

基于因子分解和光束法平差从航空序列图像恢复三维结构*

邱志强, 陆宏伟, 张小虎, 于起峰

(国防科技大学航天与材料工程学院, 湖南长沙 410073)

摘要 :从航空序列图像恢复目标三维结构是航空摄影测量的基本问题。针对现有算法中精度与计算量的矛盾,提出一种基于因子分解与光束法平差的新算法。作为因子分解与光束法平差的桥梁,提出基于“R-T”模式共线方程的概念,详细推导了其线性化公式。仿真实验和真实图像实验表明,新算法精度高、稳定性好,可应用于航空测绘、虚拟现实、军事侦察等工程实践。

关键词 :航空图像 ;序列图像 ;因子分解 ;光束法平差 ;三维结构

中图分类号 :TP391 文献标识码 :A

3D Structure from the Aerial Sequential Image Based on Factorization and Bundle Adjustment

QIU Zhi-qiang, LU Hong-wei, ZHANG Xiao-hu, YU Qi-feng

(College of Aerospace and Material Engineering, National Univ. of Defense Technology, Changsha 410073, China)

Abstract 3D structure from aerial image sequence is acquiring much attention in aerial photogrammetry. A novel method based on factorization and bundle adjustment as well as its principle and implementation is presented. Being the bridge between factorization and bundle adjustment, the notion “R-T” mode based collinearity equation is clarified as well as its linearization equations. Synthetic and real image test show the new method gives an accurate and robust result, and it can be applied to aerial surveying and mapping, virtual reality and military reconnaissance, etc.

Key words :aerial image ; image sequence ; factorization ; bundle adjustment ; 3D structure

近年来,从航空序列图像恢复目标三维结构的研究引起了航空摄影测量等领域学者的关注。基于因子分解从运动恢复结构的算法^[1],利用图像点在图像序列与目标点序列的双线性约束,采用奇异值分解法将摄像机运动与目标结构分离,算法全线性求解,速度快,但由于采用了摄像机仿射投影模型,因而精度不如光束法平差高。摄影测量学中的光束法区域网平差^[2]主要应用于空中三角测量,它能同时处理若干幅图像、若干个目标点,可以认为它也是一种从序列图像恢复三维结构的算法。由于它本质上是线性最小二乘平差法,因此速度快、精度高。经过多年的发展,光束法平差已经非常成熟,在成图、勘测等工程实践中取得了广泛的应用。但是,这种方法需要所有未知数的初始值,而且精度太低的初始值还可能导致算法不收敛。如果将光束法平差作为结果求精算法与因子分解法结合起来,用因子分解法结果作为光束法平差的初始值,则可发挥两种算法的优势,同时克服它们的缺点,这就是本文新算法的思路。

1 摄像机透视投影成像模型

摄像机成像通常采用透视投影模型,其基本关系式为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

* 收稿日期 2004 - 02 - 15
作者简介 邱志强(1972—),男,博士生。

其中 (X, Y, Z) 为目标点世界坐标 (u, v) 为对应的图像坐标, R 与 T 为世界坐标系与摄像机坐标系的旋转与平移量。这里假设摄像机焦距为 1, 主点位于图像中心, 也就是相当于图像坐标取归一化坐标^[3,4]。式(1)还可改写为^[2]:

$$\begin{cases} u = \frac{r_1 X + r_2 Y + r_3 Z + T_x}{r_7 X + r_8 Y + r_9 Z + T_z} \\ v = \frac{r_4 X + r_5 Y + r_6 Z + T_y}{r_7 X + r_8 Y + r_9 Z + T_z} \end{cases} \quad (2)$$

其中 (T_x, T_y, T_z) 为 T 的分量 (r_1, \dots, r_9) 为 R 的元素。这就是后文提到的基于“ $R \cdot T$ ”模式的共线方程。

2 算法原理

假设已经获得 n 个目标点在 m 幅图像中的对应图像点, 我们的目的是由这些图像点确定所有目标点组成的目标的三维结构。新算法可以分为两个步骤 (1)用因子分解法获得目标三维结构作为初始值 (2)用光束法平差对结果进行优化。

2.1 基于因子分解从运动恢复结构

这里的运动指摄像机在曝光时刻的光心位置与光轴姿态的变化量。如果第 j 个目标点在第 i 幅图像上的图像点坐标为 $p_{ij}(u_{ij}, v_{ij}) (i = 1, \dots, m, j = 1, \dots, n)$, 令 $\bar{u}_i = \frac{1}{n} \sum_{j=1}^n u_{ij}$ 和 $\bar{v}_i = \frac{1}{n} \sum_{j=1}^n v_{ij} (i = 1, \dots, m)$ 将图像点坐标 (u_{ij}, v_{ij}) 作变换 $u'_{ij} = u_{ij} - \bar{u}_i, v'_{ij} = v_{ij} - \bar{v}_i$ 然后组成图像测量矩阵 $W^{[1]}$ 。在无噪声条件下, W 的秩为 3, 有噪声时, W 的秩可能大于 3, 但如果按奇异值从大到小排列, 则从第 4 个开始以后的奇异值远小于前 3 个^[3]。对 W 进行奇异值分解:

$$W = LAR \quad (3)$$

L 为左奇异向量矩阵, A 为奇异值对角矩阵, R 为右奇异向量矩阵。因为从第 4 个以后的奇异值远小于前 3 个, 如果 L 取前 3 列, A 取前 3 行前 3 列, R 取前 3 行, 式(3)仍近似成立。也就是说, 如果令 \hat{M}

$$= L \begin{bmatrix} \sqrt{\sigma_1} & 0 & 0 \\ 0 & \sqrt{\sigma_2} & 0 \\ 0 & 0 & \sqrt{\sigma_3} \end{bmatrix} \text{ 和 } \hat{S} = \begin{bmatrix} \sqrt{\sigma_1} & 0 & 0 \\ 0 & \sqrt{\sigma_2} & 0 \\ 0 & 0 & \sqrt{\sigma_3} \end{bmatrix} R \text{ 其中 } \sigma_1 \geq \sigma_2 \geq \sigma_3 \text{ 且 } \sigma_1, \sigma_2, \sigma_3 \text{ 为 } W \text{ 的 3 个较大}$$

奇异值, 则有

$$W = \hat{M}\hat{S} \quad (4)$$

其中 \hat{M} 为 $2m \times 3$ 阶矩阵, \hat{S} 为 $3 \times n$ 阶矩阵。

在很多考虑摄像机成像模型的算法中, 如果以平行透视投影^[3]代替透视投影, 可大大降低算法的复杂程度, 而且引起的误差可以被接受, 特别是当被摄目标满足傍轴条件^[4], 例如像航空图像拍摄时。对于式(4)来说, 如果采用平行透视投影模型, 加上摄像机内参数等约束, 可以从 $2m \times 3$ 阶矩阵 \hat{M} 求出摄像机运动量, 从 $3 \times n$ 阶矩阵 \hat{S} 求出目标三维结构, 具体方法参见文献[3]。至此, 通过因子分解法实现了运动与结构的分离。

2.2 光束法平差

由于用到了平行透视投影来近似透视投影, 所以不可避免地引进了误差。此外, 还有很多误差因素, 如图像点检测误差、摄像机标定误差、镜头像差等, 因此考虑用平差法来精确求解, 本文采用光束法区域网平差。

首先, 对共线方程式(2)求出偏导数, 然后列出误差方程式

$$V = [B \quad C] \begin{bmatrix} X \\ t \end{bmatrix} - L \quad (5)$$

最后, 按最小二乘法组成法方程式^[2], 以因子分解法求得的三维结构结果作为初始值, 进行迭代求解, 直

至改正数小于设定的阈值为止,实践证明一般迭代 5 次就足够了。

3 基于“R-T”模式的共线方程及其线性化

3.1 基于“R-T”模式共线方程概念的提出

将共线方程式(2)改写为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (6)$$

因此可以认为,图像点齐次坐标 $(u, v, 1)$ 是三维点坐标 (X, Y, Z) 经过先旋转 R 、后平移 T 得到的,所以式(2)称为基于“R-T”模式的共线方程。与此不同,摄影测量学文献中通常的共线方程^[1]写成矩阵形式为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} r'_1 & r'_2 & r'_3 \\ r'_4 & r'_5 & r'_6 \\ r'_7 & r'_8 & r'_9 \end{bmatrix} \begin{bmatrix} X - T_x \\ Y - T_y \\ Z - T_z \end{bmatrix} \quad (7)$$

可以认为,坐标 $(u, v, 1)$ 是坐标 (X, Y, Z) 经过先平移 T 、后旋转 R 得到的,可以称之为基于“T-R”模式的共线方程。由于共线方程形式不同,因此在线性化时,不能直接搬用摄影测量学文献中的公式。

3.2 基于“R-T”模式共线方程的线性化

对共线方程线性化,也就是分别求出式(2)中 u, v 对各变量的偏导数。

令 $I = r_1X + r_2Y + r_3Z + T_x, J = r_4X + r_5Y + r_6Z + T_y, K = r_7X + r_8Y + r_9Z + T_z$ 。先求 $\partial u / \partial \omega$:

$$\begin{aligned} \partial u / \partial \omega &= (\partial I / \partial \omega - u \partial K / \partial \omega) / K \\ &= [X \partial r_1 / \partial \omega + Y \partial r_2 / \partial \omega + Z \partial r_3 / \partial \omega - u(X \partial r_7 / \partial \omega + Y \partial r_8 / \partial \omega + Z \partial r_9 / \partial \omega)] / K \\ &= [(\sin \omega \sin \kappa + \cos \omega \sin \varphi \cos \kappa - u \cos \omega \cos \varphi)Y + (\cos \omega \sin \kappa - \sin \omega \sin \varphi \cos \kappa + u \sin \omega \cos \varphi)Z] / K \end{aligned}$$

同理可得 $\partial u / \partial \varphi, \partial u / \partial \kappa, \partial u / \partial X, \partial u / \partial Y, \partial u / \partial Z, \partial u / \partial T_x, \partial u / \partial T_y, \partial u / \partial T_z, \partial v / \partial \omega, \partial v / \partial \varphi, \partial v / \partial \kappa, \partial v / \partial X, \partial v / \partial Y, \partial v / \partial Z, \partial v / \partial T_x, \partial v / \partial T_y, \partial v / \partial T_z$, 这些偏导数用于建立光束法平差中误差方程式(5)。

在式(7)中, R' 与 T' 分别指摄像机光轴与光心在世界坐标系中的位置与姿态,也就是从摄像机坐标系到世界坐标系的转换。而在式(2)中, R 与 T 分别指“摄像机坐标系与世界坐标系之间的旋转矩阵与平移矢量^[3]”,其本质是世界坐标系原点在摄像机坐标系中的位置与世界坐标系 Z 轴在摄像机坐标系中的姿态,指的是从世界坐标系到摄像机坐标系的转换。它们之间的关系为:

$$\begin{cases} R = R'^{-1} \\ T = -(R')^{-1}T' \end{cases} \quad (8)$$

4 算法的实验验证

4.1 仿真实验

如图 1 所示,假设一个立方体的边长为 4,以它的 8 个顶点为目标点,世界坐标系建立在立方体上,用一台仿真摄像机分别在如图 20 个不同位置对它进行成像,每个位置到世界坐标系原点距离保持在 400 左右,摄像机光轴基本向下,并赋以随机的光轴姿态角扰动。不失一般性,采用的摄像机内参数矩阵为 $\text{diag}(0, 0, 1)$,即主点位于图像中心,焦距为 1,相当于采用归一化图像坐标。将目标点按透视投影方式分别投影到 20 幅图上,对所得图像坐标加上方差为 0.1 像素的高斯噪声。对加噪后的图像点用本文提出的算法处理,可恢复立方体的三维结构。表 1 列出了恢复得到的立方体 12 条边长的结果。可以看出结构恢复平均误差为 0.073。

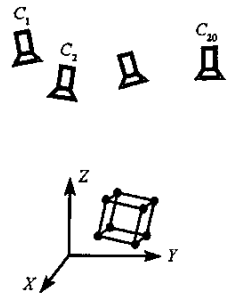


图 1 仿真实验

Fig. 1 The simulated experiment

如果不进行光束法平差,则精度下降很多,表2是只用因子分解法得到的结果,误差来源主要是摄像机成像模型误差。对比表1可知,用平行透视模型来近似透视模型会带来一定误差,而光束法平差可有效地降低误差。

表1 仿真实验结果

Tab.1 Simulated experimental results

1	2	3	4	5	6	7	8	9	10	11	12	平均误差
4.076	4.095	4.008	3.974	4.088	3.097	4.063	3.896	3.964	4.082	4.096	4.109	0.073

表2 仿真实验结果(未进行光束法平差)

Tab.2 Simulated experimental results(without bundle adjustment)

1	2	3	4	5	6	7	8	9	10	11	12	平均误差
4.735	4.739	4.702	4.735	3.197	3.296	4.708	3.396	3.595	4.515	4.608	3.496	0.647

4.2 真实图像实验

图3是一个由12幅图像组成的城堡航空图像序列中的第1、4、9和12幅,图像分辨率为512像素×480像素。整个序列跟踪了26个特征点,基本都是城墙拐角点、城中房顶角点等比较明显的特征,但限于实验条件,图像检测仅取得了整像素级精度。图3为用本文算法对所有目标点进行三维结构恢复后的侧视图,为便于看清楚,图中已将结构恢复后的特征点连接起来。从图中可以看出城墙之间的高差,以及与房顶的高差。

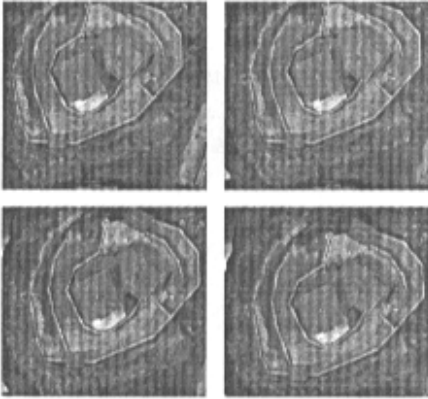


图2 城堡图像序列

Fig.2 The castle image sequence

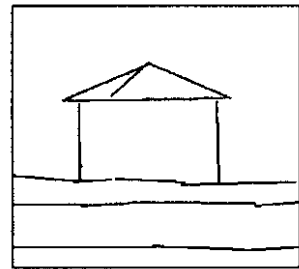


图3 城堡侧视图

Fig.3 Side-view of castle

5 结论

本文提出了一种新的从航空序列图像恢复目标三维结构的算法。运用该算法只要提取并匹配出所有目标点在图像上的位置,就能恢复出目标的三维结构。

本文重点阐述了三个要点:一是巧妙地将因子分解算法和光束法区域网平差算法结合起来处理从序列图像恢复目标三维结构问题,不但充分发挥了两种算法的优点,而且克服了各自的不足。二是提出了基于“ $R-T$ ”模式的共线方程概念,并推导了其线性化公式,从而建立了基于“ $R-T$ ”模式的法方程用于光束法平差,实现了将光束法平差手段引入到计算机视觉体系框架内,以达到两种领域算法的结合。三是阐述了摄像机坐标系与世界坐标系之间的旋转矩阵与平移向量的物理意义,并讨论了它们分别与摄像机光轴姿态、光心位置的关系。本文算法可用于航空摄影测量、三维建模、机器视觉、虚拟现实以及军事侦察、航天摄影测量等领域。

算法还存在需要进一步研究的问题。如果在拍摄序列图像过程中,摄像机的焦距发生了变化,或者部分目标点在部分图像中不可见,或目标点接近共面,这些时候都需要改进算法。另外还要开发相应的

图像点的高精度检测和自动立体匹配的算法。如果有可能,还可以考虑将因子分解算法改为直接基于摄像机透视投影模型,而不是平行透视投影模型,这样肯定会提高第二步光束法平差初始值的精度。

参考文献:

- [1] Poelman, Kanade. A Paraperspective Factorization Method for Shape and Motion Recovery [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(3).
- [2] 王之卓. 摄影测量原理 [M]. 武汉: 测绘科技大学出版社, 2000.
- [3] 马颂德, 张正友. 计算机视觉—理论与算法基础 [M]. 北京: 科学出版社, 1998.
- [4] 于起峰, 等. 基于图像的精密测量与运动测量 [M]. 北京: 科学出版社, 2002.
- [5] 邱志强, 唐力铁, 于起峰. 用神经网络变易有效焦距的摄像机标定法 [J]. 国防科技大学学报, 2002(5).
- [6] Hartley R. Euclidean Reconstruction from Uncalibrated Views, In Applications of Invariance in Computer Vision [M]. Springer-Verlag, 1994, 237-256.
- [7] Mundy J L. The Relationship between Photogrammetry and Computer Vision [J], CAD and CG Sinica, 2002, 14(1).

(上接第26页)

参考文献:

- [1] De Boer J, Bazen A M, Gerez S H. Indexing Fingerprint Databases Based on Multiple Features [A]. ProRISC 2001 Workshop on Circuits, Systems and Signal Processing, Veldhoven, the Netherlands, November 2001.
- [2] Ongun G, Halici U. Fingerprint Classification through Self-organizing Feature Maps Modified to Treat Uncertainties [J]. Proc. of the IEEE, 1996, 84(10): 1497-1512.
- [3] Ballan M, Sakarya F A. A Fingerprint Classification Technique Using Directional Images [J]. Mathematical and Computational Applications, 1999, 4(2): 91-97.
- [4] Keogh E. An Overview of the Science of Fingerprints [J]. Anil Aggrawal's Internet Journal of Forensic Medicine and Toxicology, 2001, 2(1).
- [5] Zhang Qinzhi, Huang Kai, Yan Hong. Fingerprint Classification Based on Extraction and Analysis of Singularities and Pseudoridges [A]. In Proc. Selected Papers from 2001 Pan-Sydney Workshop on Visual Information Processing, Sydney, Australia. Conferences in Research and Practice in Information Technology, 2001.
- [6] Karu K, Jain A K. Fingerprint Classification [J]. Pattern Recognition, 1996, 29(3): 389-404.
- [7] Marcialis G L, Roli F, Frasconi P. Fingerprint Classification by Combination of Flat and Structural Approaches [A]. Proc. of AVBPA 2001, 2001: 241-246.
- [8] Yao Y, Frasconi P, Pontil M. Fingerprint Classification with Combination of Support Vector Machines [A]. Proc. 3rd Int. Conf. on Audio- and Video-Based Biometric Person Authentication, 2001.
- [9] Jain A K, Prabhakar S, Hong L. A Multichannel Approach to Fingerprint Classification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(4): 348-359.
- [10] Yao Y, Marcialis G L, Pontil M, Frasconi P, Roli F. A New Machine Learning Approach to Fingerprint Classification [A]. In 7th Congress of the Italian Association for Artificial Intelligence, 2001: 57-63.
- [11] Yao Y, Marcialis G L, Pontil M, Frasconi P, Roli F. Combining Flat and Structured Representations for Fingerprint Classification with Recursive Neural Networks and Support Vector Machines [J]. Pattern Recognition, 2003, 36(2): 397-406.
- [12] Burges, C J C. A Tutorial on Support Vector Machines for Pattern Recognition [J]. Data Mining and Knowledge Discovery, 1998, 2(2): 955-974.
- [13] Jain A K, Pankanti S, Prabhakar S, Ross A. Recent Advances in Fingerprint Verification [A]. AVBPA 2001, LNCS 2091, 2001: 183-190.
- [14] Jain A, Hong L, Bolle R. On-line Fingerprint Verification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(4): 302-313.
- [15] Luo X P, Tian J, Wu Y. A Minutia Matching Algorithm in Fingerprint Verification [A]. 15th ICPR, Barcelona, 2000: 833-836.
- [16] Hong L, Wan Y, Jain A. Fingerprint Image Enhancement: Algorithm and Performance Evaluation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(8): 777-789.
- [17] Bazen A M, Gerez S H. Directional Field Computation for Fingerprints Based on the Principal Component Analysis of Local Gradients [A]. In Proceedings of ProRISC2000, 11th Annual Workshop on Circuits, Systems and Signal Processing, Veldhoven, The Netherlands, Nov. 2000.
- [18] Rao A R. A Taxonomy for Texture Description and Identification [J]. Springer-Verlag, New York, 1999.
- [19] Mehre B M. Segmentation of Fingerprint Images—A Composite Method [J]. Pattern Recognition, 1999, 22(4): 381-385.

