

基于图像匹配-点云融合的建筑物立面三维重建

王 俊¹⁾ 朱 利²⁾

¹⁾(西安交通大学电子与信息工程学院 西安 710049)

²⁾(西安交通大学软件学院 西安 710049)

摘 要 随着计算机技术的快速发展,基于图像的建筑物三维重建逐渐成为计算机图形学和计算机视觉领域的研究热点之一.由于建筑物图像背景复杂、序列长且杂乱无序,现有的三维重建算法存在耗时长、局部几何细节重建效果差的问题.文中针对这些不足提出了一种基于图像匹配实现点云融合的建筑物立面三维重建算法.首先寻找新添加的建筑物局部图像在原始图像集中的匹配图像,组成规模较小的图像集并重建出局部点云模型,然后通过匹配不同点云模型在同一幅图像上的投影点,找到点云模型之间的一致对应点集,接着求解点云集合之间的最佳对齐变换,实现整体和局部点云模型的融合,最终生成建筑物立面完整的三维模型.实验表明,采用文中算法进行三维重建,可以有效地减少重建时间,提高重建精度.

关键词 SIFT 流;图像匹配;点云融合;三维重建

中图法分类号 TP391 **DOI 号**: 10.3724/SP.J.1016.2012.02072

3D Building Facade Reconstruction Based on Image Matching-Point Cloud Fusing

WANG Jun¹⁾ ZHU Li²⁾

¹⁾(School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049)

²⁾(Software Engineering School, Xi'an Jiaotong University, Xi'an 710049)

Abstract With the rapid development of computer technology, image-based 3D building reconstruction has become a hot topic in the fields of computer graphics and computer stereo vision. Because the backgrounds of building images are typically complicated, and the sequence is very long and disorderly, existing 3D reconstruction algorithms will take a lot of time to obtain the 3D model, and possibly get poor results in the local areas. This paper proposes a novel 3D building facade reconstruction algorithm based on image matching and point cloud fusing to address those problems. Firstly, we find the best matching images in the additional new image set. Secondly, we get the 2D projection points of the 3D point cloud on the same image. The 3D points corresponding to the same 2D projection points are collected for 3D point cloud fusing. Thirdly, based on the derived 3D corresponding points, we compute the best alignment transformation between the point sets such that they can match with each other with minimum error. Finally, we merge the two point clouds with the computed best transformation to get the complete building facade model. Experiments show that our method can take much less time to reconstruct the building facade, and improve the precision as well.

Keywords SIFT flow; image matching; point cloud fusing; 3D reconstruction

1 引言

随着“数字地球”、“智慧城市”等城市信息化、智能化发展战略的实施,如何高效构建城市三维虚拟环境成为其中的关键技术问题之一,并将有助于实现公共监控、设施管理、数字生活等城市智能环境的建立。建筑物是城市的主要组成部分,各式各样的楼房是居民生活、工作、娱乐的主要场所。因此,城市建筑物的三维重建成为备受关注的重要问题,近几年来在计算机图形学和计算机视觉等领域也逐渐成为研究的热点问题。目前,城市建筑物的三维重建方法主要分为两大类^[1]:基于深度检测设备直接获取三维点云数据和基于图像的三维重建。前者通过具有深度测距功能的系统,如激光雷达(LiDAR),可以直接生成目标建筑物表面的点云模型。虽然这种方法重建的三维模型精度较高,但系统设备成本昂贵,而且不适用于复杂街道环境的建筑物重建^[2]。基于图像的三维重建技术则是通过数码摄像机等拍摄的二维图像,恢复出建筑物表面的三维信息,该方法成本低廉,仅需要普通规格的数码相机拍摄建筑物图片,而且自动化程度高,重建的三维模型包含丰富的纹理等信息,因此一直受到学者的广泛重视。本文针对基于图像的建筑物立面三维重建进行研究,提出了一种更高效的通过图像匹配实现点云融合的重建方法。

基于图像的三维重建的核心内容是 SFM (Structure-From-Motion) 技术^[3]和 MVS (Multi-View Stereo)^[4],即输入一组建筑物二维图像,通过求解对应的摄像机运动参数及几何结构,生成建筑物表面的三维模型。由于维度缺失,从二维图像反求三维运动是一个病态问题,求解困难。通常做法是将摄像机参数估计和三维坐标计算转化为图像平面投影误差极大似然估计,借助非线性最小二乘优化求解^[5]。虽然这种方法原理比较成熟,但由于建筑物图像背景复杂、图像数目较多,而且需要重建的建筑物外立面结构复杂,现有算法将所有图像作为输入的重建过程耗时长,而且重建出的三维整体模型局部效果差,不利于进一步几何细节增强处理。针对这一问题,本文提出了一种借助局部点云融合的三维重建方法,目的在于保证常规方法^[5]重建精度的同时,极大地减少重建时间。该方法首先通过寻找新添加图像在原始图像集中的匹配图像,组成规模较小的局部图像集合并重建出局部点云模型;然后求解整

体点云模型和局部点云模型之间的融合变换;最后通过融合两部分点云模型来实现建筑物立面完整的三维重建。

本文的主要贡献是提出了一种借助点云融合的三维重建算法,通过整体和局部点云的融合,实现建筑物完整三维模型的快速重建。通过这种图像匹配-点云融合方法重建建筑物立面,可以大大缩短重建时间,同时提高三维模型的重建精度。

2 相关工作

近些年来,学者围绕着建筑物的三维重建开展了大量研究。基于 LiDAR 数据的方法虽然可以直接获取建筑物表面的三维点云模型,实现建筑物的快速准确重建,但由于受到其硬件设备的制约,不适用于复杂的天气、地理环境以及具有特殊目的的应用,如军事侦察等^[2,6]。而基于图像的方法由于不受硬件制约,长期以来受到广泛关注,取得了大量研究成果。1996年,Debevec 等人^[7]最早讨论了基于照片图像的建筑物建模与绘制的方法,引起了人们对城市建筑物这一特殊模型重建技术的兴趣。该方法通过基于模型的立体视觉分析技术根据真实场景与图像的差异进行建模,可以实现对简单建筑物立面的快速重建,但是不能对包含大量重复结构的立面进行准确重建。Muller 等人^[8]借助建筑物立面图像的规则语义结构分析,半自动地从图像中分析提取规则矩形的门、窗结构,结合过程建模的方式对复杂建筑物进行重建。虽然该方法能够处理比较复杂的建筑物,但自动化程度不高,而且当重复结构不规则时会产生较大的重建误差。2008年,香港科技大学的 Xiao 等人^[9]提出一种半自动的自底而上的规则结构检测方法,通过预定义简单几何体模型实现建筑物表面结构的三维重建。这种方法对于包含复杂重复结构的建筑物立面重建效果不佳。为了解决这个问题,Xiao 等人^[10]提出一种基于线结构的对称分割方法,可以有效消除结构分析中的语义歧义,能够对大规模街景进行有效的三维重建,但对于非对称的局部几何细节重建效果较差。Micusik 和 Kosecka^[11]则将街景图像拼接成全景图,通过全景序列生成建筑物立面模型,虽然减少了图像集合的数据规模,提高了重建效率,但由于拼接过程中的局部扭曲等造成精度下降。针对互联网上越来越多的大规模无秩序的建筑物图像,Agarwal 等人^[3]提出了一种自适应的分块 SFM 求解方法,可以从大量的建筑

物图像,尤其是互联网图像重建建筑物的三维模型,这也是目前使用最广的摄像机参数估计及三维几何模型求解方法. Musialski 等人^[12]则采用交互式的方法进行立面建模,通过添加内部结构协调性约束生成更加精细的建筑物三维模型,但是这种方法的运行时间很长,不适合大规模的结构复杂建筑物重建.

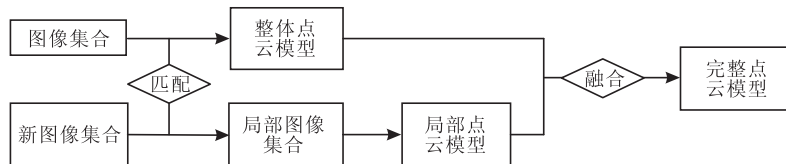


图 1 算法流程

(1) 局部图像集合的构建. 通过 SIFT 流计算图像之间相似性,从整体图像集中寻找到与新加入的局部图像相匹配的图像,然后选择匹配最佳的若干张图像,与新图像组成局部图像集合(见第 4 节).

(2) 整体和局部点云一致对应. 对整体图像集合和构建的局部图像集合分别采用常规方法^[5]重建点云模型. 通过图像匹配寻找整体和局部点云之间的一致对应点集(见第 5.1 节).

(3) 点云模型的融合. 求解对应点集的变换矩阵,实现整体和局部点云的融合,最终生成完整的建筑物立面的三维点云模型(见第 5.2 节).

4 基于 SIFT 流的图像匹配算法

对于整体点云模型中点云稀疏或者缺失的局部区域,我们重新拍摄一些只包含该局部区域的新图像(C_2),并在原始图像集中寻找与新拍摄图像匹配的图像,组成规模较小的新图像集,用以重建局部点云模型.

新拍摄的图像数目选择应遵循以下原则:第一,确保能够使用一般的传统算法^[5]恢复得到建筑物外立面的局部点云模型;第二,在确保第一条的基础上,尽量使用数目较少的图像,以减少重建所花费的时间. 实验表明,新拍摄的图像数目为 5~10 张为宜,通过新的图像实现建筑物立面几何细节的进一步重建.

本文采用基于 SIFT 流^[13]的图像匹配算法,对 C_2 中的任意一张图像 I ,寻找在 C_1 中的最佳匹配图像 J ,建立局部图像集合.

4.1 SIFT 流计算

SIFT 流^[13]是度量两幅图像之间 SIFT 特征分

3 算法简介

本文提出的建筑物立面三维重建算法流程如图 1 所示. 输入的建筑物图像包括原始图像集合 C_1 以及新拍摄的图像集合 C_2 . 通过以下 3 个步骤重建建筑物立面完整的三维点云模型.

布一致性的表示. 由于 SIFT 特征描述尺度不变和亮度独立的图像结构,匹配 SIFT 描述子允许在不同场景内容之间建立一致对应. 因此,采用 SIFT 流可以很好地度量两幅图像之间的局部相似性. 给定图像 I 和 J ,SIFT 流对 I 中的每一个像素 (x, y) 定义位移向量 $(V(x), V(y))$. 图 2(c)展示了从图像 I 到图像 J 的 SIFT 流.

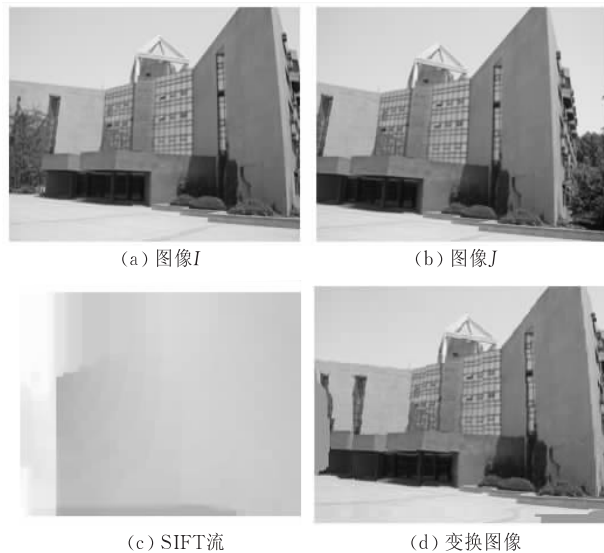


图 2 输入的两幅图像、SIFT 流及变换图像

4.2 建筑物局部图像的匹配

根据图像之间的 SIFT 流,得到集合 C_1 中每一幅图像 J 与图像 I 对应的投影图像. 投影图像的计算公式如下:

$$\begin{cases} I_{ij}(x) = I_j(x) + V(x) \\ I_{ij}(y) = I_j(y) + V(y) \end{cases} \quad (1)$$

其中 I_{ij} 为图像 I 通过与图像 J 之间 SIFT 流变换后的图像(图 2(d)).

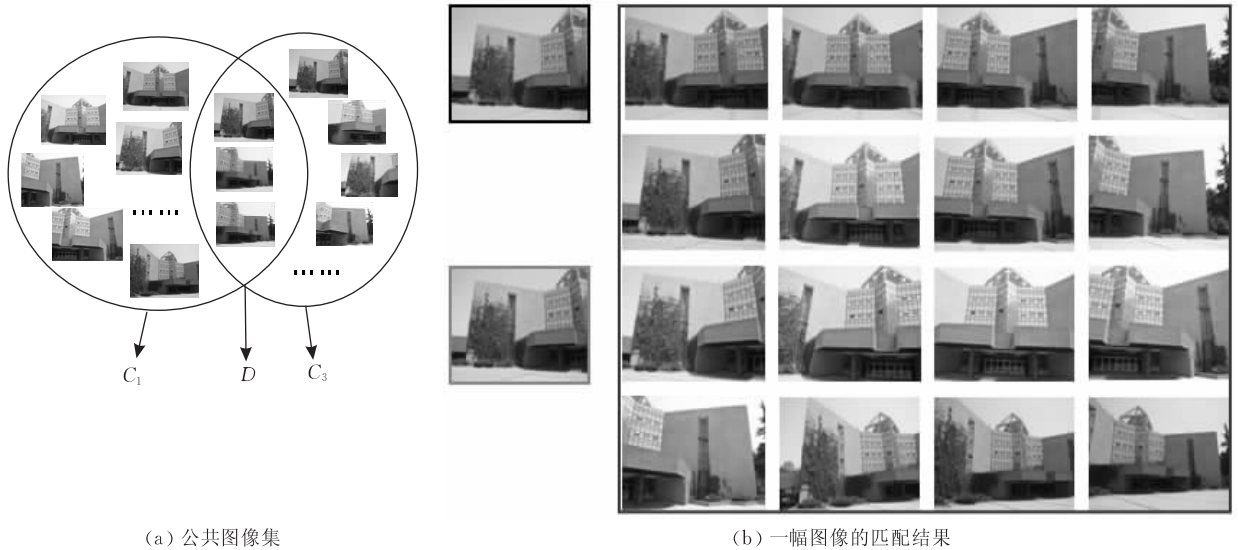
然后, 计算图像 I 和变换图像之间的 $PSNR$ (Peak Signal to Noise Ratio) 值, 并以此计算结果来表示两幅图像的相近程度. $PSNR$ 值越大, 说明两幅图像越相近. $PSNR$ 值的计算公式如下:

$$PSNR = 10 \times \lg((2^n - 1)^2 / MSE) \quad (2)$$

其中, n 为图像像素数目, MSE 为两张图像之间的均方误差.

图 3(b) 展示其中一幅图像的匹配结果. 其中, 位于左上角图像为新拍摄的图像集 C_2 中一张建筑

物的局部图像, 右侧图像集合为原始图像集 C_1 , 而最佳匹配图像为左下角图像. 对 C_2 中的每一幅图像都采用上述方法求得其匹配图像, 最后选择匹配最佳的三幅图像, 加入到 C_2 中, 组成局部图像集 C_3 . 接着采用常规方法^[5] 重建出建筑物外立面的局部点云模型 $P_2 = (X_1, X_2, \dots, X_n)^T$, 其中 $X_i = (x_i, y_i, z_i)^T$ 是点的三维坐标; 而由原始图像集合 C_1 重建的点云记为 P_1 . 由于集合 C_1 和 C_3 中都包含这 3 幅匹配图像, 我们称这 3 幅图像为公共图像集 D (图 3(a)).



(a) 公共图像集

(b) 一幅图像的匹配结果

图 3 公共图像集与图像匹配结果

5 点云融合的建筑物立面三维重建

获取整体和局部点云模型之后, 接下来需要将两部分点云进行融合. 融合算法分为两步: 第 1 步是寻找两部分三维点云模型中一致对应的点集; 第 2 步是计算变换矩阵, 实现整体点云和局部点云的融合.

5.1 局部点云模型的一致对应

我们将整体和局部点云模型分别命名为 P_1 和 P_2 . 统计公共图像 D 中每一张图像在 P_1 和 P_2 中对应的点云数目. 具体方法为, 对 D 中的每一幅图像 K , 读取图像 K 在 V_1 和 V_2 中对应的点云, 命名 P_{k1} 和 P_{k2} . 那么选择包含 V_1 和 V_2 中对应的点云数目都比较多的图像用来引导点云融合, 该图像仍记为 K . 这样选择的目的是为了提高匹配计算的精度和稳定性.

然后计算点云 P_1 和 P_2 分别投影在图像 K 上的像素位置, 记录三维点在共同图像平面上投影点位置坐标. 计算公式为

$$x_i = \mathbf{P}X_i \quad (3)$$

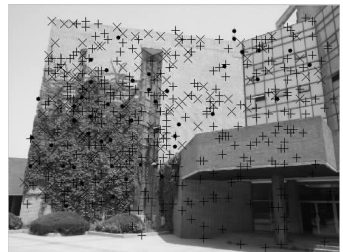
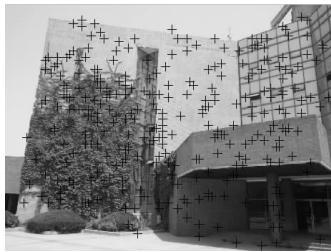
其中, X_i 是三维点坐标, \mathbf{P} 是图像 K 对应的摄像机矩阵, x_i 是 X_i 投影到图像 K 上的二维点坐标.

P_1 和 P_2 在图像 K 上的投影结果分别记为 K_1 和 K_2 , 如图 4 所示. 由式(3)计算得到的投影像素点坐标并不是整数, 首先将所有投影像素点进行坐标取整; 然后对比两张图像中投影像素点的坐标位置, 记录坐标值重合的像素点, 组成点集合 A . 我们认为这些点是匹配准确的点. 图 4 展示了坐标匹配的结果. 其中, 图 4(a) 上的十字点集合对应 K_1 中的投影像素位置, 图 4(b) 上的叉点集合为 K_2 中的投影像素位置, 而通过坐标位置匹配, 得到点云之间对应的点集 (图 4(c) 圆点所示).

为了进一步提高匹配准确度, 我们将点集 A 中的所有点在 K_1 和 K_2 中对应的像素点记为 A_1 和 A_2 . 计算 A_1 和 A_2 中每一对匹配点对之间的距离, 将所有匹配点对的距离值按照由小到大进行排序, 距离值越小, 匹配越准确. 实验中我们取阈值为 0.5, 对于距离小于 0.5 的匹配点对, 我们认为是匹配

准确的;反之,则认为是匹配不准确的,并予以去除.最后记录 A 中剩余点对在点云模型 P_1 和 P_2 中

对应的三维点坐标,我们仍用 P_{k_1} 和 P_{k_2} . 这样,我们就得到了整体和局部点云模型之间的一致点对.



(a) 整体点云的部分投影图像(十字点) (b) 局部点云的部分投影图像(叉点) (c) 不同点云在同一幅图像上的投影,圆点表示匹配正确的点

图 4 投影图像

5.2 点云融合变换

在 5.1 节我们已经找到了三维点云模型 P_1 和 P_2 中的一致点对 P_{k_1} 和 P_{k_2} . 接下来需要寻找一个相似变换 T , 实现 P_{k_1} 和 P_{k_2} 中的所有对应点集的坐标对齐. Iterative Closest Point (ICP)^[14] 方法是最常用的三维模型匹配算法. 该方法通过不断更新最近点的匹配求解变换, 是一个非线性迭代过程. 而本文中, 由于两个点云模型 P_1 和 P_2 的对应点已知, 我们接下来采用具有显式解的方法求解 T . 变换 T 由 3 个参数组成: 平移向量 L 、旋转矩阵 R 和缩放因子 s . 于是求解问题就变成最小化如下目标函数:

$$E(P_{k_1}, P_{k_2}) = \|P_{k_1} - (s \cdot P_{k_2} \cdot R + L)\| \quad (4)$$

该函数的优化是 Procrustes 问题^[15], 我们采用以下步骤逐步求解相应的变换.

(1) 计算平移量 L

首先计算 P_{k_1} 和 P_{k_2} 的几何中心, 公式如下:

$$O = \sum_{i=1}^N X_i / N \quad (5)$$

其中, O 为三维点云模型的几何中心, X_i 是三维点云模型中第 i 个点的坐标, N 是三维点云模型的点云数目. 其中, O 为三维点云模型的几何中心, X_i 是三维点云模型中第 i 个点的坐标, N 是三维点云模型的点云数目. 得到 P_{k_1} 和 P_{k_2} 的几何中心 O_1 和 O_2 后, 平移向量 L 就可以很简单地计算出来:

$$L = O_1 - O_2 \quad (6)$$

实际上, 为了计算方便, 本文采取了坐标归一化的方法, 将 P_{k_1} 和 P_{k_2} 的几何中心都平移到原点 O . 则公式转化为如下形式:

$$L = (O_1 - O_2) + (O - O_1) \quad (7)$$

(2) 计算旋转矩阵 R

通过将 P_{k_1} 和 P_{k_2} 进行坐标归一化, 并且先不考虑缩放因子, 公式可进一步简化为

$$E(P_{k_1}, P_{k_2}) = \|P_{k_1} - P_{k_2} \cdot R\| \quad (8)$$

为了满足点云模型在变换过程中保证结构的不变性, 旋转矩阵 R 需要满足刚体变换的约束条件, 即 $\det(R) = 1$. 该最小化问题可以采用 OPP (Orthogonal Procrustes Problem) 算法^[14] 进行求解, 求解出的最优旋转矩阵 R 表达式如下:

$$R = U \cdot V^T \quad (9)$$

其中, U 和 V 是对 $P_{k_2}^T \cdot P_{k_1}$ 做 SVD 分解得到的特征向量矩阵结果.

(3) 计算缩放因子 s

缩放因子 s 可以通过下面的公式直接求得

$$s = \text{tr}(P_{k_1}^T \cdot P_{k_2} \cdot R) / \text{tr}(P_{k_2}^T \cdot P_{k_2}) \quad (10)$$

其中 $\text{tr}(\cdot)$ 是矩阵的迹. 通过式 (7)、(9)、(10) 计算得到平移、旋转及缩放变换后, 使用如下公式将 P_2 中的点云变换到 P_1 坐标系下:

$$\tilde{P}_2 = s \cdot P_2 \cdot R + L \quad (11)$$

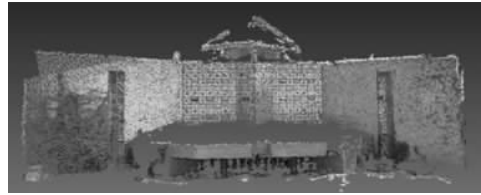
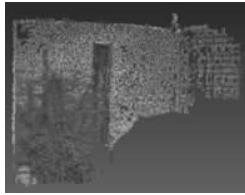
这样, 就实现了整体点云和局部点云的融合 $P_1 \cup \tilde{P}_2$, 完成了整个建筑物立面的重建.

6 实验结果及分析

为了验证本文算法的实验效果, 我们使用了多组建筑物图像集合进行测试. 这些图像都是通过普通数码相机进行拍摄, 在不同视角采集建筑物立面图像, 然后通过图像匹配实现点云融合, 进而恢复所拍建筑物的立面三维模型. 图 6 展示了部分实验结果. 实验运行环境为 Intel 双核处理器 2.27GHz, 8GB 内存, 64 位操作系统. 表 1 给出了不同实验运行的时间统计.

表 1 基于图像的建筑物立面三维模型重建时间统计

模型	整体图像 数目/幅	新拍图像 数目/幅	公共图像 数目/幅	整体点云 耗时/min	局部点云 耗时/min	匹配+融合 耗时/min	一般方法 耗时/min	节省 时间/%
科学馆(图 5)	42	5	3	29.0	6.0	2.5	35	75.7
校门(图 6(a))	36	8	3	24.0	7.5	2.5	33	69.7
化工楼(图 6(b))	44	5	3	33.5	6.5	2.5	38	76.3



(a) 局部点云通过变换与整体点云融合

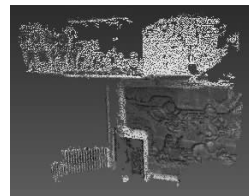


(b) 从左到右: 平移、旋转、缩放变换重建建筑物完整点云模型

图 5 科学馆三维模型重建



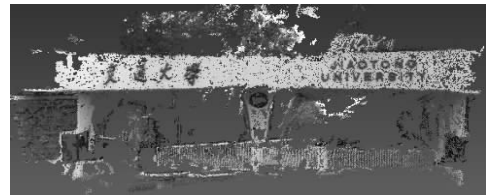
整体点云



局部点云

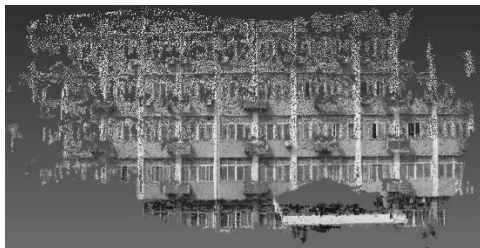


本文算法结果

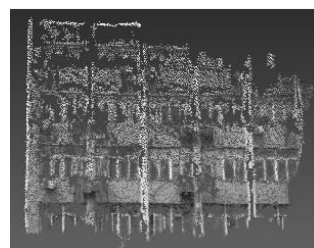


传统算法结果

(a) 校门三维模型



整体点云



局部点云



本文算法结果



传统算法结果

(b) 化工楼三维模型

图 6 校门及化工楼三维模型重建

由图 6(a)中的整体点云可以看到,校门右上角的区域点云是缺失的,这可能是由于没有拍摄到该部分的图像,或者图像特征匹配较差造成.对这部分使用本文提出的融合算法进行了修补;经过图像匹配和点云融合后,本文算法对西安交通大学北门右上角区域的点云已经进行了补充,而且局部细节纹理也更加清楚.重建的结果能够体现整个校门的几何模型.而采用一般的重建算法^[5],并没有获得更好的重建效果,重建的三维模型局部几何细节较差.分析其原因有两个:第一,由于图像数目的增多,造成了部分干扰物也重建出来,从而带来一些杂点;第二,图像数目的增多,也增加了特征点匹配和摄像机参数估计错误的概率,从而造成重建效果差.图 6(b)展示了另一个建筑物侧面的重建效果及对比.

从重建时间(表 1)来看,在整体点云模型的基础上,当加入新的图像时,采用“匹配+融合”方法进行重建比一般的方法节省 50% 以上的时间,尤其是当采集图像的数量较大时,优势更加明显.因此本文算法具有较高的运行效率,可以极大地节省建筑物立面三维模型重建时间.

7 结 论

本文针对重建出的建筑物外立面点云模型中局部点云缺失或者过于稀疏的问题,提出了一种基于二维-三维图像融合的三维重建算法,对点云模型的局部进行融合修补,得到了结构和细节更加完整丰富的建筑物外立面三维点云模型.与一般方法相比,本文提出的方法可以在保证重建精度的前提下,大大减少重建时间,提高重建效率.

当然,本文算法也存在不足,需要进一步改进.第一,本文算法依赖于 SIFT 流检测图像之间的相似性,对于视角变化剧烈的非平移运动效果较差.这一问题可以通过在拍摄新的局部图像集合时控制平移运动量解决.第二,本文算法只考虑了模型中存在点云缺失或者稀疏的问题,而没有考虑模型中存在噪音的问题.在后续工作中应考虑如何去除多余的噪音,以进一步提高重建点云模型的精度.

致 谢 衷心感谢匿名审稿专家,他们对本文提出了宝贵的修改建议!

参 考 文 献

- [1] Musialski P, Wonka P, Aliaga D G, Wimmer M, Gool L, Purgathofer W. A survey of urban reconstruction. *Computer Graphics Forum*, 2012, 31(2): 1-18
- [2] Zheng Q, Sharf A, Wan G W, Li Y Y, Mitra N J, Chen B Q, Cohen-Or D. Non-local scan consolidation for 3D urban scene. *ACM Transactions on Graphics*, 2010, 30(6): 1-9
- [3] Agarwal S, Snavely N, Simon L, Seitz S M, Szeliski R. Building rome in a day//*Proceedings of the International Conference on Computer Vision*. Kyoto, Japan, 2009: 72-79
- [4] Furukawa Y, Ponce J. Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(8): 1362-1376
- [5] Snavely N, Szeliski R. Scene reconstruction and visualization from community photos. *Proceedings of IEEE*, 2010, 98(8): 1370-1390
- [6] Shen C H, Huang S S, Fu H B, Hu S M. Adaptive partitioning of urban facades. *ACM Transactions on Graphics*, 2011, 29(4): 1-10
- [7] Debevec P E, Taylor C J, Malik J. Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach//*Proceedings of SIGGRAPH*. New Orleans, USA, 1996: 11-20
- [8] Muller P, Zeng G, Wonka P, Gool L V. Image-based procedural modeling of facades. *ACM Transactions on Graphics*, 2007, 26(3): 1-9
- [9] Xiao J X, Fang T, Tan P, Zhao P, Ofek E, Quan L. Image-based facade modeling. *ACM Transactions on Graphics*, 2008, 27(5): 1-10
- [10] Xiao J X, Fang T, Zhao P, Lhuillier M, Quan L. Image-based street-side city modeling. *ACM Transactions on Graphics*, 2009, 28(5): 1-12
- [11] Micusik B, Kosecka J. Piecewise planar city 3D modeling from street view panoramic sequences//*Proceedings of the Computer Vision and Pattern Recognition*. Miami, USA, 2009: 2906-2912
- [12] Musialski P, Wimmer M, Wonka P. Interactive coherence-based facade modeling. *Computer Graphics Forum*, 2012, 31(2): 661-670
- [13] Liu C, Yuen J, Torralba A, Sivic J, Freeman W T. SIFT flow: Dense correspondence across different scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 978-994
- [14] Besl P J, McKay N D. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 239-256
- [15] Gower J C, Dijksterhuis G B. *Procrustes Problems*. New York: Oxford University Press, 2004



WANG Jun, born in 1987, M. S. .

His research interests include image and video processing, computer graphics.

ZHU Li, born in 1968, Ph. D. , associate professor.

His research interests include image processing, data communication, multimedia and knowledge engineering.

Background

3D building reconstruction is an important issue in the fields of computer graphics and computer vision. It aims at reconstructing the 3D realistic model of facades of buildings from LiDAR data or images, of which image based reconstruction methods have drawn much attention due to its wide spectrum of applications. A lot of research on fast and accurate reconstruction is conducted in the past decades. However, existing approaches usually take all the captured facade images as input to run the two-step algorithm, bundle adjustment (BA) and multi-view stereo (MVS). Both steps are nonlinear formulations, involving complicated computation task. If the existing set of images is not sufficient to obtain accurate geometric model, new images on the local areas of the facade should be captured and added to the image set, which feed BA and MVS to perform the whole reconstruction again. Such methods consume much more time to get the 3D point cloud due to the nonlinear optimization procedure, yet lack of geometric details in the local areas.

In this paper, a 2D~3D fusion approach is proposed to solve the 3D facade reconstruction from a set of images.

Instead of taking all the images together with the newly added images as input to reconstruct the 3D model, the proposed method only reconstructs the local 3D model from the newly added images. Then, local 3D model is merged with the global one via the correspondence obtained by image matching. The method consists of four steps. Firstly, we find the best matching images in the additional new image set. Secondly, we get the 2D projection points of the 3D point cloud on the same image. The 3D points corresponding to the same 2D projection points are collected for 3D point cloud fusing. Thirdly, based on the derived 3D corresponding points, we compute the best alignment transformation between the point sets such that they can match with each other with minimum error. Finally, we merge the two point clouds with the computed best transformation to get the complete building facade model. Experimental results show that such 2D~3D fusion method can take much less time to reconstruct the building facade with better accuracy on local facade structure.

This work is supported in part by the National Basic Research Project of China (No. 2010CB327902).