

一种非定标图像高精度三维重建算法^{*}

袁建英^{1,2}, 刘甲甲², 王 凯², 李柏林²

(1. 昆明学院 自动控制与机械工程学院, 昆明 650214; 2. 西南交通大学 机械工程学院, 成都 610031)

摘要: 由非定标图像重建三维场景有着广泛的应用。给出了一种非定标多视图图像三维重建算法。该算法主要基于因子分解和光束法平差技术。首先用因子分解方法得到射影空间下相机投影矩阵和物点坐标, 以旋转矩阵的正交性以及绝对二次曲面秩为3为约束, 将射影空间升级到欧式空间, 最后用光束法平差进行优化。该方法可同时获得相机的内外参数、畸变系数和场景的三维坐标。仿真实验表明, 在1000 mm × 1000 mm × 400 mm的范围内, 当像点检测误差在0-1 pixel和0-2 pixel内, 所重建三维点的误差分别为0.1530 mm和0.6712 mm。在500 mm × 500 mm × 200 mm下, 真实实验重构三维点的误差在0.3 mm以内。所提出的算法稳定可靠, 可对实际工程进行指导。

关键词: 机器视觉; 非标定图像; 三维重建; 因子分解

中图分类号: TP391.41; TP301.6

文献标志码: A

文章编号: 1001-3695(2012)12-4786-03

doi:10.3969/j.issn.1001-3695.2012.12.101

3D reconstruction algorithm from uncalibrated images with high precision

YUAN Jian-ying^{1,2}, LIU Jia-jia², WANG Kai², LI Bai-lin²

(1. School of Automatic Control & Mechanical Engineering, Kunming University, Kunming 650214, China; 2. School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: 3D reconstruction from uncalibrated images has many applications. This paper proposed a 3D reconstruction method from uncalibrated images. The method was based on factorization and bundle adjustment algorithm. First, it got camera projection matrix and object coordinates in projective space using factorization. Then with orthogonality of rotation matrix and rank 3 of the absolute quadric as the constraint, it upgraded the projective space to euclidean space. Finally, it applied bundle adjustment to refine all the parameters. The inner, exterior and distortion parameters could be computed simultaneously. Simulation experiments show that the mean distance error are about 0.1530 mm and 0.6712 mm separately when the reconstruction field is about 1000 mm × 1000 mm × 400 mm and the image detecting error are in 0-1 pixels and 0-2 pixels. Real experiment shows that the reconstruction error is below 0.3 mm at the field of 500 mm × 500 mm × 200 mm. The proposed algorithm is robust and can be guide for real engineering application.

Key words: machine vision; uncalibrated images; 3D reconstruction; factorization

0 引言

由多幅图像恢复场景的三维结构是计算机视觉领域的研究重点和难点^[1,2], 该技术被广泛应用于三维识别^[3]、交通事故现场勘查^[4]、机器人导航^[5,6]等领域。20世纪90年代以前, 由多幅图像重构三维场景时, 都需要借助于精密的靶标来求解相机与场景之间的投影矩阵(包含相机的内外参数)^[7], 这大大地限制了三维重建技术的应用场合。进入90年代后, Faugeras^[8]和Hartley^[9]提出了摄像机自标定的概念, 使得在场景未知和摄像机任意运动的一般情形下标定成为可能。文献[8]从射影几何的角度出发, 证明了每两幅图像间存在着两个形如Kruppa方程的二次非线性约束, 通过直接求解Kruppa方程组可以解出内参数。鉴于直接求解Kruppa方程的困难, 人

们又提出了分层逐步标定的思想^[10,11], 即首先对图像序列作射影重建, 在此基础上再进行仿射重建和欧氏重建。射影重构的方法主要有基于两视图之间对极几何的方法和基于图像矩阵SVD分解的方法。前一种方法由于仅利用了两幅图像之间的信息, 当图像序列增加时, 便会产生累积误差, 致使标定精度不高, 进而影响重建效果。因子分解方法利用了多幅图像点之间的投影约束关系, 采用了鲁棒的SVD分解, 可同时获得相机的投影矩阵和场景的三维点坐标。鉴于此, 本文采用因子分解方法获得射影空间下相机投影矩阵和场景物点坐标, 再将射影空间升级到欧式空间。在求解两个空间变换矩阵时, 利用旋转矩阵的正交性及绝对二次曲面秩为3的特性, 提高求解的稳定性。为了提高重建精度, 考虑相机的畸变效应, 并利用光束法平差技术对相机的内外参数、畸变参数、物点坐标进行求

收稿日期: 2012-04-19; **修回日期:** 2012-05-25 **基金项目:** 昆明学院校级课题基金资助项目(XJ11G004); 云南省应用基础研究资助项目(2010CD094); 西南交通大学牵引动力国家重点实验室自主课题资助项目(2012TPL_T10)

作者简介: 袁建英(1982-), 女, 四川眉山人, 博士研究生, 主要研究方向为由图像进行三维场景重构(yuanjianying1-1@sina.com); 刘甲甲(1983-), 男, 安徽淮北市人, 博士研究生, 主要研究方向为图像处理、视觉检测; 王凯(1982-), 四川成都人, 博士研究生, 主要研究方向为图像分割、机器视觉; 李柏林(1962-), 男, 贵州遵义人, 教授, 博导, 主要研究方向为图形图像学、优化技术、视觉检测。

精。仿真实验和真实实验表明,本文所给的方法有较好的鲁棒性,并能得到较高重建精度。

1 本文算法介绍

本文所给的重建方法属于分层重建范畴,首先采用因子分解方法得到相机投影矩阵和物点在射影空间下的值,然后将射影空间升级至欧式空间,再求解相机的畸变系数,最后通过光束法平差技术对所有参数进行求精。

1.1 相机模型

相机采用针孔成像模型,空间中某点 X_i 在相机上的成像可用式(1)表示为

$$\lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = P_i X_i \quad (1)$$

其中:

$$P_i = K_i (R_i | T_i) \quad (2)$$

$$X_i = [X_{wi} \ Y_{wi} \ Z_{wi} \ 1]^T$$

在式(1)(2)中, λ_i 表示投影深度; $K_i = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 表示

相机的内参数矩阵, f_u, f_v 表示相机的有效焦距; s 表示相机 CCD 阵列垂直度因子; (u_0, v_0) 表示相机的主点坐标; R_i 表示第 i 个摄像机坐标系与世界坐标系之间的旋转矩阵; T_i 表示第 i 个摄像机坐标系相对于世界坐标系的平移矩阵; X_i 为物点在世界坐标系下的齐次坐标。

相机的畸变参数主要有径向畸变、切向畸变以及像平面畸变。在实际应用中,不同的相机具有不同的畸变模型,根据笔者的经验,确定所使用相机的畸变模型需要通过不同的实验组合来最终确定。

1.2 相机参数和场景三维结构初始值的获得

由因子分解实现射影的重构参见文献[12]。其基本原理如下:假定空间中有 p 个物点在 F 幅图像上成像,将 $p \times F$ 个点按如下的规则组成测量矩阵 $W_{3F \times p}$ 。

$$\underbrace{\begin{bmatrix} \lambda_{11} m_{11} & \cdots & \lambda_{1p} m_{1p} \\ \vdots & & \vdots \\ \lambda_{F1} m_{F1} & \cdots & \lambda_{Fp} m_{Fp} \end{bmatrix}}_W = \underbrace{\begin{bmatrix} P_1 \\ \vdots \\ P_F \end{bmatrix}}_Q \underbrace{\begin{bmatrix} X_1 & \cdots & X_p \end{bmatrix}}_X \quad (3)$$

在式(3)中,对 W 进行 SVD 分解,可得到 Q 和 X 。SVD (W) = VDU^T , 由于 W 的秩为 4, 故 $Q = V_0, X = D_0 U_0^T, V_0$ 表示 V 的前四列, D_0 表示 D 的前四行。在对 W 进行 SVD 分解前,需要计算投影深度 λ_{ij} , 对于投影深度的估计文献[13,14]都有阐述,本文采用文献[13]的方法。到此,便得到了射影空间下摄像机投影矩阵和场景的形状矩阵。

在欧式空间下,第 i 个相机的投影矩阵 P_{ei} 可以表示为 $P_{ei} = K_i (R_i | T_i)$, 定义非奇异的 4×4 矩阵 H 为从射影空间到欧式空间的变换矩阵,即 $P_i H = \alpha_i P_{ei}$, 由于旋转矩阵 R_i 具有单位正交性,可得到如下等式:

$$\begin{cases} P_1 H_0 H_0^T P_1^T = \alpha_1^2 K_1 K_1^T \\ \vdots \\ P_m H_0 H_0^T P_m^T = \alpha_m^2 K_m K_m^T \end{cases} \quad (4)$$

其中: H_0 表示 H 矩阵的前三列,假定拍摄图像为手持相机拍摄场景,则内参数可视 $K_1 = \cdots = K_i = K_0$ 。令 $Tran = H_0 H_0^T$, 是 4×4 的对称矩阵,含有 10 个未知数,每增加一幅图像同时增加一个未知数,因此至少需要三幅图像即可求解出 $Tran$ 的值。 $Tran$ 表示的是对偶绝对二次曲面,其秩为 3。为了在此约束下实现 H_0 的求解,采用 SVD 分解的方法,强制将 $Tran$ 的第四个奇异值设置为 0,从而获取满足约束的 H_0 矩阵。具体推导如式(5)所示。其中, $\text{diag}(\sigma_1, \sigma_2, \sigma_3, 0)$ 表示 $Tran$ 奇异值组成的对角矩阵, $\sigma_1, \sigma_2, \sigma_3$ 是 $Tran$ 的三个奇异值。

$$\begin{aligned} \text{SVD}(Tran) &= A \text{diag}(\sigma_1, \sigma_2, \sigma_3, 0) A^T = \\ &= \text{Adiag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 0) \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 0)^T A^T \\ &\quad \Downarrow \\ H_0 &= A_0 \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}) \end{aligned} \quad (5)$$

其中, A_0 为 A 的前三列。为了获得转换矩阵 H 的第四列 H_4 , 可以通过设置世界坐标系的方法实现。例如从 Q 中任选一点 x_0 作为参考点,设置欧式空间坐标系的原点与该点重合,则可以得到 $H^{-1} x_0 = (0 \ 0 \ 0 \ 1)^T$, 则可以得到 $H_4 = x_0$ 。至此,确定了欧式转换矩阵 H 的所有参数。

$$\begin{cases} P_e = PH \\ X_e = H^{-1} X \end{cases} \quad (6)$$

1.3 相机参数和场景物点坐标的优化

在实际应用中,图像的采集过程总是含有一定的噪声,这便造成了如上文方法得到的相机参数和物点坐标含有一定的误差。为了提高重建精度,需要对参数进行优化。优化目标函数为每个物点的反投影残差最小,如式(7)所示。

$$\min \sum_{i=1}^n \sum_{j=1}^m d(x_{ij}(\text{Angle}_{ei}, t_{ei}, K, Q_{ej}, D), q_{ij})^2 \quad (7)$$

其中: x_{ij} 表示欧式空间下的第 j 个点在第 i 个相机上的反投影坐标; Angle_{ei} 表示第 i 个相机对应姿态的欧拉角; t_{ei} 表示第 i 个相机对应平移向量; K 表示相机的内参数; Q_{ej} 表示欧式空间下的第 j 个三维点; D 表示图像的畸变因子; q_{ij} 表示从图像上检测到的第 j 个点在第 i 个相机上的投影坐标。在本文中,采用光束法平差对所有参数进行优化^[15]

到此,实现了场景的欧式重构与优化。所得到的三维点尺寸和真实场景尺寸存在一个比例因子,可以通过在空间布置标准杆求取该比例因子,得到与实际场景吻合的尺寸。

2 实验及结论

为了验证本文算法是否有效,设计了仿真实验和真实实验。

2.1 仿真实验

首先在空间布设仿真实相机和仿真场景,利用式(1)得到仿真像点,给像点添加随机噪声,得到含有噪声的图像点。用含有噪声的图像点进行场景三维重建。将重建物点的三维坐标与仿真物点三维坐标对齐,以场景物点重建三维坐标和仿真(真实)三维坐标的距离差的平均值作为衡量本文算法的依据。

为了模拟手持相机进行三维场景重建,仿真了 10 个内参数一致的相机、30 个随机物点。相机的内参数如表 1 所示,相机与场景物点的关系如图 1 所示。

表 1 仿真相机内参数

f_u	f_v	s	$u_0(\text{pixel})$	$v_0(\text{pixel})$
4 137.8	4 137.8	0	2 128	1 416

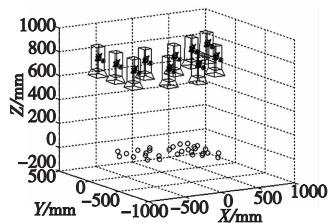


图 1 仿真相机与物点位置关系示意图

鉴于目前对图像特征点检测的定位精度一般均可达到 1pixel 内,本文在对图像添加噪声时,噪声的大小限于 1pixel 和 2pixel 内,这样仿真结论与实际更相符合。由于添加的是随机噪声,一次实验结果具有偶然性,因此每种级别的噪声实验进行 10 次。实验结果如图 2 所示。由图 2 可见,所给出的算法具有较好的稳定性,其中添加 0-1pixel 的 10 次实验平均误差为 0.153mm,其中添加 0-2pixel 的 10 次实验平均误差为 0.671mm,具有较高的标定精度。

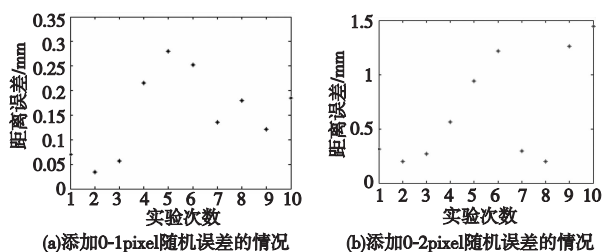


图 2 添加不同噪声实验 10 次的重建误差

2.2 真实实验

所采用的相机分辨率为 1280×1024 pixel,镜头焦距为 12 mm,围绕空间物(约 $500 \text{ mm} \times 500 \text{ mm} \times 200 \text{ mm}$)拍摄 10 组图像,每组图像 6 幅,然后重建物点坐标。图 3 为真实实验中的一张图像。以某两点的重建距离与真实距离之差作为重建误差(真实距离由双目系统测量获得),多个重建误差的平均值作为最终的重建误差。图 4 为真实实验重建距离误差。由图 4 可见,重建最大误差为 0.40 mm,最小误差为 0.075 mm,平均误差为 0.237 mm,这与仿真实验结果吻合。

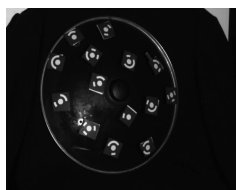


图 3 真实实验中的一张图像

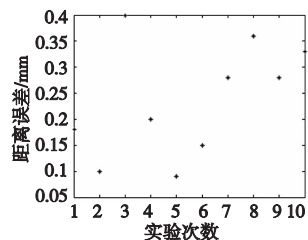


图 4 真实实验重建距离误差

3 结束语

由非标定图像重建场景精确三维信息是机器视觉的研究热点。本文在现有研究基础上,给出了一种分层重建算法。采用因子分解方法得到射影空间下相机参数及物点参数。由于

该算法利用了所有图像信息,提供了更多的约束,从原理上比仅用两幅图像的基于基本矩阵的相机外参计算方法更稳定。在将射影空间参数升级至欧式空间时,考虑绝对二次曲面秩为 3 的性质,提高求解结果的精确性。在欧式空间中获得稳定初始值的基础上,用光束法平差提高相机内外参数、物点坐标的精度。仿真实验和真实实验表明了所给方法具有较好的精度,可对实际视觉测量系统进行指导。

参考文献:

- [1] HAN Bing, PAULSON C, WU Da-peng. 3D dense reconstruction from 2D video sequence via 3D geometric segmentation [J]. *Journal of Visual Communication and Image Representation*, 2011, 22(5): 421-431.
- [2] CHEN Sheng-yong, LI You-fu, KWOK N M. Active vision in robotic systems: a survey of recent developments [J]. *International Journal of Robotics Research*, 2011, 11(30): 1343-1377.
- [3] 孙曦. 基于双目立体视觉的人脸三维重建与识别 [D]. 合肥: 中国科技大学, 2011.
- [4] 鲁光泉. 基于普通相机的交通事故现场三维重建关键技术研究 [D]. 长春: 吉林大学, 2004.
- [5] 李海峰, 傅侃, 周文晖. 视觉导航机器人三维场景重建研究 [J]. *杭州电子科技大学学报*, 2010, 30(2): 49-53.
- [6] 姚丹霖. 地面自主机器人视觉导航三维重建算法研究 [D]. 合肥: 中国科学技术大学, 2006.
- [7] ZHANG Zheng-you. A flexible new technique for camera calibration [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.
- [8] FAUGERAS O. Stratification of 3D vision: projective, affine and metric representation [J]. *Journal Optical Society of America*, 1995, 12(3): 465-484.
- [9] HARTLEY R. Euclidean reconstruction and invariants from multiple images [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1994, 16(10): 1036-1041.
- [10] TRIGGS B. Autocalibration and the absolute quadric [C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1997: 609-614.
- [11] POLLEFEYS M, VAN G L, OOSTERLINCK A. The modulus constraint: a new constraint for self-calibration [C]//Proc of the 13th International Conference on Pattern Recognition. 1996: 349-353.
- [12] HAN Mei, KANADE T. Scene reconstruction from multiple uncalibrated views, Technical Report CMU-RI-TR-00-09 [R]. Pittsburgh: Carnegie-Mellon University, 2000: 1-19.
- [13] STURM P, TRIGGS B. A factorization based algorithm for multi-image projective structure and motion [C]//Proc of the 4th European Conference on Computer Vision. 1996: 709-720.
- [14] ZHU Jun-qiao, LIANG Dong, CHENG Zhi-you. Study of a hierarchical reconstruction algorithm based on multiple views [J]. *Pattern Recognition and Artificial Intelligence*, 2003, 16(4): 407-412.
- [15] TRIGGS B, McLAUCHLAN P F, HARTLEY R I, et al. Bundle adjustment: a modern synthesis [C]//Proc of International Workshop on Vision Algorithms. London: Springer-Verlag 1999: 298-372.