的摄像机外参数, 用重投影误差的均值来间接对比 EPnP 算法、AFPnP 算法和 AFUFnP 算法的重投影精度。表 1 为实验结果。



(a)



(b)

图4 真实图像

Fig. 4 Real images

实验结果表明提出的 AFPnP 算法和 AFUFnP 算法的重投影误差均小于 EPnP 算法。

表 1 真实图像实验结果
Table 1 Results of real image experiment

噪声水平	图 4(a) 重投影误差			图 4(b) 重投影误差		
	EPnP	AFPnP	AFUFnP	EPnP	AFPnP	AFUFnP
0.5	3.618 0	2.471 6	2.451 3	4.665 0	2.877 7	2.912 0
1	4.545 0	3.104 5	3.083 4	6.036 0	3.598 4	3.627 4
1.5	5.509 9	3.724 1	3.698 5	6.973 7	4.262 3	4.320 5
2	6.455 1	4.318 9	4.297 9	8.159 1	5.001 3	5.100 6
2.5	7.378 8	4.993 9	4.990 9	9.272 0	5.715 5	5.851 1
3	8.228 0	5.612 6	5.591 6	10.420 1	6.400 6	6.546 2
3.5	9.085 7	6.193 7	6.194 2	11.590 5	7.242 7	7.442 7

4 结 论

EPnP 算法被认为是目前最好的用于已标定摄像机的高精度快速位姿估计方法之一。在 EPnP 算法的基础上,结合 POSIT 算法,提出一种新的高精度快速位姿估计算法,并可扩充应用于未知焦距摄像机的位姿估计问题,而且计算时间和计算精度均优于 EPnP 算法。由于摄像机的畸变因子和主点坐标变化均不大,这使得本文算法还可以扩展应用到未标定摄像机的位姿估计问题。本文算法的核心思想在于用最小化代数误差意义下的虚拟控制点图像坐标进行位姿估计以提高算法精度,用少量虚拟控制点及其图像点计算位姿以提高算法速度。文中理论推导和实验数据表明:

1) 由于采用少量虚拟控制点及其最小化代数误差意义下的图像坐标计算摄像机位姿,这使得在摄像机内参已知的情况下,本文算法的计算精度和时间复杂度均优于 EPnP 算法。

2) 在摄像机焦距未知的情况下,本文算法对旋转矩阵的计算精度,重投影精度仍优于已标定摄像机位姿估计算法 EPnP 算法。

3) 在摄像机内参形如式(2)且未知的情况下,本文算法对摄像机内参的计算精度优于常用于解决对应点个数 $n \geq 6$ 未标定摄像机 PnP 问题的 DLT 算法。

本文算法的源程序可以在网上 (<http://lisj.weebly.com>) 下载。

参考文献 (References)

[1] Zhang, Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11):1330-1334.

[2] Meng X, Hu Z. A new easy camera calibration technique based on circular points[J]. Pattern Recognition, 2003, 36(5):1155-1164.

[3] Haralick B M, Lee C N, Ottenberg K, et al. Review and analysis of solutions of the three point perspective pose estimation problem [J]. International Journal of Computer Vision, 1994, 13(3):331-356.

[4] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6):381-395.

[5] Hu Z Y, Lei C. A short note on P4P problem[J]. Acta Automatica Sinica, 2001, 27(6):770-776.

[6] Hu Z Y, Wu F C. A note on the number of solutions of the non-coplanar P4P problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4):550-555.

[7] Wu F C, Hu Z Y. A study on the P5P problem[J]. Journal of Software, 2001, 12(5):768-775.

[8] Wu Y, Hu Z. PnP problem revisited [J]. Journal of Mathematical Imaging and Vision, 2006, 24(1):131-141.

[9] Quan L, Lan Z. Linear n-point camera pose determination [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(8):774-780.

[10] Fiore P D. Efficient linear solution of exterior orientation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2):140-148.

[11] Li S Q, Xu C, Xie M. A robust $O(n)$ solution to the perspective-n-point problem [J]. IEEE Trans. on Pattern Anal. Mach. Intell., 2012, 34(7):1444-1450.

- [12] Lepetit V, Moreno-Noguer F, Fua P. Epnp: an accurate $O(n)$ solution to the pnp problem [J]. *International Journal of Computer Vision*, 2009, 81(2):155-166.
- [13] Yang S, Wu F C. Weighted linear methods for the camera pose estimation[J]. *Journal of Software*, 2011, 22(10):2476-2487.
- [14] Chen P, Hu G D. A simple algorithm for camera pose estimation [C]//*Proceedings of International Conference on Mechatronics and Automation*. Chengdu:IEEE, 2012:2181-2186.
- [15] Abidi M A, Chandra T. A new efficient and direct solution for pose estimation using quadrangular targets: algorithm and evaluation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(5):534-538.
- [16] Bujnak M, Kukeleva Z, Pajdla T. A general solution to the P4P problem for camera with unknown focal length [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, AK:IEEE, 2008:1-8.
- [17] Penne R, Veraart J, Abbeloos W, et al. Four point algorithm for the recovery of the pose of a one-dimensional camera with unknown focal length [J]. *Computer Vision*, 2012, 6(4):314-323.
- [18] Triggs B. Camera pose and calibration from 4 or 5 known 3D points [C]//*Proceedings of the 7th IEEE International Conference on Computer Vision*. Kerkyra, Greece: IEEE, 1999, 278-284.
- [19] Guo Y. A novel solution to the P4P problem for an uncalibrated camera [J]. *Journal of Mathematical Imaging and Vision*, 2012, 45(2): 1-13.
- [20] Wu F C, Hu Z Y. A note on the P5P problem with an uncalibrated camera [J]. *Chinese Journal of Computer*, 2001, 24(11): 1221-1226.
- [21] Dementhon D F, Davis L S. Model-based object pose in 25 lines of code [J]. *International Journal of Computer Vision*, 1995, 15(1):123-141.
- [22] Heikkilä J. Geometric camera calibration using circular control points [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2000, 22(10):1066-1077.