

基于多视图图像的摄像机自标定方法^{*}

唐秋虎, 张志毅

(西北农林科技大学信息工程学院, 陕西 杨凌 712100)

摘要:摄像机标定是三维重建时的必要步骤。传统的标定方法对设备要求高、操作繁琐,而自标定方法虽然简便,但精度不高,会严重影响三维重建的效果。因此,越来越需要一种操作简便并且精度高的自标定方法。采用 SIFT 特征点匹配算法,根据多视序列图像中对应点间的相互关系,利用光束法平差,提出了一种基于局部-全局混合优化的迭代优化方法。针对图像匹配量大的问题,提出了一种邻域内图像互匹配方法来降低时间代价。实验表明,本文提出的多摄像机自标定方法是一种有效的高精度方法,采用的邻域内图像互匹配技术能很好地降低图像匹配的时间消耗。根据多视图图像的对应点间相互关系,充分利用局部-全局优化的思想,通过混合优化的方法得到相机参数,对比现有自标定算法,本文给出的方法有较高的精度和鲁棒性。

关键词:摄像机自标定;多视图图像;图像匹配;光束平差算法

中图分类号:TP391.41

文献标志码:A

doi:10.3969/j.issn.1007-130X.2017.04.019

Camera self-calibration based on multiple view images

TANG Qiu-hu, ZHANG Zhi-yi

(College of Information Engineering, Northwest A & F University, Yangling 712100, China)

Abstract: Camera calibration is an essential step of 3D reconstruction. Conventional calibration methods require high precision equipment and sophisticated operations. Compared with them, camera self-calibration is simple but has low precision, which leads to significant performance degradation of 3D reconstruction. Therefore, there is a growing need for a simple and accurate high precision self-calibration method. By using the bundle adjustment algorithm and SIFT points matching relationship, we propose a local-global hybrid iterative optimization method. As for the large number of matching features, we propose a neighborhood image matching method, which can significantly reduce the matching time under the premise of maintaining accuracy. Experimental results show that the proposed method is effective and accurate, and it can reduce the time consumption of image matching. Based on the relationship between corresponding matching points in multi view images, our method makes full use of the local-global hybrid idea to compute the parameters of the camera. Compared with other existing methods, it is more robust with higher precision.

Key words: camera self-calibration; multiple view images; image matching; bundle adjustment algorithm

^{*} 收稿日期:2015-12-30;修回日期:2016-03-04

基金项目:国家 863 计划(2013AA10230402);国家自然科学基金(61303124)

通信地址:712100 陕西省杨凌农业示范区西北农林科技大学信息工程学院

Address: College of Information Engineering, Northwest A & F University, Yangling Agricultural Demonstration Area, Yangling 712100, Shaanxi, P. R. China

1 引言

摄像机标定是计算机视觉领域的一项关键技术,它是从二维图像提取三维空间信息必不可少的步骤,这项技术已被广泛应用于3D结构重建、导航、视觉监控等领域^[1]。传统方法对设备要求高,而且操作繁琐。Zhang^[2]的平面标定方法虽然较传统的标定方法对设备要求更加简便并且精度高,但需要一个棋盘格作为辅助设备,依然不能达到标定的自动化和简便化。利用未标定的图像序列进行三维重建是当前的热门研究领域之一^[3]。利用多视图来标定相机,孙永荣等^[4]提出了一种多视图场景下的空间点位置视觉标定方法,利用棋盘模板确定某幅图像的场景位姿参数,而后利用多视图的场景位姿参数解算空间点的三维位置坐标。该方法虽然精度比较高,由于需要使用模板进行辅助标定,导致该方法同样不够自动化和简便化。对基于多幅图像的相机自标定技术,通常是对每两幅图像进行投影矩阵的估计,然后分别对得到的投影矩阵进行优化,最后进行全局的优化得到最终的相机参数。目前常用的优化算法是More^[5]提出的LM算法(Levenberg-Marquardt Algorithm),它是使用最广泛的非线性最小二乘算法,同时保留了梯度法和牛顿法的优点。常见的基于多幅图像进行相机标定,可以通过求解Kruppa方程的方式,或者通过分层自标定的方式,或是通过绝对二次曲面的方式。Hu等^[6]提出利用图像序列结合手持摄像机实现三维重建和Nguyen等^[7]提出利用无约束无标定的图像结合手持摄像机实现三维重建,采用SIFT(Scale-Invariant Feature Transform)算法检测并匹配特征点,然后估算出投影矩阵实现相机标定,从而实现三维点云的重建以表示输入图像特征点的几何位置。这种通过对图像进行两两考虑,先局部优化,最后整体考虑进行全局优化的自标定方法虽然得到的相机参数精度较高,使得后面得到的三维重建结果比较满意,但是依然不够精确。例如,文献[7]中在计算出摄像机的相机参数后,最后通过光束平差法对所有相机参数进行优化得到一个全局的最优解。所谓光束平差法,就是能从多个视角信息中提取出三维点的坐标以及各个摄像机的相对位置和光学信息的过程。该方法将所有图像一起考虑,通过增量式的叠加,使得可以将所有摄像机的相机参数进行全局的优化得到一个全局最优解。但是,文献[7]的处理方法只是在最后采

用了全局优化,依然没有充分利用这种全局优化的思想。同样地戴嘉境^[8]提出的基于多幅图像的三维重建理论及算法研究中也采用光束平差算法进行一个最终的全局优化。

针对传统的标定方法和自标定方法的不足,本文提出了一种基于多视图图像的摄像机标定方法。它先采用局部-全局相混合的优化算法实现摄像机空间位置和姿态的自动标定,即通过序列图像得到摄像机内、外参数,最后利用光束平差法进行全局优化。根据投影原理可知,即使知道一幅图像上的所有像点的坐标,是无法得到投影矩阵的,也无法得到三维空间点。如果利用多幅图像的几何约束关系,是可以唯一确定三维点的位置的。因此,空间中的一个三维点必然和这些图像上所对应的像点相关联。所以,需要对序列图像进行特征点匹配,匹配点对的精确性对摄像机标定的准确性至关重要。Lowe^[9]提出的SIFT算法对图像旋转、尺度变化和亮度变化具有不变性,对视角变化、仿射变换、噪声也同样有很好的稳定性,使得该算法在基于图像序列的三维重建方面有显著优势,因此本文使用SIFT匹配算法进行图像间的特征点匹配。

2 自标定算法

2.1 双视图图像的对极几何

本质上,双视图图像之间的对极几何是图像平面与以基线为轴的平面束的相交几何^[10]。由江泽涛等^[11]提出的基于本质矩阵的摄像机自标定方法可知,本质矩阵 E 包含在物理空间中两个摄像机相关的旋转和平移信息,基础矩阵 F 除了包含 E 的信息外,还包含了两个摄像机的内参数。由于 F 包含了这些内参数,因此它可以在像素坐标系将两台摄像机关联起来。

假定两幅图像是同一摄像机在不同方位拍摄的,并且世界坐标系和左图对应的摄像机坐标系对齐。假设左图上一个点 $p=(x_p, y_p, 1)^T$ 和右图上一个点 $q=(x_q, y_q, 1)^T$ 是一对匹配点,有: $q^T F p=0$,展开得到:

$$(x_q x_p, x_q y_p, x_q, y_q x_p, y_q y_p, y_q, x_p, y_p, 1) f = 0 \quad (1)$$

其中, $f=(f_{00}, f_{01}, f_{02}, f_{10}, f_{11}, f_{12}, f_{20}, f_{21}, f_{22})^T$ 。

n 对匹配点对就可以构成一个线性方程组,通常采用多于8对匹配点对即可计算出基础矩阵 F 。在模型估计中,要解出准确的模型,就要消除外点

(outliers)对估计的影响,要达到这个目的,最直观的想法就是从原始数据中找出一组不包含 outliers 的数据抽样来进行参数估计^[12],这就是 RANSAC (RANdom SAmple Consensus)算法的思想。本文通过改进 RANSAC 算法优化下的 8 点法^[13]来求基础矩阵 F 。算法步骤如下:

Step 1 根据置信概率设置抽样停止条件。

Step 2 随机从两图像的匹配特征点中抽取 8 对匹配点组成一个抽样。

Step 3 在步骤 2 选择的随机抽样中,用这 8 个点采用线性最小二乘法得到一个基础矩阵 F_{tmp} 。

Step 4 用两图像所有的匹配对来检验每个 F_{tmp} ,获得每个 F_{tmp} 所对应的内点(inliers)数量。重复步骤 2 到步骤 4,直到满足置信概率为止。

Step 5 根据 inliers 数量选择对应的基础矩阵 F 。

Step 6 找出这个最优的基础矩阵 F 对应的所有 inliers,并用这些 inliers 采用非线性最小二乘法计算得到最终的基础矩阵 F 。

根据成像原理,投影矩阵 P 是一个 3×4 阶矩阵,可以分解为: $P = K [R | T]$,其中, R 和 T 为外部参数矩阵,分别表示空间点相对摄像机的旋转和平移分量, K 是内部参数矩阵,表示为:

$$K = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

其中, (u_0, v_0) 为主点坐标,表示图像中心点坐标, f_x 和 f_y 分别是 x 和 y 方向上的焦距值。

由基本矩阵和本质矩阵的关系: $E = K^T F K$,可以很方便地求得本质矩阵 E 。此外,基础矩阵 F 还将用于 SIFT 特征点匹配时去除错误的匹配。

对两幅图像,假设以第一幅图像作为主相机,即 $[R_1 | T_1] = [I | 0]$ 。Hartley^[14]提出对本质矩阵 E 进行奇异值分解(SVD),可得到第二幅图对应相机的旋转和平移参数,即:

$$E = UDV^T \quad (3)$$

其中, U 和 V 都是 3×3 的正交矩阵, D 是一个 3×3 的矩阵,则第二幅图对应的相机有四种取值:

$$\begin{aligned} [R_r | T_r] &= [UWV^T | u_3] \\ [R_r | T_r] &= [UW^T V^T | u_3] \\ [R_r | T_r] &= [UWV^T | -u_3] \\ [R_r | T_r] &= [UW^T V^T | -u_3] \end{aligned} \quad (4)$$

其中, u_3 为 U 的最后一列,且:

$$W = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

再通过其他的匹配特征点,就能够唯一确定式(4)中的一个,就确定了第二幅图像对应的相机的旋转矩阵 R 和平移矩阵 T 。

2.2 图像匹配

即使是一幅分辨率比较小的图像,如果要匹配所有点,工作量也会非常大。如果能够在图像中找到一些特殊的点,这些点相比它周围的点有明显的特征,这样就会很大程度地降低图像匹配的复杂度。为降低 SIFT 算法的误匹配率,本文使用双向 SIFT 算法来实现图像匹配。采用 K 邻近算法来得到每个关键点在另一幅图像中的欧氏距离最短的两个点,即 NN(Nearest Neighbor)和 SCN(Second Closest Neighbor)。假设 key_1 和 key_2 分别表示两幅图像中的同一个 SIFT 特征点,其欧氏距离表示为:

$$D(key_1, key_2) =$$

$$\|key_1 - key_2\| = \sqrt{\sum_{i=1}^{dim} (key_{1i} - key_{2i})^2} \quad (6)$$

其中, dim 表示特征向量的维度(SIFT 特征向量为 128 维)。

图 1 显示了一个 SIFT 特征点及其匹配示例,用箭头表示特征点的尺度、方向和位置。一幅图像可以和任意其他图像间匹配特征点对,若对每两幅图像都匹配特征点对,则计算量将会随图像的增加而急剧增加。当绕着物体拍摄多幅图像时,对于其中一幅图像来说,其邻域内图像所对应的相机的旋转、平移都和该图像对应的相机的旋转和平移比较接近,图像间包含更多相同的信息,有更多的相似度,特征匹配点对也更多,对于远离该幅图像的那些图像则相似度较低。因此,只对一幅图像邻域内的图像进行匹配可节约时间。



Figure 1 An example of SIFT features and matching
图 1 SIFT 特征点及匹配示例

2.3 光束平差法

光束平差算法的初始值分别是相机投影矩阵的初始值和三维点的初始值。通过对数码相机拍

摄的图像提取 EXIF (Exchangeable Image File) 信息可以获取每幅图像对应相机的初始焦距值 f_0 。此外还需要获得相机的旋转矩阵、平移矩阵和三维点的初始值。对于初始相机对,通过上述方法计算得到旋转平移矩阵作为初始值,用 RANSAC 优化下的最小二乘法计算得到对应特征匹配点的三维点并用 LM 算法进行优化,将优化后的三维点作为初始三维点。

假设一共有 A 幅图像和 B 对特征匹配点对。因为每幅图像的投影矩阵的自由度为 11,每个三维空间点的自由度为 3,所以该算法需要计算 $11A+3B$ 个参数。定义 $\mathbf{P}_i (i=0,1,\dots,A-1)$ 表示第 i 幅图像所对应的相机投影矩阵;定义三维点集 $(X_j, Y_j, Z_j), j=0,1,\dots,B-1$, 表示 B 个三维点;定义 $(x_{ij}, y_{ij}), (x'_{ij}, y'_{ij})$ (其中, $0 \leq i \leq A-1, 0 \leq j \leq B-1$) 分别表示第 j 个三维点在第 i 幅图像上对应的特征点坐标和投影点坐标。然后定义投影误差 $error$ 表示一个三维点通过对应的投影矩阵投影到对应图像上的距离误差;定义重投影误差 $ERROR$ 为投影到每幅图像的所有投影距离误差的平方和相加之和,即:

$$ERROR = \sum_{i=0}^{A-1} \sum_{j=0}^{B-1} Invisible_{ij} [(x'_{ij} - x_{ij})^2 + (y'_{ij} - y_{ij})^2] \quad (7)$$

其中,BOOL 变量 $Invisible_{ij}$ 表示第 j 个三维点在第 i 幅图像中是否可见。

光束平差算法的基本原理是通过反复迭代,修正 $\mathbf{P}_i (i=0,1,\dots,A-1)$ 和 $(X_j, Y_j, Z_j) (j=0,1,\dots,B-1)$, 使重投影误差 $ERROR$ 收敛到一个很小的值。使用非线性最小二乘法 LM 算法来最小化重投影误差 $ERROR$ 。LM 算法需要对每一个待估参数求偏导,通过梯度来求取最小值,属于“爬山”法的一种。它的好处是同时具有牛顿法以及梯度法的长处。具体地,当 λ 很大时,步长就大致等于梯度下降法的步长,而当 λ 很小的时候,步长直接取牛顿法的步长。

由以上分析可知,一共有 $11A+3B$ 个参数需要修正,由于 $\mathbf{P}=\mathbf{K}[\mathbf{R}|\mathbf{T}]$, 可以定义焦距 f 的修正值为 Δf , 主点 (u_0, v_0) 的修正值为 $(\Delta u_0, \Delta v_0)$, 平移矩阵 \mathbf{T} 的修正值为 $\Delta \mathbf{T}=(\Delta t_0, \Delta t_1, \Delta t_2)^T$, 一共 6 个参数。因此,旋转矩阵 \mathbf{R} 需要修正 $11-6=5$ 个参数。注意到旋转矩阵 \mathbf{R} 是正交矩阵,即 $\mathbf{R}\mathbf{R}^T=\mathbf{I}$, 得到三个约束方程,又正交矩阵的每列的平方和为 1,因此旋转矩阵 \mathbf{R} 有 3 个自由度。定义 $\Delta \mathbf{R}$ 为

旋转矩阵 \mathbf{R} 的修正值,那么有:

$$(\mathbf{R} + \Delta \mathbf{R})(\mathbf{R}^T + \Delta \mathbf{R}^T) = \mathbf{I} \quad (8)$$

其中, \mathbf{I} 是 3×3 的单位矩阵。

展开得到:

$$(\Delta \mathbf{R}\mathbf{R}^T)^T = -\Delta \mathbf{R}\mathbf{R}^T \quad (9)$$

可见 $\Delta \mathbf{R}\mathbf{R}^T$ 是一个反对称矩阵,可以写成如下形式:

$$\Delta \mathbf{R}\mathbf{R}^T = \begin{bmatrix} 0 & -\varphi_2 & \varphi_1 \\ \varphi_2 & 0 & -\varphi_0 \\ -\varphi_1 & \varphi_0 & 0 \end{bmatrix} \quad (10)$$

令矩阵 $\psi=(\varphi_0, \varphi_1, \varphi_2)^T$, 有:

$$\psi \times \mathbf{I} = \begin{bmatrix} 0 & -\varphi_2 & \varphi_1 \\ \varphi_2 & 0 & -\varphi_0 \\ -\varphi_1 & \varphi_0 & 0 \end{bmatrix} = \Delta \mathbf{R}\mathbf{R}^T \quad (11)$$

因此有:

$$\Delta \mathbf{R} = \Delta \mathbf{R}\mathbf{R}^T \mathbf{R} = (\psi \times \mathbf{I}) \mathbf{R} = \psi \times \mathbf{R} \quad (12)$$

因此,旋转矩阵 \mathbf{R} 调整的时候只需要计算 ψ , 故只需要调整 3 个参数即可。所以,总共需要调整 $9A+3B$ 个参数即可。

2.4 算法分析与描述

对于多视图图像序列,各个图像可以看作是多个相机从不同角度对目标物体拍摄的照片。通过选取两个最好相机对开始标定算法,然后根据所有图像之间的特征点匹配和空间三维点的几何关系,依次挑选图像进行相机标定。一般地,在得到初始图像对和初始三维点后,对于其他相机图像,根据初始三维点集,通过最小二乘法分别计算得到新加入图像的相机参数并用 LM 算法进行优化,最后将得到的所有相机参数用光束平差算法进行全局优化。

一个更好的方法是,在得到初始相机对和初始三维点时,对于其他相机图像,采取按某种挑选机制每次挑选一张相机图像,根据已知的三维点集,计算得到对应的相机参数并用 LM 算法优化,并将每次计算新相机后对应的三维点加入到三维点集中,用于计算下一个新相机的参数,这样就可以充分利用各个相机图像对应的三维点。随着三维点的增多,对于新加入的相机来说,计算得到的相机参数将更为精确。同样更为精确的相机参数能得到其对应的更精确的三维点集,并加入到已知三维点集中。最后,将所有得到的相机参数通过光束平差算法进行最后的全局优化得到最优解。

根据以上分析可知,前一种方法没有采用一次加入相机图像的思想,计算每个相机的参数时,没

有充分利用各个图像所对应的空间三维点的几何关系,忽视很多有价值的三维点,计算结果会有较大误差。虽然采用了 LM 算法进行优化,但也是一种在局部范围内进行的优化,局部误差的累积会影响最终的整体优化;后一种方法采用依次加入相机图像,这种思想利用每次加入的图像扩充三维点集,使得在计算后一个即将加入的相机图像时,有更多的输入值保证结果的精确性。这样虽然减少了每次计算出的相机参数的误差值,但是每次采用 LM 算法进行优化时依然是在局部进行的,还是会有局部误差的累积。对于多视图,由于图像数量比较多,每次的局部优化次数就会很多,最后累积的误差也会较大,会影响最后的全局优化。假若在每次加入新的相机后,对所有已经计算出参数的相机进行一次全局的光束平差算法优化,这样可以更好地消除局部优化过程中的累积误差。整个过程可以看作,在每次加入新相机时采用 LM 进行局部优化得到新相机的相机参数,然后对所有已经求得的相机,用光束平差算法进行一个全局的优化,直到所有的相机都加入并且计算得到相机参数。这个过程中,在每次加入新的相机时,采用光束平差算法进行全局优化可以理解为在依次加入图像时,是一种将各个标定结果进行合并的行为,从而得到优化后的相机参数,这就是本文的局部-全局混合迭代优化的思想。算法步骤如下:

Step 1 对每幅图像进行 SIFT 特征点提取。

Step 2 对每幅图像,和其 K 邻域内的所有图像之间进行 SIFT 特征点双向匹配。

Step 3 图像匹配完成后,对所有匹配特征点进行追踪,这样同一个空间点在各个图像上对应的特征点就生成一个轨迹。

Step 4 选取两张图像,对应的相机作为初始相机对,选取标准是,这两幅图像在邻域内,匹配的特征点要足够多,同时两幅图像之间的基线足够宽。

Step 5 根据前文所述方法计算初始相机对光束法平差的初始参数,并利用光束平差法优化得到初始相机对的相机参数和三维点集。

Step 6 添加相机,根据步骤 3 中得到的特征点轨迹找到与三维点集中对应最多的图像作为下一个标定相机。根据三维点集,利用 RANSAC 优化下的最小二乘法计算得到新加入相机的相机参数并用非线性最小二乘法进行优化。

Step 7 根据新加入相机的参数和此前已矫正相机图像的匹配特征点,计算得到新的三维点并加入到三维点集中。

Step 8 将三维点集与新加入相机的参数作为光束平差法初始值,利用光束平差法进行全局性优化。

Step 9 重复步骤 6 到步骤 8 直到所有相机都矫正完成。

3 实验结果

文献[15,16]是经典的自标定算法,有着较高的精度和鲁棒性,重构效果相当好^[11]。本文采用从 <http://www.robots.ox.ac.uk/~vgg/data/data-mview.htm> 下载的分辨率为 512×768 的 Valbonne church 图像序列,即与文献[11]和文献[15-18]相同的图像序列,将此图像序列作为测试图像,即对同一摄像机进行标定。图 1 中的图像即是 Valbonne church 示例图。比较本方法对相机进行标定得到的结果与文献[11]和文献[15-18]所给的结果,如表 1 所示。

Table 1 Estimation result comparison between the proposed method and other methods
表 1 不同算法的标定结果对比

算法来源	f_x	f_y	u_0	v_0
文献[11]	670.457	679.232	248.648	392.561
文献[15]	682.84	682.843	255.999	383.999
文献[16]	679.285	681.345	258.802	383.188
文献[17]	667	693	241	398
文献[18]	619	699	234	372
本文	673.285	673.285	256.787	392.182
误差比较(1)/%	1.399	1.399	-0.308	-2.131
误差比较(2)/%	0.883	1.183	0.779	-2.347

文献[11]中通过将文献[15,16]的结果作为参考标准进行对比,得出其结果要好于文献[17,18]的结论。参照文献[15,16],由表 1 可知,本文所得结果比文献[11]和文献[17,18]更加接近标准结果。表 1 中误差比较(1)和误差比较(2)分别是将文献[15,16]作为标准结果时本文结果的相对误差,具体计算是分别用文献[15,16]所给结果和本文方法所得结果求差,然后比上文献[15,16]所给结果。可以看出,焦距 f 的相对误差大概在 1% 左右, u_0 的误差大概在 $\pm 0.5\%$ 左右, v_0 的相对误差大概在 -2.2% 左右,因此标定结果非常令人满意,说明了本文算法的精确性和有效性。

为进一步验证本文方法的精确性和有效性,根据华盛顿大学西雅图分校的 Furukawa^[19]提出的基于面片的多视图三维重建算法,分别对文献[11]和文献[15-18]所给的相机参数对 Valbonne church 序列图像进行三维重建,得到彩色的 3D 点云三维可视化和本文方法进行比较。图 2 和图 3 分别显示了将所得 Valbonne church 的 3D 点云的侧面和正面进行可视化的比较结果。由图 2 和图

3 的彩色点云可视化结果可见,文献[11]所得结果明显好于文献[17,18]所得结果,文献[15,16]所得结果和文献[11]所得结果差别不大。与文献[11]和文献[15-18]相比,本文方法所得结果明显有更好的效果。本文实验条件采用的奔腾 4、2.8 GHz 的处理器,大概耗时 116 s 内。其中时间耗费主要在 SIFT 特征点的提取和匹配上。

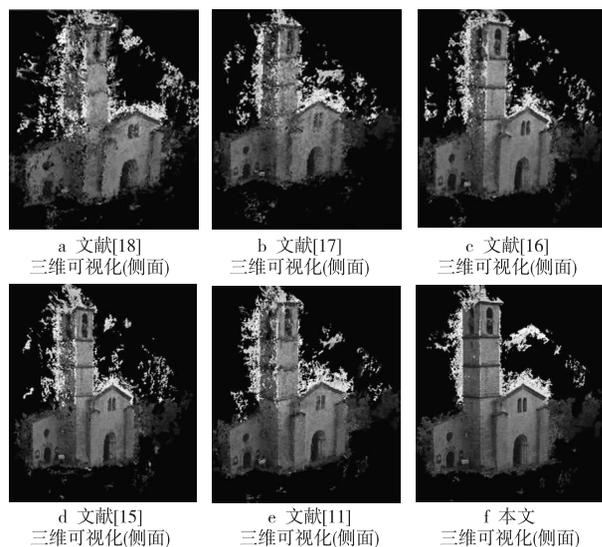


Figure 2 3D visualization comparison of the Valbonne church between the proposed method and other methods (Flank)

图 2 不同方法下的 Valbonne church 三维可视化对比(侧面)

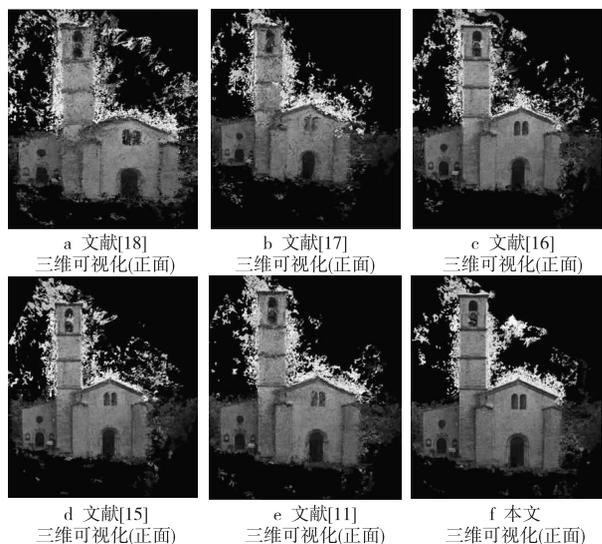


Figure 3 3D visualization of Valbonne church with the proposed method and other methods (Front)

图 3 不同方法下的 Valbonne church 三维可视化比较(正面)

相比文献[7]的方法,本文方法主要区别在于局部优化的引入,通过结合局部-全局的混合优化迭代方法得到更好的效果。图 4 显示了文献[7]方

法下的 Valbonne church 序列图像三维可视化结果,通过比较图 2f 和图 4a、图 3f 和图 4b 可知,本文采用的方法效果明显要好于文献[7]采用的方法。因为本文方法在相机多视图标定时,在依次加入相机图像时,不仅充分利用局部上的空间几何关系,同时在此过程使用全局优化,达到在过程中合并各组标定的结果从而优化相机标定的最终结果。

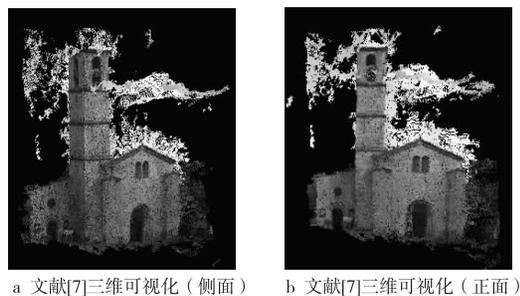


Figure 4 3D visualization of the Valbonne church of the reference [7] method

图 4 文献[7]所给方法对 Valbonne church 的三维可视化

根据 2.2 节分析可知,若只匹配每个图像邻域内的图像,则可以减少特征点匹配耗时,为了验证其有效性,同时考虑到多视图图像数量的不同对标定速度和精度的影响,本文采用 SONY DSC-W310 数码相机对实验室的一个毛绒兔玩具环绕一周拍摄了 70 张照片,图像分辨率为 2592×1944 ,焦距为 2 777 像素,图 5 为被实验对象的一幅示例图。



Figure 5 An image of toy rabbit

图 5 一幅玩具兔子示例图像

通过将照片分三种情况进行实验,分别对应 30 幅、50 幅和 70 幅图像,进行摄像机标定并进行 3D 点云可视化以验证标定结果。对于每组情况,也分别给出三种情况的 K 邻域图像匹配,对应为 4 邻域、8 邻域和全匹配。用三组不同数目的图像进行实验,来比较多视图情况下不同数目的图像所得结果,得出标定结果和图像数目的关系。分别对三个实验组采用不同邻域的匹配策略,通过实验结果来验证,当采用邻域内匹配策略时,是否能减少特征点匹配时间消耗,并可以同时保证特征点匹配结果。表 2 分别给出了不同数量图像和不同邻域匹配数的图像匹配时间比较结果。

Table 2 Estimation results with different number of images and different neighbors
表 2 不同图像数量和邻域匹配下的标定结果

	f	焦距差	f 误差 /%	u_0	v_0
4 邻域	2 649.80	127.2	4.580	1 240.97	930.724
30 幅 8 邻域	2 697.95	79.05	2.846	1 119.86	839.893
全匹配	2 708.21	68.79	2.477	1 349.95	1 012.460
4 邻域	2 659.06	117.94	4.247	1 360.87	1 020.650
50 幅 8 邻域	2 713.56	63.44	2.284	1 292.52	969.389
全匹配	2 718.73	58.27	2.098	1 220.45	915.335
4 邻域	2 693.47	83.53	3.008	1 301.41	976.057
70 幅 8 邻域	2 727.24	49.76	1.792	1 316.51	987.382
全匹配	2 760.49	16.51	0.595	1 247.90	935.921

将已知焦距和实验得到的焦距求焦距差,由表 2 可知,实验所得的焦距误差可以进一步表明本文方法的有效性和精确性。可以看出,图像数量越多,焦距误差越低,本次实验最小误差是 0.595%。

图 6 表示的是不同图像数和不同邻域匹配下所得到的特征点匹配数和时间消耗。由图 6 可以看出,随着图像数的增多,匹配点数急剧增加。若不考虑邻域匹配,特征点匹配的时间会随着图像数量的增多而大幅增加。

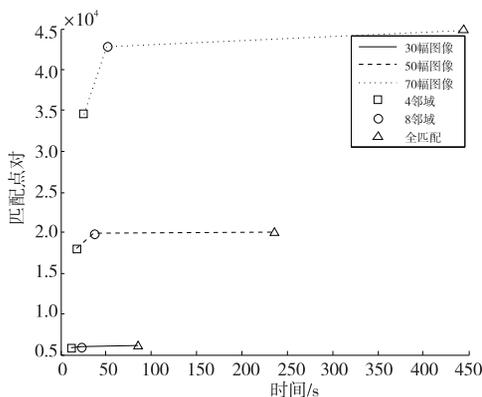


Figure 6 Time-consumption and matches with different number of images and neighbors

图 6 不同图像数和邻域匹配下的匹配数和时间消耗

从图 6 可见,采用邻域匹配策略时,当图像数量大幅增加时,特征点的匹配时间消耗增加并不明显,同时特征点匹配数目则大幅增多,这个结果是非常可观的。例如,70 幅图像全匹配时间大约是 30 幅图像全匹配时间 5 倍,当对 70 幅图像采取 8 邻域匹配时,所用时间大概为 30 幅图像全匹配时间的一半,同时标定的精度也好于 30 幅图像全匹配。从表 2 也可以看出,采用 8 邻域时的精度和全匹配时精度相差不大,匹配点对的数目也相差不

大,但匹配时间却明显少于全匹配时间,这印证了一幅图像的匹配图像几乎都在其邻域内这一假设。图 7 显示了对应的 3D 点云可视化结果。

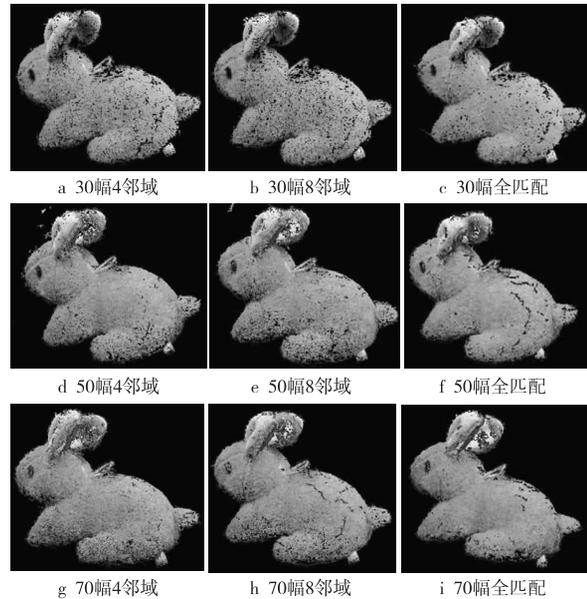


Figure 7 3D visualization of the rabbit with different number of images and different neighbors

图 7 兔子不同图像数不同邻域匹配的三维可视化结果

4 结束语

本文提出了一种基于多视图序列的摄像机自标定的方法,主要通过结合局部-全局相混合的迭代优化方法计算得到相机参数。实验表明,通过仅对互为邻域的图像进行匹配可以有效降低图像匹配的时间消耗。针对多图像匹配效率慢的问题,本文根据相似信息主要集中在相邻图像的特点,提出邻域图像特征点匹配的改进方法,实验表明对图像匹配效率提升效果显著。

综上所述,本文充分利用局部优化和全局优化的思想,采用局部-全局混合的方法求得各个相机的内外参数的系列化方法,经实验验证,不仅精度较高,而且操作简便。

参考文献:

- [1] Mi Tao, An Ping, Liu Su-xing, et al. Multi camera calibration based on multi view registration algorithm[J]. Journal of Image and Graphics, 2008, 13(10): 1922-1925. (in Chinese)
- [2] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [3] Wu Qing-shuang, Fu Zhong-liang, Meng Qing-xiang. New camera self-calibration method based on image sequences[J]. Journal of Computer Applications, 2011, 31(11): 3010-3014.

- (in Chinese)
- [4] Sun Yong-rong, Huang Bin, Yang Bo-wen, et al. Vision calibration method of spatial point position under multi-view scene determination [J]. Journal of Nanjing University of Aeronautics & Astronautics, 2015, 47(3): 343-347. (in Chinese)
- [5] Moré J J. The Levenberg-Marquardt algorithm: Implementation and theory[M]//Watso G A. Numerical Analysis. Berlin:Springer,1978;105-116.
- [6] Hu S, Qiao J, Zhang A, et al. 3D reconstruction from image sequence taken with a handheld camera[C]//Proc of International Society for Photogrammetry and Remote Sensing, 2008;559-562.
- [7] Nguyen M H, Wunsche B, Delmas P, et al. Modelling of 3d objects using unconstrained and uncalibrated images taken with a handheld camera[C]//Proc of the 7th International joint Conference on Gomputer vision, Image and Computer graphics Theory and Applications, 2013;86-101.
- [8] Dai Jia-jing. Research on the theory and algorithms of 3D reconstruction from multiple images[D]. Shanghai: Shanghai Jiao Tong University, 2012. (in Chinese)
- [9] Lowe G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-100.
- [10] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry [J]. Artificial Intelligence, 1995, 78(1): 87-119.
- [11] Jiang Ze-tao, Wu Wen-huan. An essential matrix-based camera self calibration method[J]. Journal of Image and Graphics, 2010, 15(4): 565-569. (in Chinese)
- [12] Gong Sheng-rong, Zhao Wan-jin, Liu Chun-ping. Matched points purify algorithm based on gradient of disparity constraint[J]. Journal of System Simulation, 2008, 20(S1): 407-411. (in Chinese)
- [13] Liu Xue-yi, Zhao Ya-ling. Practical method for solving fundamental matrix[J]. Science & Technology Information, 2007(7): 254-255. (in Chinese)
- [14] Hartley R I. Estimation of relative camera positions for uncalibrated cameras[C]//Proc of the 2nd European Conference on Computer Vision(ECCV'92), 1992;579-587.
- [15] Lourakis L, Deriche R. Camera self-calibration using the svd of the fundamental matrix; Technical report 3748[R]. Sophia Antipolis, France: The France National Institute for Research in Computer Science and Control, 1999.
- [16] Zeller C, Faugeras O D. Camera self-calibration from video sequences: The Kruppa equations revisited; Research Report 2793[R]. Sophia Antipolis, France; The France National Institute for Research in Computer Science and Control, 1996.
- [17] Habed A, Boufama B. Camera self-calibration from bivariate polynomials derived from Kruppa's equations[J]. Pattern Recognition, 2008, 41(8): 2484-2492.
- [18] Fusiello A, Benedetti A, Farenzena M, et al. Globally convergent autocalibration using interval analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(12): 1633-1638.
- [19] Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(8): 1362-1376.

附中文参考文献:

- [1] 宓桃, 安平, 刘苏醒, 等. 基于多视定位算法的多摄像机标定 [J]. 中国图象图形学报, 2008, 13(10): 1922-1925.
- [3] 吴庆双, 付仲良, 孟庆祥. 基于序列图像的摄像机自标定方法 [J]. 计算机应用, 2011, 31(11): 3010-3014.
- [4] 孙永荣, 黄斌, 杨博文, 等. 多视图场景确定下的空间点位置视觉标定方法 [J]. 南京航空航天大学学报, 2015, 47(3): 343-347.
- [8] 戴嘉境. 基于多幅图像的三维重建理论及算法研究 [D]. 上海: 上海交通大学, 2012.
- [11] 江泽涛, 吴文欢. 基于本质矩阵的摄像机自标定方法 [J]. 中国图象图形学报, 2010, 15(4): 565-569.
- [12] 龚声蓉, 赵万金, 刘纯平. 基于视差梯度约束的匹配点提纯算法 [J]. 系统仿真学报, 2008, 20(S1): 407-411.
- [13] 刘学义, 赵雅玲. 一种求解基础矩阵的实用方法 [J]. 科技资讯, 2007(7): 254-255.

作者简介:



唐秋虎(1991-), 男, 湖北汉川人, 硕士生, 研究方向为图形学与虚拟技术。E-mail: tangqiu-hu@qq.com

TANG Qiu-hu, born in 1991, MS candidate, his research interest includes graphics, and virtual technology.



张志毅(1974-), 男, 山西运城人, 博士, 副教授, CCF 会员 (E200033209M), 研究方向为计算机辅助设计与图形学。E-mail: 815802490@qq.com

ZHANG Zhi-yi, born in 1974, PhD, associate professor, CCF member (E200033209M), his research interests include computer aided design, and computer graphics.