

文章编号 1004-924X(2015)02-0557-09

面向场景重构的多序列间配准

杨 磊^{1,2}, 李桂菊^{1*}, 王丽荣³

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;
2. 中国科学院大学, 北京 100039; 3. 长春大学 电子信息与工程学院, 吉林 长春 130022)

摘要:针对传统的单序列扩展式场景重构方法易丢失场景信息, 数据利用率低下等问题, 提出一种面向场景重构的多序列间配准的方法。该方法采用相似变换作为配准模型, 完成对不同参考坐标系下尺度、旋转、平移相分离的初始配准; 然后通过拟合观测平面来抑制噪声点并筛选公共可见点; 最终对不同尺度、方位的序列进行配准, 配准结果可直接用于后续的重构中。实验表明: 通过相应噪声抑制, 使得各序列的系统初始重投影误差降低了 17.69% 到 46.86%, 终止重投影误差降低了 27.5% 到 71.96%。配准后从相同的 47 幅图像中可重构 10 596 个场景点, 相比传统单序列方法的 3 893 个场景点, 该方法更充分有效地利用了观测图像, 使最终拟合的场景曲面包含了更多的场景细节。

关键词:三维重构; 点云; 配准变换模型; 初始配准; 噪声抑制; 公共可见点

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.3788/OPE.20152302.0557

Registration between multiple sequences for scene reconstruction

YANG Lei^{1,2}, LI Gui-ju^{1*}, Wang Li-rong³

(1. Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China;

2. University of Chinese Academy of Sciences, Beijing 100039, China;

3. School of Electronic and Information Engineering, University of Changchun, Changchun 100022, China)

* Corresponding author, E-mail: lgjciom666@126.com

Abstract: As traditional single sequence scene reconstruction method is easy to loss scene information and has a lower data utilization, this paper proposes an approach to the registration between multiple sequences for scene reconstruction. The method uses a similarity transformational as the registration model and proposes the separating initial registration method to deal with the scale, rotation and translation respectively. Then, it fits the observed plane to suppress the noise data and select the co-visible points. Finally, the registration between multiple sequences with different scales, positions is performed and the registration results are applied in successive subsequent reconstruction. The experiment shows that the initial error has been reduced from 17.69% to 46.86%, and the terminated error reduced from 27.5% to 71.96% after suppression of noise data along the oriented orientation. Moreover, the 10 596 vertices are reconstructed from 47 frame images, more better than 3 893 vertices re-

收稿日期: 2013-11-15; 修订日期: 2013-12-23.

基金项目: 国家自然科学基金资助项目 (No. 61172111)

constructed by traditional single sequential method. The proposed method full makes use of the observation images and allows the ultimate fitting surface of the scene to have more scene details.

Key words: three-dimensional reconstruction; point cloud; transformational model of registration; initial registration; noise suppression; co-visible point

1 引言

长期以来,对场景中被观测目标的几何结构进行三维重构一直被认为是视觉计算的最终目标。作为该领域的代表性理论, Marr 的 3D 视觉理论提出视觉计算的目标是建立一个可独立于观察者的场景三维模型。这一理论最终发展为几何视觉的基础理论,其研究内容包含三类几何问题^[1]:对应性几何、运动几何、结构几何。其中:对应性几何也称特征几何,常见于特征匹配;运动几何也称摄像机几何,常见于运动姿态估计;结构几何也称场景几何,常见于场景目标的三维结构重建。结构几何问题被认为是几何视觉的最终问题,其解决方法往往基于对应性几何与运动几何。

当观测目标尺寸超出视场范围时,所成图像只包含目标全貌的一部分;当目标自身存在自遮挡或局部重复性结构时,将造成不同观测位置下重建结果出现歧义性,这一现象在大型目标三维重构过程中时常出现。此外,受成像清晰度等因素的限制,基于图像的三维重构方法多用于近景测量,但由于透视投影所造成的“近大远小”的成像效果会导致近景测量中目标成像的尺度相对较大,从而加剧了自遮挡及重复性结构的影响^[2]。因此,较大型目标全貌的细致重建,以及如何避免自遮挡及重复性结构带来的重建歧义目前已成为三维重构研究的重点。

长期实践证实,基于多视图几何^[3]的大型目标重构是可靠的,而这类方法中最成功的是“由运动恢复结构”^[4-7]。在此类方法中,目标的三维场景模型与其二维图像特征完全依靠相对目标与观测者的相对运动姿态进行关联,因此,无论是初始的结构重建,还是后期的模型优化,都需反复估计相对运动姿态。较为常见的方法是根据运动顺序建立单一的观测图像序列,以进行序列式场景重构^[4,7]。由于单一序列只存在单一的参考坐标系,因此这类方法相对运动姿态的估计精度高;其主要缺点是重建特征的数量随观测序列的增加而

增长放缓,甚至减少,对于利用率较差,重构结果较为稀疏的观测图像,极易丢失细节。

若对同一目标采用多个观测序列进行重构,则可大大提升对观测图像的利用率,但由于不同观测序列基于不同的参考坐标系,重构结果之间的相对位置、方向、尺度均不同,无法进行直接叠加。针对这个问题,本文提出了一种可以对多个观测序列的观测姿态进行配准的方法,以更充分地利用观测图像。该方法可以实现多重序列重构,得到更密集的重构效果。

2 多序列场景重构原理

典型的单序列式重构的基本步骤是:(1)观测者在不同位置提取场景中感兴趣目标的图像特征,并建立不同图像间的特征匹配,记为 $\{u_i \leftrightarrow u_j\}$; (2)根据图像间的特征匹配关系选定参考坐标系原点 q 建立观测序列,并估计各观测点处摄像机与目标之间的相对运动姿态,记为 \mathbf{P} ; (3)由图像中的匹配特征及摄像机姿态重建目标的三维结构模型,记为 \mathbf{X} ,通过初始姿态 $\{\mathbf{P}_i, \mathbf{P}_j\}$ 、 \mathbf{X} ,以及匹配集 $\{u_i \leftrightarrow u_{j+k}\}$ 可预测 $j+k$ 处的运动姿态,以对结构进行序列式扩展。

上述场景的三维重构过程也可建模为: $\mathbf{X}_n = \mathbf{P}_m^+ \mathbf{u}_{mn}$; $1 \leq m \leq M, 1 \leq n \leq N$ 。其中, \mathbf{P}^+ 是姿态投影矩阵 \mathbf{P} 的伪逆,不唯一(应求最优解), m 和 n 标记为第 m 个观测位置处的第 n 个观测点。理想状态下,序列中的 3D 点到 2D 图像平面的重投影可与其相应的 2D 特征点重合,即 $\mathbf{u}_{mn} = \mathbf{P}_m \mathbf{X}_n$,因此,当重建结果为最优时,3D 点对 2D 特征的重投影误差应最小^[1],即式(1)成立:

$$\{\hat{\mathbf{P}}, \hat{\mathbf{X}}\} = \underset{\{\mathbf{P}, \mathbf{X}\}}{\operatorname{argmin}} \sum_m \sum_n \|\mathbf{u}_{mn} - \mathbf{P}_m \mathbf{X}_n\|. \quad (1)$$

式(1)中,观测姿态的投影矩阵 \mathbf{P} 可由摄像机的内参 \mathbf{K} 与外参 \mathbf{E} 联合表示。在本文中,摄像机采用针孔模型,并已进行标定。因此,不同观测点处投影矩阵的差别由尺度 λ 及外参 \mathbf{E} 决定,其中 \mathbf{E} 由旋转矩阵 \mathbf{R} 和平移向量 \mathbf{t} 组成,即: $\mathbf{E} =$

$[\mathbf{t}]_x \mathbf{R}$ 。通常,对于单一序列,可规定初始重建的图像对中基准图像(本文规定观测序列的基准图像为初始图像对中的左侧图像)的初始投影矩阵为 $\mathbf{P}_0 = \mathbf{K}[\mathbf{I}, \mathbf{0}]$;而对于多个序列,由于其基准图像观测位置不同,其配准过程可归结为先选定一个序列作为基准,再将其余序列的姿态及结构对准到该序列,即: $\mathbf{P}_0 = \mathbf{P}_1 \mathbf{M}_1 = \mathbf{P}_2 \mathbf{M}_2 = \dots = \mathbf{P}_k \mathbf{M}_k$, $\mathbf{X}_0 = \mathbf{T}_1 \mathbf{X}_1 = \mathbf{T}_2 \mathbf{X}_2 = \dots = \mathbf{T}_n \mathbf{X}_n$,理想状态下, \mathbf{T} 为 \mathbf{M} 的伪逆。然而,多序列间存在大量的非公共点,因此 $\mathbf{T}_k \neq \mathbf{M}_k^+$ 。当多序列重构的结果最优时,序列间的公共可见点处可满足 3D 对 2D 的重投影误差最小,即有式(2):

$$\{\hat{\lambda}, \hat{\mathbf{T}}, \hat{\mathbf{M}}, \hat{\mathbf{P}}, \hat{\mathbf{X}}\} = \underset{(\lambda, \mathbf{T}, \mathbf{M}, \mathbf{P}, \mathbf{X})}{\operatorname{argmin}} \sum_k \sum_m \sum_n \| \mathbf{u}_{kmi} - \lambda_k (\mathbf{P}_m \mathbf{M}_k) (\mathbf{T}_k \mathbf{X}_n) \| . \quad (2)$$

上式中,引入了表征不同序列间尺度差异的因子,且配准变换矩阵 \mathbf{M} 和 \mathbf{T} 难以同时估计,因此应假设 $\mathbf{T}^+ = \mathbf{M}$,然后通过最小化式(2),最终求出最优解。

这一问题有时也被称为“最优观测路径”或“公共可见性”问题,是这一领域近年的研究热点^[8-10]。解决这一问题的典型方法是假定序列在公共节点处的尺度相近,以刚体模型进行初始旋转配准后,通过式(3)最优化尺度和平移配准^[11-14]。这类方法的主要缺点是:(1)当两序列在公共节点处的观测尺度存在明显差异时,其配准结果易产生较大的重投影误差,甚至病态解;(2)受噪声点影响较大,且配准误差容易向后扩散,导致后续配准精度较先前明显下降,当各序列自身包含较多重构点数时,配准后容易出现模糊效果甚至错误。

针对上述问题,本文主要进行了两点改进:(1)对于序列间存在明显的尺度差异问题,采用相似变换模型来减小序列间变换模型的配准误差;(2)在序列的观测过程中引入有针对性的噪声抑制以及公共可见点的筛选方法来提高重构及配准精度。

3 多序列间配准

序列间配准即对两序列进行结构上的对准,其可分为初始阶段和优化阶段。在优化阶段,由变换模型决定其误差函数形式,而且受初值的影

响较大,因此,配准变换模型及相应的初值估计显得尤为重要。此外,需配合相应的噪声抑制及公共可见点筛选方法,以提高配准精度。按上述顺序,本节首先讨论了配准变换的建模;然后研究了相应的噪声抑制以及序列间公共可见点的选取方法;最后探讨了去噪后模型下的初始配准方法。

3.1 配准变换模型

不同观测序列间的差异体现在以下 3 个方面:尺度 λ 、旋转 \mathbf{R} 、平移 \mathbf{t} ,当序列 i 与 j 间的观测尺度 λ 相近时,变换模型 \mathbf{T} 可表示为刚体变换,即: $\mathbf{T}_{ij}^E = [\mathbf{R}_{ij}, \mathbf{t}_{ij}]$ 。 \mathbf{T}^E 包含 6 自由度,分别是绕 x, y, z 轴的旋转分量以及沿 x, y, z 轴的平移分量。其中,旋转矩阵 \mathbf{R} 为正交矩阵,其尺度因子可视作公因子约去。然而,当序列 i 与 j 的参考坐标原点差别较大时, i 与 j 之间将具有明显的尺度差异,按照投影变换分层的方式^[10],配准重建过程将由刚体变换模型转变为相似变换模型。尺度因子 λ 对旋转矩阵及平移向量的尺度影响相同,可作为公共系数提出, \mathbf{R}_{ij}' 的各分量仅在其矢量方向上具有物理意义,故 \mathbf{R}_{ij}' 为单位正交矩阵,其配准模型满足:

$$\mathbf{T}^E = [\lambda \mathbf{R}_{ij}', \mathbf{t}_{ij}']$$

$$\text{subject to: } \mathbf{R}_{ij}' = \mathbf{U}_{ij} \operatorname{diag}(1, 1, 1) \mathbf{V}_{ij}^T. \quad (3)$$

上式表明:对 \mathbf{R}_{ij}' 的分解及规范化过程可由奇异值分解实现^[2]。对尺度、旋转、平移分别进行初始化配准,然后再根据配准后公共可见点处的重投影误差进行联合优化。对于各序列间的公共可见点,成像过程应具有统一形式,可描述为:

$$\begin{bmatrix} f_x & 0 & p_x & 0 \\ 0 & f_y & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda \mathbf{R}' & \mathbf{t}' \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{X}^w = \begin{bmatrix} f_x & 0 & p_x & 0 \\ 0 & f_y & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{X}^h = \mathbf{U}^1, \quad (4)$$

式(4)得到任意两序列 i 和 j 间的公共可见点的重投影过程,也即 $\mathbf{P}_i \mathbf{X} = \mathbf{u}_i = \mathbf{u}_j = \mathbf{P}_j \mathbf{X}$,则对于序列 i 和 j ,以 i 为参考序列进行配准,即 $\mathbf{X}_i = \mathbf{T}_{ij}^E \mathbf{X}_j$,可按照尺度、旋转、平移的分离顺序进行初始配准。

3.2 噪声抑制

空间 3D 点在成为曲面或体的顶点前,不具备拓扑属性(即连接关系),只具备坐标属性(即空间位置),噪声点即是空间位置错误的点。也即忽

略拓扑关系时,坐标值明显偏离场景表面的点。由式(4)可知,空间中 3D 顶点 $\mathbf{X}=(x, y, z, 1)^T$ 向图像中 2D 特征点 $\mathbf{u}=(u, v, 1)^T$ 投影的过程可通过两次变换完成。第一次变换为 $\mathbf{X} \rightarrow \mathbf{X}_h$, 通过外参,其中 $\{\mathbf{X}^h | \mathbf{X}_k^h=(x^h/z^h, y^h/z^h, 1), z^h>0\}$, 将 \mathbf{X} 从世界坐标系向以摄像机光心为原点的摄像机欧式坐标系投影,该过程可逆,其可表示为 $P^3 \Leftrightarrow P^3$ (等价于 $R^3 \Leftrightarrow R^3$); 第二次变换为 $\mathbf{X}_h \rightarrow \mathbf{u}_h$, 是通过内参将 \mathbf{X} 从摄像机欧式坐标系向图像坐标系投影的过程,不可逆,可表示为 $P^3 \Rightarrow P^2$ (等价于 $R^3 \Rightarrow R^2$)。

序列间配准属于在 P^3 空间内变化,即 $P^3 \Leftrightarrow P^3$, 与第一次变换涉及到的外参模型紧密关联,故应在 P^3 (或 R^3) 内滤除噪声点。2D 图像保留了目标的 x 和 y 轴的坐标信息,丢失了 z 轴坐标信息,因此图像三维重构的关键在于重构 3D 点的 z 坐标值。由于 z 轴正方向总是指向观测点位置,理想的摄像机投影平面应垂直于观测方向,观测顶点组成的空间曲面应分布在该理想投影平面两侧,曲面法向量与 z 轴方向夹角应小于一定值(本文为 45°)。对点集 \mathbf{X} 中心化,即 $\mathbf{X}'=\mathbf{X}-\bar{\mathbf{X}}$, 从 \mathbf{X}' 中随机抽取 3 个观测点,组成点集 $\mathbf{X}_s=\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$, 由 $\mathbf{X}_s^T \mathbf{\Pi}=0$ 解得观测平面 $\mathbf{\Pi}$ 。该观测平面的法向量 $\mathbf{n}_\Pi=(\mathbf{X}_1-\mathbf{X}_2) \times (\mathbf{X}_1-\mathbf{X}_3)$, 观测点处 z 轴方向为 \mathbf{n}_z , 则法向量与观测方向的观测夹角可定义为 $\theta_k(\mathbf{n}_\Pi, \mathbf{n}_z)=\arccos \|\mathbf{n}_\Pi-\mathbf{n}_z\|$ 。

实际观测中,若观测夹角过大,则说明至少有一个观测顶点的深度值较其他各顶点有较大差异。而对于单一视点处的局部观测面而言,可能是由观测误差造成。计算其余顶点到该观测面的距离之和,则最优观测面可最小化为点面距离的均方和,该最小化过程可通过反复抽样迭代完成,即:

$$\hat{\mathbf{\Pi}}=\operatorname{argmin}_{\mathbf{\Pi}_k} \sum_k d(\mathbf{X}, \mathbf{\Pi}_k)$$

$$\text{subject to: } \theta_k(\mathbf{n}_{\mathbf{\Pi}_k}, \mathbf{n}_z)=\arccos \|\mathbf{n}_{\mathbf{\Pi}_k}-\mathbf{n}_z\|<\theta_w, \quad (5)$$

式中:夹角阈值 θ_w 在本文实验中取 45° , $d(\mathbf{X}, \mathbf{\Pi}_k)$ 为欧式范数下的点到平面的距离。求解中心化点集 \mathbf{X}' 到最优观测平面 $\hat{\mathbf{\Pi}}$ 的距离平均值及标准差,根据“3 倍标准差法则”将距离均值的 3 个标准差的顶点作为噪声点。这一去噪过程需在观测阶段的序列式扩充重构中进行,以降低系统误差,

并可明显降低初始平移配准的难度。

3.3 公共可见点选取

各序列在公共节点处存在可共同观测的 3D 顶点,也即公共可见点。由于初始尺度及平移配准对其依赖程度较高,因此应进一步筛选去噪后的公共可见点。从式(4)、式(5)可知,无论外参变换模型的具体形式如何,都有 $\mathbf{X}^h=\mathbf{T}\mathbf{X}$, 且点集 \mathbf{X}^h 处于摄像机投影平面上,点集在该平面 x 与 y 方向的最优分布范围可通过对 $\mathbf{T}\mathbf{X}$ 进行奇异值分解估计。为减小数据规模,采用外参 $\mathbf{E}=\lambda[\mathbf{t}]_r \mathbf{R}$ 代替 \mathbf{T} , 点集 \mathbf{X} 采用欧式坐标并中心化。若 \mathbf{X} 共包含 K 个点,点集中心为 $\bar{\mathbf{X}}_{3 \times 1}$, 则该过程可表示为:

$$\mathbf{U}_{3 \times 3} \mathbf{D}_{3 \times 3} \mathbf{V}_{3 \times K}=\mathbf{E}_{3 \times 3}(\mathbf{X}_{3 \times K}-\bar{\mathbf{X}}_{3 \times 1}), \quad (6)$$

上式中,单位正交矩阵 $\mathbf{U}=(\mathbf{u}^x, \mathbf{u}^y, \mathbf{u}^z)$ 的 3 个分量指出空间坐标系三坐标轴方向,对角阵 $\mathbf{D}=\operatorname{diag}(\sigma^x, \sigma^y, \sigma^z)$, 3 个对角线因子表示三轴的坐标尺度,当且仅当公共顶点 \mathbf{X}_k 的坐标满足 $\|\mathbf{X}_k\|^2<\|\sigma^x \mathbf{u}^x\|^2+\|\sigma^y \mathbf{u}^y\|^2+\|\sigma^z \mathbf{u}^z\|^2$ 时,才被选为序列间公共可见点,参与序列间配准。

3.4 初始配准方法

3.4.1 尺度配准

首先进行尺度配准,令针孔相机的主点 (p_x, p_y) 位于图像中心,将 \mathbf{U}^l 改为非齐次形式 $\mathbf{u}=\{u, v\}=\{\frac{f_x x^h}{z^h}+p_x, \frac{f_y y^h}{z^h}+p_y\}$, 则点集 $\{\frac{x^h}{z^h}, \frac{y^h}{z^h}, 1\}$, $z^h \neq 0$, 将位于摄像机单应性平面上。为简化实现,可令 $f=(f_x+f_y)/2$ 。该针孔模型中,场景的 3D 顶点全部位于摄像机前方,其深度坐标值为正,对于序列 i 与序列 j 间的尺度估计可由满足深度测试的 K 个公共可见点的平均深度 $\{\bar{z}_i, \bar{z}_j\}$ 以及两序列各自的焦距 $\{f_i, f_j\}$ 进行放缩,指定序列 i 为参考序列:

$$\hat{\mathbf{X}}_{jk}=\operatorname{argmin}_{\mathbf{X}_j} \sum_{k=1}^K \|\bar{z}_{jk} f_j \mathbf{X}_{jk}-\bar{z}_{ik} f_i \mathbf{X}_{jk}\|$$

$$\text{subject to: } \mathbf{X}_{ik} \in \{(x_k, y_k, z_k) | u_{ik}=P_i X_i, z_k>0\}$$

$$\mathbf{X}_{jk} \in \{(x_k, y_k, z_k) | u_{jk}=P_j X_j, z_k>0\}. \quad (7)$$

若存在满足式(7)的点 $\hat{\mathbf{X}}_{jk}$, 则可对序列 j 的所有 3D 顶点进行尺度放缩,即 $\mathbf{X}_j^s=\frac{\bar{z}_{jk} f_j}{z_{ik} f_i} \mathbf{X}_j$ 。对于放缩后的顶点 $\{\mathbf{X}_j^s\}$, 已初步消除式(3)模型中的尺度差异 λ_{ij} , 此时可进行初始旋转配准。

3.4.2 旋转配准

式(3)中旋转矩阵应具备单位正交性,其各分

量尺度并无物理意义,因此其配准过程存在以下两点特殊性:(1)旋转配准的初始配准可直接由各序列自身在公共节点处的旋转姿态计算;(2)由于其不直接基于公共可见点进行估计,若噪声影响较小时,可在后期的全局优化中将旋转矩阵固定。依然指定序列 i 为参考序列,序列 j 为待配准序列,在公共节点 \check{x}_j 处,序列 i 的旋转矩阵为 \mathbf{R}_i ,序列 j 的旋转矩阵为 \mathbf{R}_j ,则 i 和 j 的相对旋转矩阵 \mathbf{R}_{ij} 应满足:

$$\hat{\mathbf{R}}_{ij} = \underset{\mathbf{R}_{ij}}{\operatorname{argmin}} (\| \mathbf{R}_j \mathbf{R}_i^T - \mathbf{R}_{ij} \|)$$

$$\text{subject to: } \det(\hat{\mathbf{R}}_{ij}) = 1. \quad (8)$$

可通过将旋转矩阵改为四元数分量形式求解上式,即

$$\dot{\mathbf{r}}_{ij} = r_{ij}^0 + i \cdot r_{ij}^x + j \cdot r_{ij}^y + k \cdot r_{ij}^z,$$

其中:实部表示尺度,3个虚部表示围绕 x , y , z 方向三轴的独立旋转分量,由于最优解 $\det(\hat{\mathbf{R}}_{ij}) = 1$,因此式(8)的解应满足 $\| \dot{\mathbf{r}}_{ij} \| = 1$ 。由于各序列自身的旋转矩阵为正交阵,则式(8)可表示为四元数形式下的求解线性最小二乘系统:

$$[\dot{\mathbf{r}}_{ij}]_x \dot{\mathbf{r}}_i = \dot{\mathbf{r}}_j$$

$$\text{subject to: } (r_{ij}^0)^2 + (r_{ij}^x)^2 + (r_{ij}^y)^2 + (r_{ij}^z)^2 = 1. \quad (9)$$

为求外积,采用四元数向量 \mathbf{r}_{ij} 的反对称阵形式 $[\dot{\mathbf{r}}_{ij}]_x$, $\operatorname{rank}([\dot{\mathbf{r}}_{ij}]_x) = 4$ 。对顶点坐标的变换与对参考坐标系的变换互为逆变换。序列的旋转配准为 $\mathbf{X}_j^r = \mathbf{R}_{ij}^{-1} \mathbf{X}_j^s$,其中 \mathbf{X}_j^s 为尺度配准后的序列顶点。在实际的计算过程中存在少量误差,使得 $|\det(\dot{\mathbf{r}}_{ij}) - 1| = \delta$, δ 为一个较小的数值。如前文所述,旋转配准并不直接依赖于公共可见点,但当公共可见点中的噪声点所占比例较小时, δ (可取 1×10^{-5}) 的影响可忽略不计,因此在各序列独立观测阶段进行有针对性的噪声抑制可确保旋转序列间配准的成功。

3.4.3 平移配准

根据式(3)所建模型,相对平移向量 $\mathbf{t}_{ij}' = \lambda_{ij} \mathbf{t}_{ij}$,其尺度因子由序列尺度差异决定,且 $\mathbf{t}_{ij} = -\mathbf{R}_{ij} \mathbf{c}_{ij}$, \mathbf{c}_{ij} 为公共节点处以序列 i 为参考的序列 j 的摄像机相对坐标,因此初始的平移配准需在初始的尺度与旋转配准完成后才能进行。不同于旋转配准,若序列间已进行尺度与旋转配准,则序列间的初始平移配准可通过公共可见点进行。与文献[12-14]中将平移配准直接并入后期全局优化

的传统方法不同,本文采取的结构与运动相分离,初值估计与后期优化相分离的配准方法可明显简化实现过程。

在公共节点处,序列 i 与 j 的可观测顶点集分别为 $\{\mathbf{X}_i | (x_i, y_i, z_i)\}$, $\{\mathbf{X}_j | (x_j, y_j, z_j)\}$,公共可见点集为 $\mathbf{X}_{ij}^{\text{com}} \subseteq \{\mathbf{X}_i \cap \mathbf{X}_j\}$, $\mathbf{X}_{ij}^{\text{com}} \neq \emptyset$ 。序列 i 在公共可见点集中的点集为 $\check{\mathbf{X}}_i = \{x_i \cap \mathbf{X}_{ij}^{\text{com}}\}$,相应地, $\check{\mathbf{X}}_j = \{x_j \cap \mathbf{X}_{ij}^{\text{com}}\}$ 。通过分别计算点集之间的相对平移 $\mathbf{t}_{ij} = \check{\mathbf{X}}_i - \check{\mathbf{X}}_j$ 可进行平移配准,该平移可通过点集中心计算,可表示为:

$$\mathbf{t}_{ij} = \frac{1}{K} \sum_{k=1}^K (\check{\mathbf{X}}_{ik} - \check{\mathbf{X}}_{jk})$$

$$\text{subject to: } \check{\mathbf{X}}_{ik} \neq \emptyset, \check{\mathbf{X}}_{jk} \neq \emptyset. \quad (10)$$

序列的平移配准即: $\mathbf{X}_j^{\text{st}} = \mathbf{X}_j^r + \mathbf{t}_{ij}$,此时即完成了序列间的初值配准过程。值得注意的是:公共可见点的点集中心易受噪声点影响,因此后期仍需对平移进行优化。但是,结合特定的噪声抑制方法并筛选公共可见点后,可明显改善初始的尺度及平移配准效果,极大减轻了后期优化的负担。

4 实验及结果

本次实验对象为观测场景中无法由单一视角观察全貌的大型目标。由4个初始方位进行4序列扩展式观测与重构后,最终进行序列间配准及重构。

4.1 实验步骤

首先对于同一场景的多个不同观测序列进行相互独立的单一序列式重构;然后,指定参考序列,进行多序列间配准;最后,基于配准结果进行拟合场景曲面,具体步骤为:

(1)首先对各序列进行经典的单序列扩展式重构,主要步骤见第二节。需要说明的是:本文采用尺度不变特征变换算法(SIFT)进行特征提取。另外,根据初始 $\{\mathbf{P}_i, \mathbf{P}_j\}$ 、 \mathbf{X} 以及匹配集 $\{u_i \leftrightarrow u_{j+k}\}$,可建立 3D \rightarrow 2D 的对应,通过透视 N 点法(Perspective N Points),可对各序列进行扩展式重构,本文采用了 EPNP(Efficient Perspective N -Point)^[15];

(2)各序列中,每当扩展一幅新的观测图像(即得到新的 \mathbf{P} 及 \mathbf{X})时,需首先利用式(5)结合常规的重投影误差抑制方法去除噪声点,然后对运动 \mathbf{P} 及结构 \mathbf{X} 进行全局优化,本文采用光束平差法;

(3)当各序列扩展均完成后,根据需要(通常由所希望的观测位置决定)指定参考序列,在各序列间的公共节点处,根据式(6)求得判据并筛选公共可见点,选择公共可见点最多的节点作为配准节点。根据式(7)、(8)、(9)、(10)进行序列间的初始配准,再对全部运动 \mathbf{P} 及结构 \mathbf{X} 进行全局优化,完成多序列配准下的场景重构。

执行步骤(1)时,可得到单一序列的重构结果,以及包含的 3D 顶点数,可在实验的最终步骤中用于对比分析观测数据的利用效率。在执行步骤(2)时,可通过采用本文去噪方法前后的结果直观检验方法性能,也可通过重投影误差量化地检验噪声抑制效果。在执行步骤(3)时,可通过观察配准后的结构点直观检验配准效果,也可通过对比序列在公共可见点处系统的重投影误差量化地检验初始配准误差。最终对比步骤(1)的结果,可知数据的利用率。

4.2 实验结果

序列 1 为全图像集上单序列扩展式重构,序列 2 至 5 为分段式序列,用于多序列配准式重构。仅采用常规去噪方法与本文去噪方法的系统重投影误差对比结果如图 1 所示。当序列 1 仅采用常规的重投影噪声抑制时,初始误差为 48 409.222 905,终止误差为 32 391.880 55;当结合本文式(9)去噪时,初始误差为 30 273.256 862,终止误差为 18 977.437 456。由此可知,序列 1 采用本文方法去噪后,初始误差已低于常规的去噪并全局优化后的误差。可见,对于较长的观测序列,若存在大量的错误投影点,无法通过重投影检验时,本文去噪方法将尤为有效。各序列误差降幅如下:序列 1 初始误差降幅为 37.46%,终止误差降幅为 45.07%;序列 2 初始误差降幅为 28.97%,终止误差降幅为 33.06%;序列 3 初始误差降幅为 30.39%,终止误差降幅为 27.5%;序列 4 初始误差降幅为 17.69%,终止误差降幅为 36.54%;序列 5 初始误差降幅为 46.86%,终止误差降幅为 71.96%。

图 2 为两种噪声抑制方法的重构结果。由图 2 可知,使用 2 种方法剔除噪声的同时也将损失部分正确的场景点,从而使重构细节稀疏,这可通过多序列配准使其“密集化”。采用本文去噪方法并完成步骤(4),可得到噪声点较少的 4 组重构序列,其中:4 组为配准且未去噪的点集序列结果,

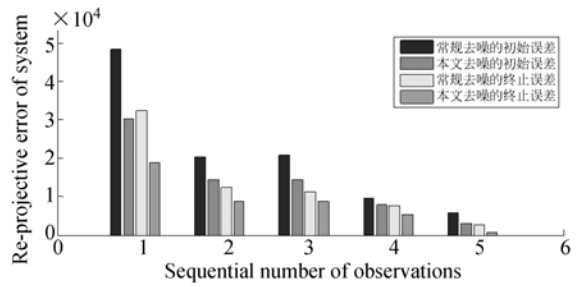
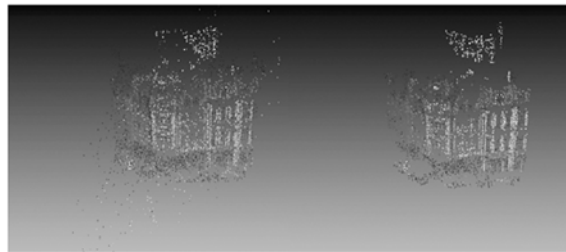


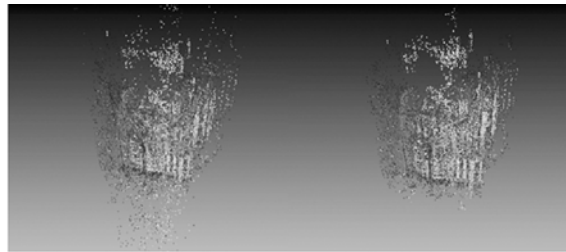
图 1 常规去噪方法与本文去噪方法的系统重投影误差对比

Fig. 1 Comparison of system reprojective errors between conventional filtering methods and proposed method



(a)全图像的单序列式重构

(a)Single sequence reconstruction for full image



(b)序列 2~序列 5 未经配准直接叠加结果

(b)Directly superposition from the 2nd sequence to 5th sequence

图 2 两种噪声抑制方法的重构结果对比

Fig. 2 Comparison of reconstruction results of two methods

见图 2(b)左图,4 组点集序列经去噪后的未配准结果见图 2(b)右图。而后可任意指定其中一个序列为基准序列进行序列间配准,以下指定序列 4 为参考序列,并执行步骤(5)。4 个序列配准前的结果可见图 2(b)右图,采用本文方法配准后的结果如图 4 右图。需指出:本次实验中,各配准步骤将统一采用本文的初始尺度,这是因为各待配准序列间存在明显的尺度差异,若使用传统的初始配准方法将产生较大误差,使对比失去意义。

基于这一原因,步骤(5)将对采用本文方法前后结果。配准路径为:序列 2->序列 3,序列(2+3)->序列 4,序列 5->序列 4。3 次初始配准的对

比结果如图 3(彩图见期刊电子版)所示。统一规定当前参考序列(基准序列)点集为蓝色,待配准序列的点集为黄色,即黄色点集向蓝色点集配准。

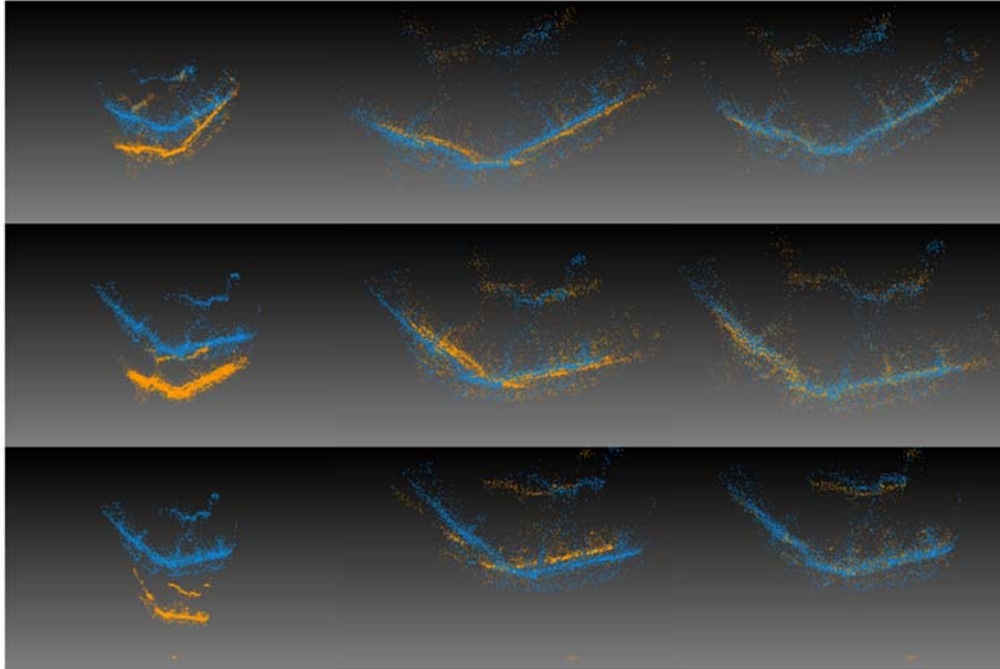


图 3 3 次初始配准序列的结果(左侧为初始未配准状态;中间为未采用本文方法结果;右侧为采用本文方法结果)

Fig. 3 Registration results of 3 times of initial registration sequences(The left sides are initial states without registration,ones the middles of the figures are registration results before using the proposed method and the right ones are registration results after using the proposed method)

由图 3 可知,每行中间的一列由于使用了相似变换作为配准模型,故其尺度配准较为准确,但未采用本文的初始旋转和平移配准方法,其配准结果存在少量误差,如第二行中可见平移误差,第三行中可见旋转误差。完成序列配准后,将两序列的公共可见点向公共节点处重投影,并将其值作为量化配准误差的依据,3 次配准误差如表 1 所示。

较高,例如 1+2 向 3 配准时。因此,需通过相似变换模型进行全局尺度优化,在全局优化前需对初始配准结果进行去噪。全局优化后,两种多序列配准的重构结果如图 4 所示。

表 1 序列间初始配准的公共节点处重投影误差

Tab.1 Reprojective errors on common nodes of initial registration between sequences

	1 向 2 配准	1+2 向 3 配准	4 向 3 配准
未采用本文初始配准	38.907 8	54.568	48.884 9
采用本文初始配准	19.570 5	49.982 3	15.489 2

由表 1 可见:当尺度配准不佳时,即便旋转和平移配准较精确,其节点处的重投影误差也依然

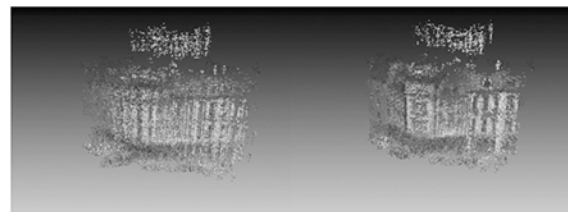


图 4 采用本文的旋转及平移配准前后的结果

Fig. 4 Registration results before/after using the proposed methods

由图 4 可见,左侧虽采用了本文方法进行初始尺度配准,但由于后续优化中仍采用传统刚体模型,因此无法进一步优化尺度,最终严重影响了旋转和平移的配准精度,导致重构点云较为散乱,产生视觉模糊。相比之下,本文的配准结果经优

化后模糊感较为微弱,也即配准偏移较小,因此对场景细节的重构效果较好。重构的点云所包含的点数决定了对场景细节的再现程度,各序列(包含配准序列)重构所包含的顶点数及观测图像数对比结果如表 2 所示。

表 2 各序列重构的点云点数及观测图像数

Tab. 2 The number of points of reconstructed point clouds and observed images in each sequence

	序列 1	序列 2	序列 3	序列 4	序列 5	配准序列
重构顶点数	3 893	3 477	3 508	3 640	3 083	10 596
观测图像数	47	13	13	14	11	47

由表 2 可见:当观测图像超过一定数量后,传统单一序列重构的点数不再有明显增长。这是由于为了抑制噪声、保证系统稳定性,需反复优化去噪,新点加入的同时,大量旧点被舍弃。经过二次去噪、全局配准优化等步骤后,配准序列的点云所包含点数并不等于各子序列点数的和,但仍远多于典型的单序列扩展式重构。对比图 2(a)右图(单序

列重建)和图 4 右图(多序列配准后重建结果)可知,在观测图像数量相同时,多序列配准对于场景的重构更为细致,对图像数据的利用率更高。

5 结 论

本文对经典的单序列扩展式重构进行了多序列式扩展,结合相应的噪声抑制及公共点筛选方法,通过合适的配准变换模型实现较准确的初始配准,从而完成了多序列间配准,并达到多序列场景重构的目的。由表 2 可知,当分段子序列总数为 4,各子序列包含图像数为 11 到 14 时,在相同数量的观测数据下,传统方法下的序列 1 的重构顶点数与本文方法的配准序列的重构顶点数分别为 3 893 和 10 596,重构的 3D 点数是传统方法的 2.72 倍,具体效果如图 2(a)右图和图 4 右图所示。在试验场景下,本文方法以更高的数据利用率重构了更多场景的细节。若能进一步优化配准结果或改进配准模型,则可进一步改善配准效果。

参考文献:

- [1] HARTLRY R, ZISSERMAN A. *Multiple View Geometry in Computer Vision* [M]. Cambridge: Cambridge University Press, 2003: 151-360.
- [2] 李畅, 刘亚文, 胡敏, 等. 面向街景立面三维重建的近景影像直线匹配方法研究[J]. 武汉大学学报: 信息科学版, 2010, 35(12): 1461-1465.
LI CH, LIU Y W, HU M, *et al.*. Study on street elevation 3D reconstruction oriented straight line matching for close-range images[J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(12): 1461-1465.
- [3] 冯萍, 魏振忠. 光笔式大视场三维视觉测量系统[J]. 光学精密工程, 2013, 21(9): 2217-2224.
FENG P, WEI ZH ZH. Light probe based large FOV 3D vision measurement system [J]. *Opt. Precision Eng.*, 2013, 21(9): 2217-2224.
- [4] STURM P F. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 1997: 1100-1105.
- [5] SNAVELY N, SEITZ S M, SZELISKI R. Modeling the world from internet photo collections[J]. *International Journal of Computer Vision*, 2008, 80(2): 189-210.
- [6] 黄凤山, 刘恩福, 方忆湘, 等. 基于智能三坐标测量机的零件位姿单目立体视觉识别[J]. 光学精密工程, 2013, 21(5): 1326-1332.
HUANG F SH, LIU E F, FANG Y X, *et al.*. Single camera stereo vision recognition for Parts' pose based on intelligent three coordinate measuring machine [J]. *Opt. Precision Eng.*, 2013, 21(5): 1326-1332.
- [7] 陈庭旺, 王庆. 面向大规模场景的三维线段检测算法[J]. 计算机辅助设计与图形学学报, 2011, 23(5): 790-796.
CHEN T W, W Q. A 3D line segment detection algorithm for large-scale scenes [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2011, 23(5): 790-796.
- [8] GOESELE M, SNAVELY N, CURLESS B, *et al.*. Multi-view stereo for community photo collections [C]. *IEEE 11th International Conference on*

- Computer Vision, ICCV 2007, 2007: 1-8.*
- [9] CORNELIS N, LEIBE B, CORNELIS K, *et al.*. 3D urban scene modeling integrating recognition and reconstruction[J]. *International Journal of Computer Vision*, 2008, 78(2-3): 121-141.
- [10] NEWCOMBE R A, DAVISON A J. Live dense reconstruction with a single moving camera[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2010, 2010: 1498-1505.*
- [11] PULLI K. Multi-view registration for large data sets [C]. *Proceedings of Second International Conference on 3-D Digital Imaging and Modeling*, 1999: 160-168.
- [12] KAHL F, HARTLEY R. Multiple-view geometry under the L1-Norm[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(9): 1603-1617.
- [13] 陈辉, 龙爱群, 彭玉华. 由未标定手持相机拍摄的图片构造全景图[J]. *计算机学报*, 2009, 32(2): 328-335.
- CHEN H, LONG A Q, PENG Y H. Building panoramas from photographs taken with an uncalibrated hand-held camera [J]. *Chinese Journal of Computers*, 2009, 32(2): 328-335.
- [14] HARTLEY R, KAHL F, OLSSON C, *et al.*. Verifying global minima for L2 minimization problems in multiple view geometry [J]. *International Journal of Computer Vision*, 2013, 101(2): 288-304.
- [15] LEPETIT V, MORENO F, FUA P. EPNP: An accurate $o(n)$ solution to the pnp problem[J]. *International Journal of Computer Vision*, 2009, 81(2): 155-166.

作者简介:



杨磊(1986—),男,湖北武汉人,博士研究生,2009年、2012年于吉林建筑大学分别获得学士、硕士学位,主要从事视觉三维重构方面的研究。E-mail: nemoleiyang@gmail.com

导师简介:



李桂菊(1964—),女,吉林吉林人,硕士,研究员,1985年于吉林大学获得学士学位,1988年于中国科学院长春光学精密机械与物理研究所获得硕士学位,主要从事数字图像处理及DSP应用方面的研究。E-mail: lgjciom666@126.com

(版权所有 未经许可 不得转载)