

文章编号: 0253-2239(2007)06-1004-7

基于仿射近似从序列图像重建目标三维结构*

邱志强^{1,2} 梁永辉¹ 于起峰¹

(¹ 国防科学技术大学光电科学与工程学院光信息科学与技术系, 长沙 410073)
(² 中国人民解放军 76321 部队, 广州 510500)

摘要: 从序列图像重建目标三维结构是摄影测量学和计算机视觉的研究热点。针对传统算法中精度与计算量的矛盾, 提出一种基于仿射近似的新算法, 首先用仿射三线性张量合成序列图像中的缺失图像点, 重新建立图像测量矩阵, 然后用仿射矩阵分解从图像特征直接获得目标三维结构的初始值, 最后用基于“ $R-T$ ”模式的光束法平差进行优化。仿真数据和真实图像实验表明, 新算法精度高、稳定性好, 有效解决了传统算法中精度、算法复杂性与结果数值稳定性的矛盾, 为序列图像三维解析中的关键技术问题提供了解决方案。算法可广泛应用于航空测绘、工业检测、虚拟现实、动画制作以及军事侦察等领域。

关键词: 光学测量; 仿射近似; 序列图像; 矩阵分解; 光束法平差; 三线性张量; 三维结构

中图分类号: TP391 文献标识码: A

Three-Dimensional Structure Reconstruction from Image Sequence Based on Affine Approximation

Qiu Zhiqiang^{1,2} Liang Yonghui¹ Yu Qifeng¹

(¹ Department of Optical Information Science and Technology, College of Opto-Electronic Science and Engineering, National University of Defense Technology, Changsha 410073)
(² Unit 76321, Peoples Liberation Army, Guangzhou 510500)

Abstract: Three-dimensional structure reconstruction from image sequence is acquiring much attention in photogrammetry and computer vision. A new algorithm based on affine approximation is put forward to resolve the contradiction of accuracy and robustness in conventional algorithms. Firstly, it synthesizes those missing points in image sequence based on tri-linear tensor, and then rebuilds a new measurement matrix without missing elements, which is decomposed by the bilinear affine factorization method to recover the initial data for the three-dimensional structure of objects. Finally, the initial three-dimensional reconstruction is optimized by the “ $R-T$ ”-model-based bundle adjustment. Synthetic and real image tests show the new method presents more accurate and robust result and it can be applied to aerial surveying and mapping, industry inspection, virtual reality, cartoon making, military reconnaissance, etc..

Key words: optical measurement; affine approximation; image sequence; matrix factorization; bundle adjustment; tri-linear tensor; three-dimensional structure

1 引 言

在军事侦察、飞行器导航、航空制图、机器人控制、工业检测、灾害监测等领域, 从摄像机获取的目标或场景序列图像重建出用户感兴趣的三维信息得

到日益广泛的应用。从技术上说, 这些实际应用都涉及到一个重要课题——如何从序列图像重建目标三维结构。

近年来, 从序列图像重建目标三维结构的研究

* 国家自然科学基金(10472133)和中国博士后科学基金(2005038230)资助课题。

作者简介: 邱志强(1972-), 男, 湖南浏阳人, 博士后, 工程师, 主要从事精密光测图像处理、计算机视觉、摄影测量、自适应光学方面的研究。E-mail: qiuzhiqiang@hotmail.com

导师简介: 于起峰(1958-), 男, 山东威海人, 教授, 博士生导师, 主要从事精密光测图像处理技术、智能光信息处理技术方面的研究。E-mail: yuqifeng@vip.sina.com

收稿日期: 2006-08-22; 收到修改稿日期: 2006-11-14

引起了摄影测量学、计算机视觉等领域学者的关注,取得了一系列研究成果,其中有两种典型算法,分别是矩阵分解和光束法平差。

矩阵分解法^[1~3]利用图像点在图像序列与目标点序列中的双线性约束,建立图像测量矩阵,采用奇异值分解(SVD)将摄像机运动与目标结构分离,算法全线性求解,速度快。按摄像机投影模型形式,矩阵分解法可分为透视矩阵分解和仿射矩阵分解。Bill Triggs^[1], Peter Sturm^[2]分别提出了一种透视矩阵分解法。鉴于透视投影模型的非线性, Poelman 等^[3]提出了基于仿射投影模型的矩阵分解,即仿射矩阵分解。不管是透视矩阵分解还是仿射矩阵分解,其优点是相对非线性算法速度快,而且不需要未知数的初始值。缺点是由于线性求解受噪声影响比较大,特别是透视矩阵分解中有时由于噪声比较大而不能正确求得投影深度。同时,仿射矩阵分解采用仿射投影模型来近似透视投影模型,由此带来的模型误差也会使得结果精度偏低。因此矩阵分解法的应用受到一定限制,主要用在三维动画制作等重视视觉效果而非测量精度的场合。

在矩阵分解中,如果部分图像点缺失,将无法对图像测量矩阵直接进行奇异值分解。在解决图像点缺失问题上, Brandt^[4]和 Jacobs 等^[5]作了相应研究,大多集中于寻找测量矩阵的子阵或直接对存在缺失元素的测量矩阵进行近似分解。可以说,图像缺失点问题并未得到有效解决。

摄影测量学中的光束法平差^[6~8]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重建目标三维结构的算法。这种算法通过共线方程线性化,组成法方程,基于 Levenberg-Marquardt 优化准则求解,本质上是线性最小二乘平差法。经过多年的发展,光束法平差已经非常成熟,且速度快、精度高,在成图、勘测等工程实践中取得了广泛的应用,但它需要所有未知数的初始值。如果初始值选择不好,很容易收敛到错误解导致算法失败。特别是对于航空图像,线方位元素与角方位元素存在强相关^[7],算法的成功很大程度上依赖于有一个较好的初始值。

本文针对序列图像的三维解析问题,提出基于仿射近似的思路。

在新算法中,我们假定图像是基于仿射投影模型形成的,从而建立快速线性算法获得初始值。然后,考虑其它因素(如透视畸变、镜头像差、噪声等),也就是基于精确的透视投影模型,对初始值进一步

求精。这就是所谓仿射近似思想。

2 算法简介

本文主要通过摄影测量学与计算机视觉两大学科的结合而实现。摄影测量学已经建立了很成熟的体系和方法。而计算机视觉是 20 世纪 80 年代以后发展起来的。但从本质上说,摄影测量学和计算机视觉在应用目的与数学框架上是一致的,都是解决如何从二维图像数据获取目标三维信息的科学。本文将光束法平差作为结果求精算法与矩阵分解法结合起来,用矩阵分解法结果作为光束法平差的初始值,充分发挥两种算法的优势,同时克服它们的缺点,为从序列图像重建目标三维结构提供完整的解决方案。如图 1 所示,该算法可以分为 3 个步骤。

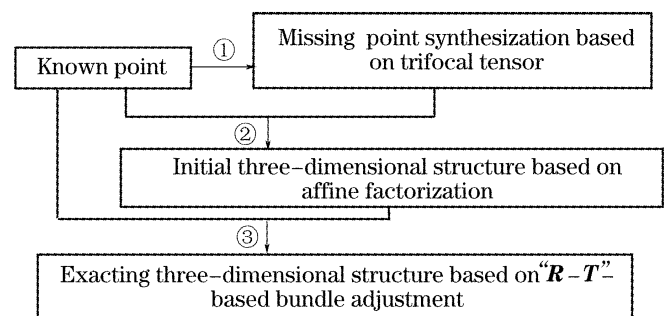


图 1 本文新算法主要步骤

Fig. 1 Primary steps of this new algorithm

1) 基于仿射三线性张量补齐缺失图像点

由于特征提取、目标遮挡等原因,在序列图像中,部分目标的图像点在某些图像中缺失是不可避免的,此时不能直接进行奇异值分解。与传统算法不同,新算法将图像合成中的三线性张量^[9~11]应用到图像点缺失问题中,基于仿射三线性张量合成图像点,补齐图像测量矩阵的所有缺失元素。仿射三线性张量比透视三线性张量约束和未知量更少,求解更稳定。

2) 用仿射矩阵分解法获得的目标三维结构作为初始值

以补齐的图像测量矩阵为对象,按经典的奇异值分解进行仿射矩阵分解,得到三维结构和摄像机运动的初始值。仿射矩阵分解采用摄像机几何成像的近似模型——仿射投影模型,建立多图多点约束,避免了投影深度的引入,特别适用于航空或远景图像的分析。与透视矩阵分解算法相比,仿射矩阵分解不仅算法简单,而且数值上更加稳定。

3) 用基于“ $R-T$ ”模式的光束法平差对结果进行优化。

由于补齐的图像点毕竟不是已知信息,存在一定程度的误差,因此在光束法平差中,对合成图像点不予考虑。光束法平差本身对于图像点是否缺失是没有影响的。

可以看出,为了提高算法的稳定性,新算法不仅在矩阵分解中采用的是仿射矩阵分解,而且采用自由度更小的基于仿射的三线性张量来取代通常的透视三线性张量。

2.1 数学推导

摄像机成像通常采用透视投影模型^[7,12~15],其基本关系式为

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

其中 $(X, Y, Z)^T$ 为目标点世界坐标, $(u, v)^T$ 为对应的图像点坐标, \mathbf{R} (3×3 阶正交矩阵)与 \mathbf{T} 分别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^[16]。(1)式可改写为

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

其中 $(T_x, T_y, T_z)^T$ 为 \mathbf{T} 的分量, (r_1, \dots, r_9) 为 \mathbf{R} 的元素。这就是后文将要提到的基于“ $\mathbf{R}-\mathbf{T}$ ”模式的共线方程。

2.1.1 基于仿射三线性张量补齐缺失图像点

如果给定某目标点在3幅图像中的图像点分别为 (u, v) , (u', v') 和 (u'', v'') 。采用仿射投影模型可列方程如下^[16]

$$\begin{cases} k_5 u + l_5 v - m_8 u' + m_5 = m_6 v'', \\ k_4 u + l_4 v - m_7 u' + m_4 = m_6 u'', \\ k_2 u + l_2 v - m_8 u' + m_2 = m_3 v'', \\ k_1 u + l_1 v - m_7 u' + m_1 = m_3 u'', \end{cases} \quad (3)$$

其中 $(k_1, \dots, k_5, l_1, \dots, l_5, m_1, \dots, m_6)$ 称为仿射三线性张量。从上式可线性求解仿射三线性张量的各元素。本算法采用仿射三线性张量而非透视三线性张量的原因是,仿射三线性张量只有3个约束条件和15个待求未知量,大大少于透视三线性张量的约束条件数和未知量个数,这对减小结果误差是大有好处的。求解透视三线性张量,需要的点数为 n ,则应满足 $4n \geq 26$,最少需要7点。而对仿射三线性张量,则应满足 $4n \geq 15$,最少只需要4组对应点就可求解

三线性张量。因此,采用仿射三线性张量,方程中的未知数更少,因此数值上更稳定。

求得仿射三线性张量后,利用下式可直接求解缺失图像点的坐标

$$\begin{cases} u'' = \frac{m_3(k_1 u + l_1 v - m_7 u' + m_1)}{m_3^2 + m_6^2} + \frac{m_6(k_4 u + l_4 v - m_7 v' + m_4)}{m_3^2 + m_6^2}, \\ v'' = \frac{m_3(k_2 u + l_2 v - m_8 u' + m_2)}{m_3^2 + m_6^2} + \frac{m_6(k_5 u + l_5 v - m_8 v' + m_5)}{m_3^2 + m_6^2}. \end{cases} \quad (4)$$

2.1.2 基于仿射矩阵分解获得目标三维结构的初始值

在序列图像中,假设图像幅数和目标点数分别为 m 和 n 。补齐缺失图像点后,可以认为 n 个目标点在 m 幅图像中均有对应的图像点。令第 j 个目标点在第 i 幅图像上的图像点坐标为 $p_{ij}(u_{ij}, v_{ij})$ ($i = 1, \dots, m, j = 1, \dots, n$),我们的目的是从这些图像点确定所有目标点组成的三维结构。

令 $\bar{u}_i = \frac{1}{n} \sum_{j=1}^n u_{ij}$, $\bar{v}_i = \frac{1}{n} \sum_{j=1}^n v_{ij}$ ($i = 1, \dots, m, j = 1, \dots, n$),将图像点坐标 (u_{ij}, v_{ij}) 作变换, $u'_{ij} = u_{ij} - \bar{u}_i$, $v'_{ij} = v_{ij} - \bar{v}_i$,按如下方式组成 $2m \times n$ 阶矩阵

$$\mathbf{W} = \begin{bmatrix} u'_{11} & \cdots & u'_{1n} \\ \vdots & \ddots & \vdots \\ u'_{m1} & \cdots & u'_{mn} \\ v'_{11} & \cdots & v'_{1n} \\ \vdots & \ddots & \vdots \\ v'_{m1} & \cdots & v'_{mn} \end{bmatrix}, \quad (5)$$

\mathbf{W} 称为图像测量矩阵。在无噪声条件下, \mathbf{W} 的秩为3,有噪声时, \mathbf{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 \mathbf{W} 进行奇异值分解,

$$\mathbf{W} = \mathbf{L} \mathbf{\Lambda} \mathbf{B}, \quad (6)$$

其中 \mathbf{L} 为左奇异向量矩阵, $\mathbf{\Lambda}$ 为奇异值对角矩阵, \mathbf{B} 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 \mathbf{L} 取前3列, $\mathbf{\Lambda}$ 取前3行前3列, \mathbf{B} 取前3行,式(6)仍近似成立。也就是说,如果令

$$\hat{\mathbf{M}} = \mathbf{L} \begin{bmatrix} \sqrt{\sigma_1} & 0 & 0 \\ 0 & \sqrt{\sigma_2} & 0 \\ 0 & 0 & \sqrt{\sigma_3} \end{bmatrix}$$

和

$$\hat{\mathbf{S}} = \begin{bmatrix} \sqrt{\sigma_1} & 0 & 0 \\ 0 & \sqrt{\sigma_2} & 0 \\ 0 & 0 & \sqrt{\sigma_3} \end{bmatrix} \mathbf{R},$$

其中 $\sigma_1 \geq \sigma_2 \geq \sigma_3$ 为 \mathbf{W} 的 3 个较大奇异值, 则有

$$\mathbf{W} = \hat{\mathbf{M}}\hat{\mathbf{S}}, \quad (7)$$

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

如果采用仿射投影模型, 加上摄像机内参量等约束, 可以从 $2m \times 3$ 阶矩阵 $\hat{\mathbf{M}}$ 求出摄像机运动量, 从 $3 \times n$ 阶矩阵 $\hat{\mathbf{S}}$ 求出目标三维结构, 具体方法参见文献[3]。至此, 通过仿射矩阵分解实现了运动与结构的分离(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), 从而获得了目标三维结构的初始值。

2.1.3 基于“ $\mathbf{R}-\mathbf{T}$ ”模式的光束法平差

由于用到了仿射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, 如镜头像差、图像点检测误差等^[17~19], 因此考虑用光束法平差来精确求解。

在通常的摄影测量学文献[6,7]中, 共线方程可改写成矩阵形式

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{R}'(\mathbf{X} - \mathbf{T}') = \begin{bmatrix} r'_1 & r'_2 & r'_3 \\ r'_4 & r'_5 & r'_6 \\ r'_7 & r'_8 & r'_9 \end{bmatrix} \begin{bmatrix} X - T_x \\ Y - T_y \\ Z - T_z \end{bmatrix}. \quad (8)$$

可以认为, 图像点齐次坐标 $(u, v, 1)$ 是目标点世界坐标 (X, Y, Z) 经过先平移 \mathbf{T} 后旋转 \mathbf{R} 得到的, 本文称之为基于“ $\mathbf{T}-\mathbf{R}$ ”模式的共线方程。与此不同, (1) 式可以认为 $(u, v, 1)^T$ 是 $(X, Y, Z)^T$ 经过先旋转 \mathbf{R} 后平移 \mathbf{T} 得到的, 因此, 本文称式(2)为基于“ $\mathbf{R}-\mathbf{T}$ ”模式的共线方程。在将共线方程式(2)线性化时, 不能直接搬用摄影测量学文献中的公式, 应该根据(2)式分别求出 u, v 对各变量的偏导数。

求得偏导数后, 可以建立误差方程式,

$$\begin{aligned} \delta u &= \frac{\partial u}{\partial T_x} \Delta T_x + \frac{\partial u}{\partial T_y} \Delta T_y + \frac{\partial u}{\partial T_z} \Delta T_z + \\ &\quad \frac{\partial u}{\partial \omega} \Delta \omega + \frac{\partial u}{\partial \phi} \Delta \phi + \frac{\partial u}{\partial \kappa} \Delta \kappa + \\ &\quad \frac{\partial u}{\partial X} \Delta X + \frac{\partial u}{\partial Y} \Delta Y + \frac{\partial u}{\partial Z} \Delta Z, \end{aligned} \quad (9)$$

$$\begin{aligned} \delta v &= \frac{\partial v}{\partial T_x} \Delta T_x + \frac{\partial v}{\partial T_y} \Delta T_y + \frac{\partial v}{\partial T_z} \Delta T_z + \\ &\quad \frac{\partial v}{\partial \omega} \Delta \omega + \frac{\partial v}{\partial \phi} \Delta \phi + \frac{\partial v}{\partial \kappa} \Delta \kappa + \\ &\quad \frac{\partial v}{\partial X} \Delta X + \frac{\partial v}{\partial Y} \Delta Y + \frac{\partial v}{\partial Z} \Delta Z, \end{aligned}$$

用矩阵表示上式,

$$\mathbf{V} = [\mathbf{B} \quad \mathbf{C}] \begin{bmatrix} \mathbf{X} \\ \mathbf{t} \end{bmatrix} - \mathbf{L}, \quad (10)$$

按最小二乘法组成法方程式

$$\begin{bmatrix} \mathbf{B}^T \mathbf{B} & \mathbf{B}^T \mathbf{C} \\ \mathbf{C}^T \mathbf{B} & \mathbf{C}^T \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix} - \begin{bmatrix} \mathbf{B}^T \mathbf{L} \\ \mathbf{C}^T \mathbf{L} \end{bmatrix} = 0, \quad (11)$$

以仿射矩阵分解法求得的三维结构结果作为初始值, 从上式可计算出 \mathbf{B}, \mathbf{C} 和 \mathbf{L} , 而后进行迭代求解 \mathbf{X} 和 \mathbf{T} , 直至改正数小于设定的阈值则停止, 实践证明一般迭代 4~5 次就足够了。详细过程可参见文献[6~8]。

在(8)式中, \mathbf{R}' 与 \mathbf{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, \mathbf{R} 与 \mathbf{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”^[13,15], 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

$$\mathbf{R} = \mathbf{R}'^{-1}, \quad \mathbf{T} = -(\mathbf{R}')^{-1} \mathbf{T}'. \quad (12)$$

2.2 算法步骤总结

- 1) 根据已知图像点用(3)式计算仿射三线性张量;
- 2) 根据仿射三线性张量用(4)式合成缺失图像点;
- 3) 补齐缺失元素后, 重新建立图像测量矩阵(5)式;
- 4) 根据(7)式进行奇异值分解得到仿射结构 $\hat{\mathbf{X}}$, 并求出目标三维结构的初始值;
- 5) 剔除合成的图像点, 建立误差方程式(10)和法方程式(11), 并进行迭代求解。

3 算法的实验验证

为了验证本文算法, 分别设计了真实图像实验和仿真实验。真实图像实验可以观察算法结果的合理性, 主要用来验证算法的正确性。仿真实验由于具有理论值作为参照, 还可以改变可能的影响因素来观察算法结果的变化趋势, 可以全面检验算法的精度、稳定性等性能。

3.1 真实图像实验 1

作者用数码相机从不同角度拍摄了一栋房屋的 28 幅图像, 组成一个图像序列, 用最小核值相似区(SUSAN)角点检测算法^[15]提取了每幅图像中的角点作为特征点。在图 2 中列出了其中的第 2 和第 20 幅, 角点已用“十”字丝标出。然后用本文新算法

进行处理,求得屋顶的三维结构图,其俯视图和侧视图表示在图 2(c)和图 2(d)中。可以看出,重建结果与真实情形大致相符。

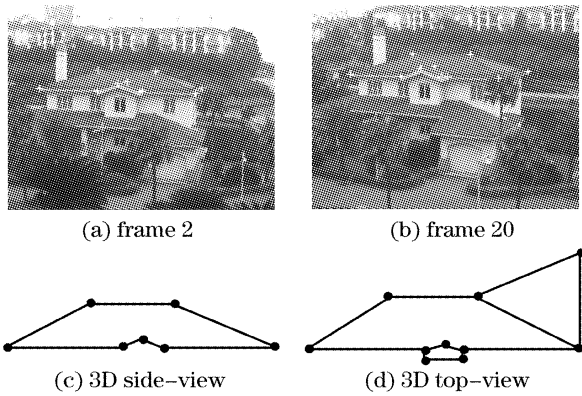


图 2 房屋序列图像及三维重建结果

Fig. 2 House image sequence and its three-dimensional structure

3.2 真实图像实验 2

作者用数码相机从不同角度拍摄了所在单位一栋教学楼的一角,如图 3 所示,以墙面装饰瓷砖角点为目标点作为测量对象。图 3(a)和图 3(b)分别表示了序列图像中的第 1 和第 11 幅。首先用“十”相关法^[15]检测瓷砖缝横纵线交点作为特征点,然后用本文算法进行处理,可以得到这些点的三维结构。图 3(c)和图 3(d)分别为重建结果的侧视图和俯视图,用直线拟合法可得到左右墙面夹角为 136° ,与实际测量值 139° 的相对误差为 2.1%。

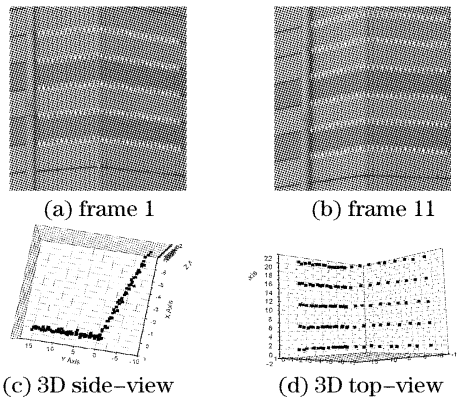


图 3 教学楼墙角序列图像及三维重建结果

Fig. 3 School building corner image sequence and its three-dimensional structure

3.3 仿真实验

为了检验本文算法对远景图像(如航空图像)的适用性,按如下航空成像条件设计仿真实验。如图 4 所示,有一个由 2 块矩形网格板以夹角 135° 拼接而成的目标,矩形网格板长宽均为 100 m,网格板上有正交等距网格线,纵横网格线交点就是观测的目标点(如点 P)。用仿真摄像机在 25 个不同位置

(C_1, \dots, C_{25})对这些点投影。摄像机焦距为 1 m,主点位于图像中心。光心位于平面 $Z = -1000$ m 的平面内(相当于物距为 1000 m),随机散布于边长为 100 m 的正方形区域,光轴欧拉角按均值为 0,方差为 0.15 rad 的高斯随机数产生。根据透视投影计算出图像点坐标,并添加方差为 0.5 pixel 的高斯噪声,然后用本文新算法进行处理获得目标点的三维位置。

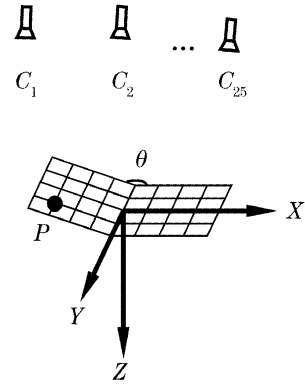


图 4 仿真实验

Fig. 4 Simulation experiments

关于误差的评价准则,分别采用以下两种。一是三维定位误差。由三维结构加上绝对定向和全局比例缩放后,就可得到每个目标点的三维位置。为便于研究误差的变化规律,实验中均将得到的三维结构转化为每个目标点的三维位置。首先进行重建,得到三维位置与实际三维位置的距离,就可以获得三维定位误差量。二是二维重投影残差和。采用重投影图像点与原图像点距离的残差和,如果所恢复得到的目标点结构和摄像机参量都足够准确,则按恢复的参量重投影得到的图像点将与原图像点重合。

表 1 列出了分别用本文算法和传统算法得到的结果。可以看出,本文算法得到的三维定位误差要小于传统算法的误差。

3.4 对算法性能的进一步分析

为了进一步研究算法性能,分别将可能对算法误差有影响的因素,如图像噪声水平、目标点数、摄像机个数、目标点间距、摄像机间距、目标点共面程度和航高按一定规律进行变化,观察误差的变化规律。考虑到随机因素的影响,每个条件下均重复实验 50 次,取平均值作为实验结果。实验结果分别按两种准则统计误差。

图 5(a)和图 6(a)分别为三维定位误差和二维重投影残差和随图像点噪声的变化曲线。可以看出,随着图像点噪声增加,误差也增加,但本文算法误差要小于传统算法的误差。在实际处理中应设法提高图像点的检测精度。

表 1 仿真实验结果

Table 1 Results of simulation experiments

Point	Theoretical value			Results of this new algorithm				Results of traditional algorithm			
	X	Y	Z	X	Y	Z	Errors	X	Y	Z	Errors
1	0.000	-40.000	0.000	0.0063	-40.070	-0.004	0.070	0.016	-39.952	0.207	0.213
3	0.000	0.000	0.000	-0.002	-0.001	0.000	0.003	-0.002	-0.001	0.000	0.003
5	0.000	40.000	0.000	-0.006	40.071	0.006	0.071	-0.018	39.951	-0.311	0.315
7	18.478	-20.000	-7.654	18.480	-20.036	-7.680	0.045	18.467	-20.037	-7.515	0.144
9	18.478	20.000	-7.654	18.473	20.036	-7.678	0.043	18.451	19.911	-7.779	0.156
11	36.955	-40.000	-15.307	36.960	-40.070	-15.342	0.079	36.937	-40.070	-15.007	0.309

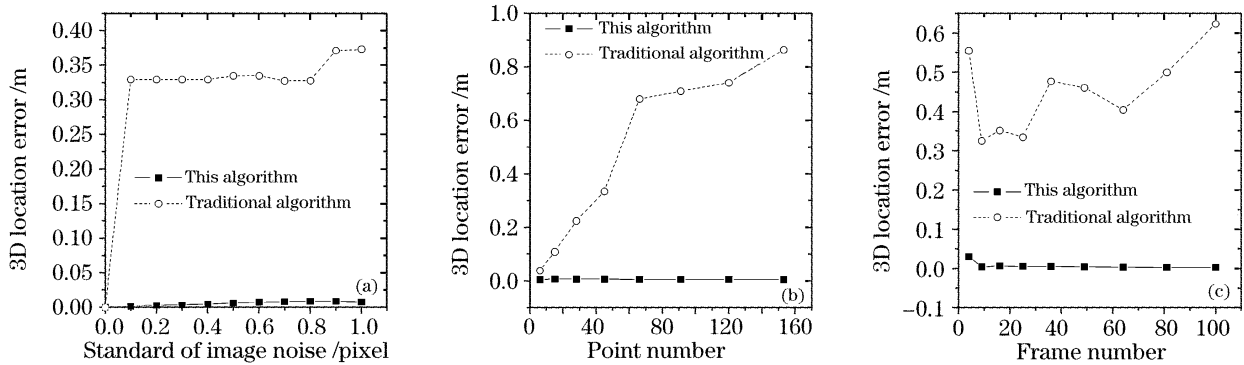


图 5 三维定位误差随(a)图像点噪声,(b)目标点个数以及(c)摄像机个数的变化曲线

Fig. 5 Curve of three-dimensional location error versus (a) standard-error of image noise, (b) point number and (c) frame number

图 5(b)和图 6(b)分别为三维定位误差和二维重投影残差和随目标点个数的变化曲线。可以看出,随着目标点个数增加,本文算法的误差变化不大,但传统透视矩阵分解法的误差会增大。这说明目标点个数太多,不仅不能提高信息量,反而会影响稳定性,这一现象在传统算法中表现得尤为明显。而本文算法采用仿射矩阵分解,减小了模型的复杂

性,从而在一定程度上改善了算法的数值稳定性。

图 5(c)和图 6(c)分别为三维定位误差和二维重投影残差和随摄像机个数(图像幅数)的变化曲线。可以看出,随着摄像机个数增加,本文算法的误差变化不大,但传统算法的误差会增大。其原因分析同上述目标点数的分析。

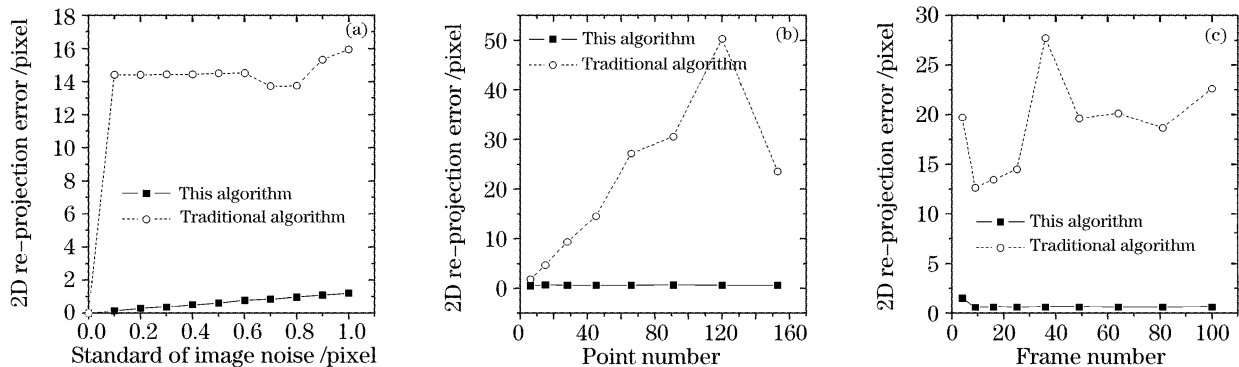


图 6 重投影残差和随(a)图像点噪声,(b)目标点个数以及(c)摄像机个数的变化曲线

Fig. 6 Curve of re-projection residual versus (a) standard-error of image noise, (b) point number and (c) frame number

4 结 论

本文提出了一种基于仿射近似从序列图像重建目标三维结构的方法。只要提取并匹配出图像特征点在序列图像上的位置,建立多图多点约束,就能获得物体的空间三维结构信息。采用仿射近似思想,

将矩阵分解法与光束法平差有机结合,得到了满意的精度与数值稳定性。

主要成果:

1)仿射近似的运用。在本文中,有两处用到仿射近似,一是基于仿射三线性张量合成缺失图像点,

二是基于仿射矩阵分解获得目标三维结构的初始值。仿射近似的运用大大提高了算法的数值稳定性；

2) 摄影测量学和计算机视觉中算法的结合。将矩阵分解与光束法平差结合起来,这样做的好处是将矩阵分解法的结果作为光束法平差的初始值,而通过光束法平差又可以减小仿射近似引入的误差,不但充分发挥了两种算法的优点,而且克服了各自的不足；

3) 将广泛用在图像合成中的三线性张量原理运用到图像缺失点的合成中,形成了比传统算法简单得多的新算法；

4) 基于“ $R-T$ ”模式共线方程概念的提出,并推导了其线性化公式,从而建立了基于“ $R-T$ ”模式的光束法平差,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有机结合；

5) 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。

本文算法可广泛用于航空摄影测量、三维建模、机器视觉、虚拟现实、动画制作以及军事侦察、航天摄影测量等领域。

参 考 文 献

- 1 Bill Triggs. Some notes on factorization methods for projective structure and motion[J]. <http://www.inrialpes.fr/lear/people/triggs/pub/Triggs-factorization-notes-oct98.ps.gz>
- 2 Peter Sturm, Bill Triggs. A factorization based algorithm for mult-image projective structure and motion[C]. *4th European Conference on Computer Vision*, 1996. 709~720
- 3 C. J. Poelman, T. Kanade. A paraperspective factorization method for shape and motion recovery[J]. *IEEE Transactions on Pattern Analysis And Machine Intelligence*, 1997, **19**(3): 206~218
- 4 Sami Brandt. Closed-form solutions for affine reconstruction under missing data [C]. In *Proceedings of the Statistical Methods for Video Processing Workshop*, in conjunction with ECCV, 2002. 109~114
- 5 D. W. Jacobs. Linear fitting with missing data for structure-from-motion[J]. *Computer Vision and Image Understanding*, 2001, **82**(1): 57~81
- 6 Zhu Zhaoguang, Sun Hu, Cui Bingguang. *Photogrammetry* [M]. Beijing: Surveying and Mapping Press, 1995. 221~246 (in Chinese)
朱肇光,孙 护,崔炳光. 摄影测量学[M]. 北京: 测绘出版社, 1995. 221~246
- 7 Li Deren, Zheng Zhaobao. *Analytic Photogrammetry* [M]. Beijing: Surveying and Mapping Press, 1992. 159~173 (in Chinese)
李德仁,郑肇葆. 解析摄影测量学[M]. 北京: 测绘出版社, 1992. 159~173
- 8 Chen Xiru. Historical backgrounds and present state of the least squares method [J]. *J. Graduate School, Academia Sinica*, 1998, **15**(1): 4~11 (in Chinese)
陈希儒. 最小二乘法的历史回顾与现状[J]. 中国科学院研究生院学报, 1998, **15**(1): 4~11
- 9 Shen Peyi, Wang Wei, Wu Chengke. Trilinear constraint and its application in image synthesis[J]. *J. Xidian University*, 1999, **26**(4): 431~435 (in Chinese)
沈沛意,王 伟,吴成柯. 三线性约束及其在图像合成中的应用[J]. 西安电子科技大学学报, 1999, **26**(4): 431~435
- 10 Li Baozhou, Deng Yanpin, Li Jiegu. Image rendering method based on tri-linear tensor[J]. *J. Data Acquisition & Processing*, 2000, **15**(3): 320~323 (in Chinese)
李保洲,邓雁萍,李介谷. 基于三线性张量生成图像的方法[J]. 数据采集与处理, 2000, **15**(3): 320~323
- 11 Li Baozhou, Deng Yanpin, Li Jiegu. Algebraic constraint-based view rendering[J]. *J. Shanghai Jiaotong University*, 2000, **34**(5): 699~701 (in Chinese)
李保洲,邓雁萍,李介谷. 基于代数约束的图像生成[J]. 上海交通大学学报, 2000, **34**(5): 699~701
- 12 Yu Qifeng, Qiu Zhiqiang, Sun Xiangyi *et al.*. Three dimensional movement measurement for the wing of a flying airplane[J]. *Acta Mechanica Sinica*, 2003, **19**(6): 575~578
- 13 Qiu Zhiqiang, Yu Qifeng. 3D building structure from image sequence[J]. *Engineering J. Wuhan Univeristy*, 2003, **36**(3): 13~15 (in Chinese)
邱志强,于起峰. 从建筑物序列图像恢复三维结构[J]. 武汉大学学报(工学版), 2003, **36**(3): 13~15
- 14 Zeng Jiyong, Su Xianyu. A new rectification method of stereo pairs without calibration for camera [J]. *Acta Optica Sinica*, 2004, **24**(5): 628~632 (in Chinese)
曾吉勇,苏显渝. 一种无相机标定的立体图像对校正新方法[J]. 光学学报, 2004, **24**(5): 628~632
- 15 Yu Qifeng, Lu Hongwei, Liu Xiaolin. *Image Based Precision Measurement and Movement Measurement* [M]. Beijing: Science Press, 2002. 116~120 (in Chinese)
于起峰,陆宏伟,刘肖琳. 基于图像的精密测量与运动测量[M]. 北京: 科学出版社, 2002. 116~120
- 16 Qiu Zhiqiang. *Three-Dimensional Structure Reconstruction from Aerial Images Based on Affine Approximation* [D]. Changsha: National University of Defense Technology, 2004. 10, 13~16 (in Chinese)
邱志强. 基于仿射近似从航空图像重建目标三维结构[D]. 长沙: 国防科学技术大学, 2004. 10, 13~16
- 17 Yu Qifeng, Shang Yang, Wang Ying *et al.*. Analytical system of trajectory camera photo plates and accurate calibration of common flat scanners[J]. *Acta Optica Sinica*, 2004, **24**(1): 104~109 (in Chinese)
于起峰,尚 洋,王 颖等. 弹道相机图像分析系统与普通扫描仪精确标定[J]. 光学学报, 2004, **24**(1): 104~109
- 18 Zhao Chuangxin, Xu Jinliang, Zhang Yongli *et al.*. Three-dimensional reconstruction of free-flying insect based on single camera [J]. *Acta Optica Sinica*, 2006, **26**(1): 61~66 (in Chinese)
赵创新,徐进良,张永立等. 基于单摄像机的昆虫自由飞行参量三维重构[J]. 光学学报, 2006, **26**(1): 61~66
- 19 Liu Zhangwen, Gu Tianxiang. Three-dimensional measurement of object by using gray gradient of CCD image[J]. *Acta Optica Sinica*, 2003, **23**(11): 1384~1388 (in Chinese)
刘章文,古天祥. 利用 CCD 图像的灰度梯度实现物体三维测量[J]. 光学学报, 2003, **23**(11): 1384~1388