

# 一种提升3维重建精度和效率的图像选择方法

王莹, 徐振宇, 郑豫楠, 史雪蕾, 刘怡光

(四川大学 计算机学院, 四川 成都 610065)

**摘要:**大型场景3维重建常需融合各局部深度图,故局部3维数据精确度直接关系整个场景重建精确度,且计算量巨大。针对此问题,提出一种旨在提高3维计算效率和精度的图像选择方法。为提高计算效率,提出一种基于位运算的参考图像选择方法,选取可覆盖场景全部3维点的最小标定图像集子集作为参考图像集;根据影响重建精度的尺度评价因子、参考与邻近图像对应相机主光轴夹角及3维点与相机主光轴夹角建立邻近图像选择的评价函数;在邻近图像的具体选择上,先根据参考图像与邻近图像的覆盖率缩小搜索范围,再根据评价函数对剩余图像进行降序排序,并选择前3张图像作为该参考图像的邻近图像。有效的参考图像选择能够极大地减少大场景多视立体3维重建过程中深度图计算量,有效的邻近图像选择兼顾了重建结果的精度和完整度。将本文方法在两组具有真实3维点云数据的公开数据集和两组室外真实场景拍摄的数据集上与另外两种选图方法以及COLMAP方法进行了对比。实验结果表明:本文方法相比于另外两种选图算法时间节省53%~59%,完整性提高4%~6%,精度提高4%~7%;相比于先进的COLMAP算法时间节省16%以上,完整性提高3%,精度提高2.7%左右。实验结果证明了本文方法的有效性。

**关键词:**多视3维重建;深度图;图像选择;重建精度;计算效率

中图分类号:TP391

文献标志码:A

文章编号:2096-3246(2019)05-0178-07

## An Image Selection Method for Efficient and Accurate 3D Scene Reconstruction

WANG Ying, XU Zhenyu, ZHENG Yunan, SHI Xuelei, LIU Yiguang

(College of Computer Sci., Sichuan Univ., Chengdu 610065, China)

**Abstract:** Large scale 3D scene reconstructions are usually realized via merging local depth maps with heavy computational burden, so that the accuracy of local 3D data will affect the accuracy and the efficiency of scene reconstruction. To tackle this problem, in this paper, an image selection method to improve the computation efficiency and accuracy of 3D reconstruction was proposed. To improve the efficiency, a reference image selection method based on bit operation was proposed. Specially, a minimum calibrated images subset covering all three-dimensional points of the scene was selected as the reference image set, then an evaluation function of neighbor image selection was defined according to the scale evaluation factor, the angle between two camera main optical axes and the angle between 3D points and camera main optical axes. In implementation of the neighbor image selection algorithm, the method firstly removed most of the neighbor image candidates by the overlapping rate of reference image and candidates, and then sorted the remaining candidates by the evaluation indicator with descending order. Effective reference images selection can greatly reduce the amount of depth map calculation in multi-view 3D reconstruction of large scene, while effective adjacent images selection took into account the reconstruction accuracy and integrity. The proposed method was compared with another two methods and the COLMAP algorithm on two public data sets with ground truth and two data sets collected from real outdoor scenes. Experimental results demonstrated that the proposed method saved computing time by 53%~59%, improved the completeness by 4%~6% and the accuracy by 4%~7% in contrast to the other two image selection methods. Besides, the proposed method saved computing time more than 16%, improved the completeness by 3% and the accuracy about 2.7% in contrast to the state-of-the-art COLMAP method. Experimental results on multiple data sets demonstrated the effectiveness of the proposed method.

收稿日期:2018-04-11

基金项目:国家自然科学基金重点项目(61860206007);国家自然科学基金面上项目(61571313);四川省科技创新苗子工程项目(2018048);成都市科技项目(2018-YFYF-00061-GX)

作者简介:王莹(1996—),女,博士生。研究方向:视觉计算;图形图像处理。E-mail: yingwang@stu.scu.edu.cn

网络出版时间:2019-09-04 09:21:41 网络出版地址: <http://kns.cnki.net/kcms/detail/51.1773.TB.20190903.0959.001.html>

**Key words:** multi-view 3D reconstruction; depth map; image selection; reconstruction accuracy; computational efficiency

高效获取大场景精确3维模型在增强现实、文物保护、城市规模建模<sup>[1]</sup>等领域具有广泛的应用前景。多视立体视觉(multi-view stereo, MVS)从多视点图像重建密集的3维模型,以低廉的成本获得与激光扫描仪相当的精度<sup>[2]</sup>且适用性广泛,成为计算机视觉领域一个主要研究方向。基于MVS的大型场景3维重建,涉及数以万计的输入图像,计算量巨大。传统MVS算法认为图像间外观变化较小并且视点分布均匀,以简化计算复杂度,但这与实际数据集相差较大,从而存在较大的重建误差,故有效选择重建图像在保证重建精度及完整性的情况下提高计算效率是亟需解决的问题。

根据文献<sup>[3]</sup>可知,基于深度图融合的3维重建大致分为4个流程:1)无序图像经过从运动恢复结构(structure-from-motion, SFM)算法获得相机位置与场景稀疏3维点坐标;2)参考图像与邻近图像的选择;3)深度图计算;4)深度图融合、点云优化、3维模型生成。本文算法主要研究第2个流程,即参考图与邻近图像的选择。目前大多数的算法都关注于如何提高深度图计算效率和精度,而如何对参考图像和邻近图像进行有效选择的算法较少。Goesele等<sup>[4]</sup>将全局选图和局部选图相结合实现了自适应选图过程,但此方法局部细节重建误差大。在此基础上,Bailer等<sup>[5]</sup>改进了选图全局目标函数,提出一种针对单张采样率变化较大的大型无结构数据集的多视立体方法,该方法不仅考虑了大尺度差异,也考虑了任意小部分最优局部重建的加权合并,但计算开销巨大。Shen等<sup>[3]</sup>将近邻图像选择看作一个组合优化问题,并使用量子进化算法寻求最优解,但该方法仅考虑深度误差,未考虑重建误差。Schönberger等<sup>[6]</sup>提出给每个像素选择一个邻近图像,在一定程度上提升了重建完整性,但由于需要将大量图像同时调入内存,其计算量和内存消耗都较大。

相对于传统深度图融合的算法给场景中每张图像计算深度图,作者提出一种参考图像选择方法:选取少量可覆盖全部场景的图像进行计算,一方面减少了计算量,另一方面也可以防止多张深度图因为误差而导致的深度不一致的现象。传统的邻近图像选择算法,只考虑了深度图的误差,但深度图的误差与空间中3维点之间的误差并不完全一致。为提供减小重建误差的更优图像选择,作者提出一种基于空间3维点误差分析的邻近图像选择算法:根据邻近图像相对于参考图像的覆盖率及相应位置关系计算邻近图像的评价函数,选择评价函数值最高的前3张图像作为邻近图像。

## 1 算法描述

针对无序多视图图像集,重建流程如下:从运动恢复结构(SFM)步骤采用bundle<sup>[7]</sup>获取相机参数及点云信息;应用改进的参考图像与邻近图像选择算法,选择最佳参考及邻近图像进行深度图计算;深度图计算使用PatchMatch算法,PM算法主要分为随机初始化、传播扩散、随机扰动3步,在深度图计算场景下,效果较好<sup>[8]</sup>。深度图融合采用文献<sup>[6]</sup>中方法,最终获得稠密3维点云。

### 1.1 参考图像选择

深度图融合首先需要计算单张深度图,单张深度图对应的2维图像即参考图像。参考图像的选择依赖于3维重建第1个流程中SFM算法<sup>[9-11]</sup>的结果。SFM算法的输出可以看成两个集合(已标定的图像集与3维点集)及其之间的关系。标定的图像是指标定了图像对应相机参数;3维点集是先通过图像之间的特征点匹配,再通过多视几何关系求得的3维点,故一个3维点对应两张或者以上图像中的一个2维点。设图像集合为 $I$ 包含 $n$ 张图像,3维点集合为 $P$ 包含 $m$ 个3维点。两个集合间的关系可以描述如下:

$$P = P_{I_1} \cup P_{I_2} \cup \dots \cup P_{I_n} \quad (1)$$

式中, $P$ 代表3维点集, $P_{I_n}$ 代表图像 $I_n$ 中可见的3维点集。由于一个3维点是通过多张图像中相匹配的2维点求出,故对于任意图像 $I_j$ ,必然存在多个 $k(k \neq j)$ ,使得:

$$P_{I_j} \cap P_{I_k} \neq \emptyset \quad (2)$$

根据式(2)可知图像集合中存在大量的冗余图像,即去掉式(1)中等号右边某些项,等号依旧成立。不同于传统基于深度图融合的算法给标定图像集合中每张图像计算深度图,本文方法为提高计算效率,只选择标定图像集合中的一个子集作为参考图像集计算深度图。为提高计算效率,该参考图像集越小越好,同时,需要考虑场景的完整性,即所有参考图像对应的3维点的并集需完整地覆盖场景中全部的3维点。记向量 $S = [x_1, x_2, \dots, x_n]^T$ 代表参考图像选择的结果,其中, $x_i \in \{0, 1\} (1 \leq i \leq n)$ , $x_i = 1$ 表示图像 $I_i$ 被选中, $x_i = 0$ 表示没有被选中。记向量 $P_j = [p_j^0, p_j^1, \dots, p_j^m]^T$ ,其中, $p_j^k \in \{0, 1\} (1 \leq k \leq m)$ 表示第 $j$ 张图像是否包含3维点 $k$ , $p_j^k = 1$ 表示第 $j$ 张图像包含3维点 $k$ , $p_j^k = 0$ 则表示不包含。记向量 $N = [1, 1, \dots, 1]^T$ ,其中,每个分量都是1。故求解一个包含全部场景的最少图像选择结果等价于以下优化问题:

$$\begin{aligned} s^* &= \arg \min \|S\|_1 \\ \text{s.t. } & [P_0, P_1, \dots, P_n] S \geq N \end{aligned} \quad (3)$$

式(3)所描述的问题是一个0-1规划问题,该类问题是非凸的。对于该类问题,可行的解决方法有枚举法、隐枚举法、割平面方法等<sup>[12-14]</sup>。由于在一个典型的SFM输出结果中包含数万个3维点及成百上千张图片,故在式(3)中约束条件存在数万项,优化参数数量也较大。传统方法难以在可控的时间内求解。作者给出一种简洁快速的近似最优求解方法。

记集合 $R_i$ 代表经过第 $i$ 张图像 $I_i$ 选择的参考图像集, $B_i$ 、 $A_i$ 是二进制比特串,其中,每个比特位与集合 $P$ 中3维点一一对应, $B_i$ 代表 $R_i$ 中已经包含的3维点, $A_i$ 代表图像 $I_i$ 中包含的3维点。则可以用式(4)求得待选的第 $i+1$ 张图像 $I_{i+1}$ :

$$\begin{aligned} x^* &= \operatorname{argmax} \operatorname{counts}(B_i | A_x \odot B_i) \\ \text{s.t. } I_x &\in I \setminus R_i, x = i + 1 \end{aligned} \quad (4)$$

式中, $\operatorname{counts}()$ 为计算一个比特串中包含1的个数的函数, $|$ 为比特串的逻辑或操作, $\odot$ 为比特串的逻辑异或操作。现代计算机的CPU大都集成了以上操作的快速运算指令集,如SSE等。完整的算法流程如算法1所示。

#### 算法1 计算参考图像集

输入:  $P, I, P_{I_x} (x \in [1, n])$

输出:  $R_i$

初始化:  $R_0 = 0, B_0 = 0, i = 1$

Do

$x^* = \operatorname{argmax} \operatorname{counts}(B_i | A_x \odot B_i)$

s.t.  $I_x \in I \setminus R_i, x = i + 1;$

$R_{i+1} = R_i + I_{x^*};$

$B_{i+1} = B_i | A_{x^*}$

While  $B_i \neq P$

针对一个场景进行选图的结果如图1所示,其中,灰色和白色方块代表图像数据集,白色方块代表所选择的参考图像集。从图1中可以看出,只选择较少的一部分图像作为参考图像,即可包含完整的3维场景。

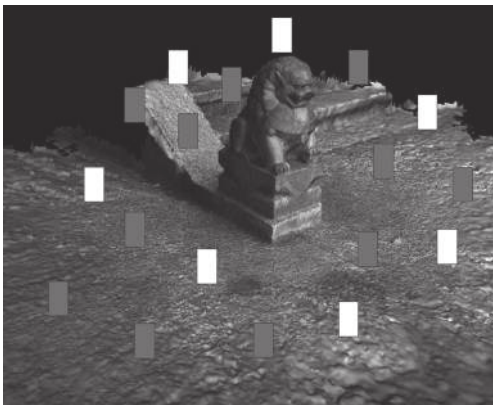


图1 参考图像选择

Fig. 1 Reference images selection

## 1.2 邻近图像选择

通常来说,重建精度主要受到深度图计算方法、3维场景本身的性质(如场景局部表面离相机的远近、场景局部物体表面的连续性等),以及参考图像的邻近图选择等因素的影响。由于多视3维重建中涉及大量计算,基于PatchMatch的多视立体匹配是唯一能够在计算效率和重建精度上都满足条件的一种方法。因此,作者提出的深度图重建方法采用基于PatchMatch的多视立体匹配方法,具有精度高、速度快的优点,且能计算浮点数精度的深度。给定参考图像 $I_r$ 及某一邻近图像 $I_l$ ,由这二者计算的深度图 $D_r$ 中点 $(x, y)$ 的深度为:

$$d^* = \operatorname{argmin} \operatorname{Cost}(d \mathbf{R}^{-1} \mathbf{K}^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + \mathbf{C}) \quad (5)$$

式中: $\mathbf{R}$ 为参考图像的旋转矩阵; $\mathbf{K}$ 为参考图像的摄像机内参矩阵; $\mathbf{C}$ 为参考图像的摄像机中心; $\operatorname{Cost}()$ 为聚合匹配代价函数,定义了参考图像一个区域与邻近图像一个区域之间的相似性,而这两个区域通过3维空间中某个平面映射得到。关于使用PatchMatch计算多视深度图的方法,更加具体的细节可以参考文献[15]。根据式(5)可以看出,计算参考图像上某个点的深度,等价于将一个3维点沿着以参考图像摄像机中心为起点的一条射线移动,在移动过程中通过与邻近图像上的某个区域进行相似度计算,相似度最大时得到该点深度值。由于图像区域之间的匹配误差难以避免,特别是多视3维重建中计算量大且无法采用太复杂的图像匹配方法,故深度图计算也必然存在误差。在多视图图像区域相似性计算中归一化互相关算法(normalized cross-correlation, NCC)是最常用的方法,它能达到整像素级的精度,且计算效率较高,对线性的光照变化不敏感,一般可通过设定相似度阈值使得NCC的匹配误差不超过2个像素。重建误差的形成如图2所示。

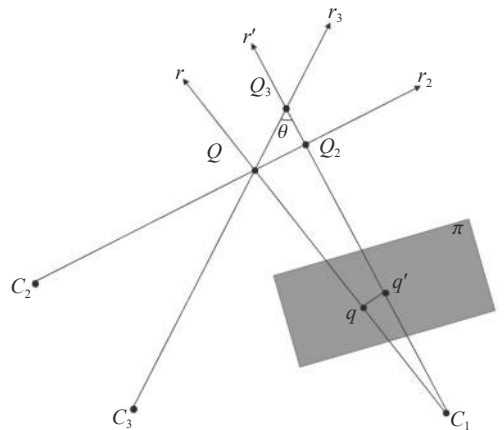


图2 重建误差分析

Fig. 2 Error analysis of reconstruction

图2中,  $C_1$ 为参考图像 $C_2$ 的邻近图像,  $Q$ 为图像 $C_2$ 中某个点对应的正确3维点,  $q$ 为图像 $C_1$ 中与之对应的正确的2维点。但由于存在匹配的误差, 使求得的匹配代价最小的点为 $q'$ , 故实际求得的3维点为 $Q_2$ ,  $\|QQ_2\|$ 即为重建误差, 邻近图像的选择的标准就是使该误差尽可能小。

由于深度图的计算依赖于参考图像与其邻近图像之间的重叠区域, 故给一个参考图像选择邻近图像可以首先根据两个图像之间的覆盖率去除大部分候选者以缩小搜索范围。两个图像之间的覆盖率定义为:

$$\alpha = \frac{\text{counts}(A_i \& A_j)}{\text{counts}(A_i)} \quad (6)$$

本文算法首先删除 $\alpha < 0.3$ 的候选图像。重建误差一方面是由于匹配误差导致, 另一方面是由于参考图像与邻近图像对应的摄像机位置关系引起的。若参考图像的某个区域与邻近图像的某个区域是一个刚性变换, 则因采样频率不一致以及插值误差等导致的匹配误差可以缩减至最小。由文献[16]可知, 仅当两个相机的主光轴方向一致且尺度一致即焦距与3维点到成像平面的距离满足反比关系时, 3维空间中对应的两个图像区域存在一个刚性变换。故对邻近图像的选择, 首先给出基于匹配误差的评价函数:

$$E_m = E_s \times E_d \quad (7)$$

式中,  $E_s$ 表示选图的尺度评价因子,  $E_d$ 表示两个相机主光轴夹角的权重因子。 $E_s$ 定义如下:

$$E_s = \exp\left(-\frac{1}{|P_l \cap P_r|} \times \sum_{p_i \in P_l \cap P_r} \left(1 - \frac{d_r^i f_l}{d_l^i f_r}\right)^2\right) \quad (8)$$

式中:  $f_x$ 为图像 $I_x$ 以像素为单位的焦距;  $d_x^i$ 代表3维点 $p_i$ 到图像 $I_x$ 的像平面的距离, 可由式(9)求得:

$$d_x^i = \left[ \det(\mathbf{R}_x) \times \mathbf{R}_x \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right] \cdot \mathbf{p}_i \quad (9)$$

式中, “ $\cdot$ ”为向量内积运算,  $\mathbf{R}_x$ 为图像 $I_x$ 对应的摄像机旋转矩阵。 $E_d$ 定义如下:

$$E_d = \exp\left(\frac{-\arccos(\mathbf{R}_r^3 \cdot \mathbf{R}_l^3)}{\sigma_d^2}\right) \quad (10)$$

式中,  $\mathbf{R}_x^y$ 代表图像 $I_x$ 的旋转矩阵 $\mathbf{R}_x$ 的第 $y$ 行,  $\sigma_d^2$ 取为 $\frac{\pi}{6}$ 。

图像的重建误差还受到参考图像与邻近图像关系的影响。经过匹配误差筛选条件, 作者选出的参考图像与3维点的距离以及参考图像与邻近图像相机主光轴的方向尽可能一致, 此时匹配误差最小化后

且相机主光轴的夹角以及邻近图像与3维点的距离都大致确定, 要进一步减少重建误差需考虑3维点与参考图像和邻近图像之间的夹角。如图2所示, 固定匹配误差 $\|qq'\|$ 当且仅当从参考图像相机中心发出的与3维点连接的射线 $r_2$ 垂直于邻近图像发出的射线 $r'$ 时重建误差最小。重建误差与 $\theta$ 相关, 易知 $\theta = \angle C_3QC_1 - \angle qC_1q'$ 。由于匹配误差一般都较小, 即 $\|qq'\|$ 一般小于2个像素; 而 $\|C_1q\|$ 大于等于焦距, 即通常情况下会大于1 000个像素单位长度, 故 $\theta$ 近似等于参考图像相机中心与3维点的连线和邻近图像相机中心与3维点的连线的夹角。基于此, 给出邻近图像选择关于3维点与相机中心夹角的权重因子为:

$$E_a = \frac{1}{|P_l \cap P_r|} \sum_{p_i \in P_l \cap P_r} \exp\left(-\frac{\left(\angle C_r p_i C_l - \frac{\pi}{2}\right)^2}{\sigma_a^2}\right) \quad (11)$$

式中,  $\sigma_a^2 = \frac{\pi}{18}$ 。从而, 给一个参考图像选择 $k$ (作者取 $k=3$ )个邻近图像, 可以将由覆盖率筛选过的图像, 根据尺度因素、相机主光轴夹角和3维点与相机中心夹角因素给出式(12)评价函数, 对该函数值进行降序排序并选择前 $k$ 个即可:

$$E = E_s \times E_d \times E_a \quad (12)$$

## 2 实验

为了验证本文方法的性能, 将本文方法与文献[3, 5]选图算法, 以及文献[6]中COLMAP算法在重建时间、场景重建精度、场景重建完整度上进行了对比; 并且, 在两组有真实值的公开数据集Fountain\_P11、Herz-Jesu\_P25<sup>[17]</sup>和两组室外真实场景拍摄的数据集Lion、Stone上, 将本文方法与文献[3, 5]及文献[6]COLMAP算法的重建结果进行对比。在与文献[3, 5]的选图算法进行对比时, 为使实验对比更为公平, 将基于深度图融合的3维重建的4个流程中, 除图像选择以外的其余流程都设置为一致的算法。其中, SFM流程使用bundle<sup>[7]</sup>求出摄像机内外参及3维点信息, 深度图计算和融合采用文献[5]方法。实验运行在一台PC上, CPU为E3 1230(4核8线程主频3.5 GHz), 内存32 GB, 显卡为gtx1080ti。其中深度图的计算启用CPU多线程和GPU运算法。

### 2.1 效率对比

本文方法、文献[3, 5]选图算法, 以及文献[6]中COLMAP算法在Herz-Jesu\_P25数据集上的重建时间对比结果, 如表1所示。重建时间的统计从图像选择开始至3维点云生成为止。由表1可知, 由于本文方法只需要选择能覆盖全部场景的最小图像集进行深度图计算, 其重建时间相对于文献[3, 5]算法减少了

53%~59%，相对于目前较先进文献[6]中COLMAP算法的重建时间减少16%以上。

表 1 Herz-Jesu\_P25数据集上各选图方法重建时间

Tab. 1 Reconstruction time of each image selection method on the Herz-Jesu\_P25 dataset

| 方法   | 本文方法       | 文献[3]方法     | 文献[5]方法     | 文献[6]方法    |
|------|------------|-------------|-------------|------------|
| 重建时间 | 5 min 33 s | 11 min 40 s | 13 min 10 s | 6 min 39 s |

## 2.2 精度和完整性对比

根据文献[17]，重建精度被定义为重建点云集合中所有点到真实点云集合中最小距离的平均值，记录集合 $H$ 为重建的点云， $G$ 为真实点云，则重建点云的精度定义如下：

$$acc = \frac{1}{|G|} \sum \min(\|p_i - p_j\|_2) \quad (13)$$

式中， $p_i$ 、 $p_j$ 分别为点云集合 $G$ 、 $H$ 中的3维点。完整性是指场景被还原的程度，即是否被完整的重建出来，其被定义为：

$$com = \frac{|G'|}{|G|} \quad (14)$$

式中， $G'$ 被定义为：

$$G' = \{p_i | \min(\|p_i - p_j\|_2) < 1.5 \times acc\} \quad (15)$$

4种方法在数据集Herz-Jesu\_P25上的重建精度和完整性如表2所示。本文方法相比于文献[3, 5]方法的完整性提高4%~6%，精度提高4%~7%，这得益于本文方法中更加精确的3维空间重建误差分析，而另外两种方法的误差分析只局限于深度图。相比于目前先进的文献[6]COLMAP算法，本文方法的重建结果完整性提高3.0%，精度提高2.7%左右。

表 2 Herz-Jesu\_P25数据集上各选图方法完整性和精度

Tab. 2 Completeness and accuracy of each image selection method on the Herz-Jesu\_P25 dataset

| 方法      | 完整性/% | 精度/cm |
|---------|-------|-------|
| 本文方法    | 93    | 1.41  |
| 文献[3]方法 | 89    | 1.48  |
| 文献[5]方法 | 87    | 1.52  |
| 文献[6]方法 | 90    | 1.45  |

## 2.3 重建效果对比

为验证本文方法有效性，将本文重建结果与文献[3, 5]方法和文献[6]COLMAP重建结果进行比较。图3为Fountain\_P11图像集在4种方法下的完整重建结果对比。由图3可知，本文方法的重建效果和完整性明显优于其他3种方法。图4为图3中白框部分细节重建结果，在细节信息丰富的区域，本文方法可以保留更多信息。

图5为4种方法在公开数据集Herz-Jesu\_P25上的完整重建结果对比。

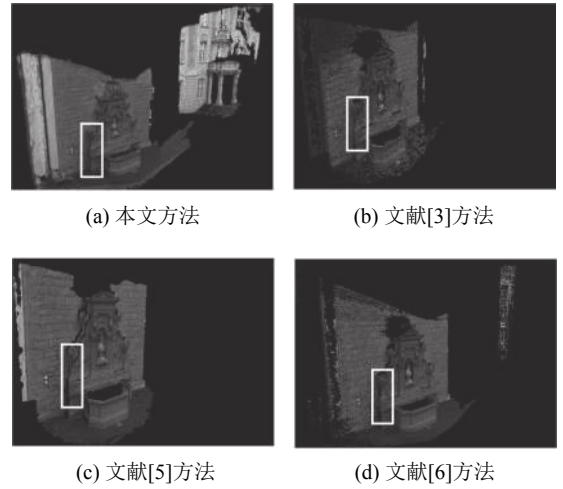


图 3 Fountain\_P11图像集完整重建结果对比

Fig. 3 Comparison of reconstruction results of each method on the Fountain\_P11 dataset

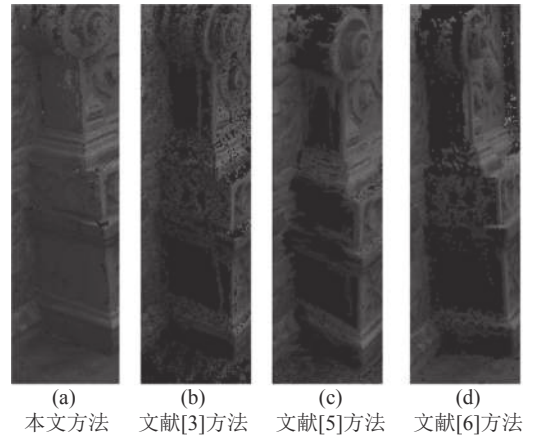


图 4 Fountain\_P11图像集细节重建结果对比

Fig. 4 Comparison of detailed reconstruction results of each method on the Fountain\_P11 dataset

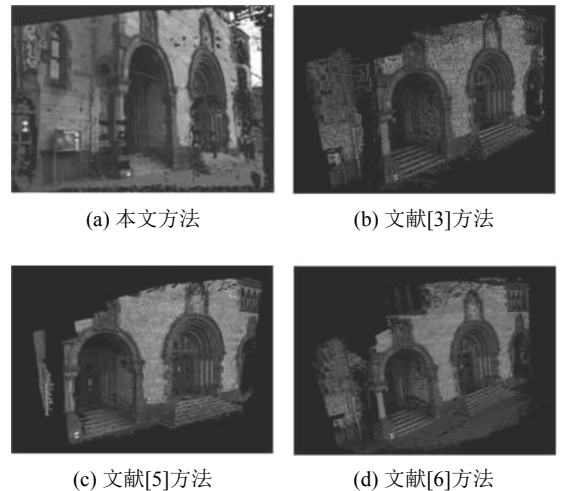


图 5 Herz-Jesu\_P25图像集重建结果对比

Fig. 5 Comparison of reconstruction results of each method on the Herz-Jesu\_P25 dataset

由图5可知,本文方法相比于其他3种方法,在提高重建速度的同时,可以重建出更完整的建筑墙体,并且具有更丰富的纹理。

图6和7是4种方法在户外两组真实场景数据集Lion、Stone的重建结果。由图6可知,本文方法在Lion数据集上对狮子及阶梯的重建完整性明显优于文献[3,5]的方法,在细节纹理上优于文献[6]的COLMAP算法。由图7可知,本文方法在Stone数据集重建出模型的汉字能够清晰识别,对草丛部分的重建效果也优于其他方法。

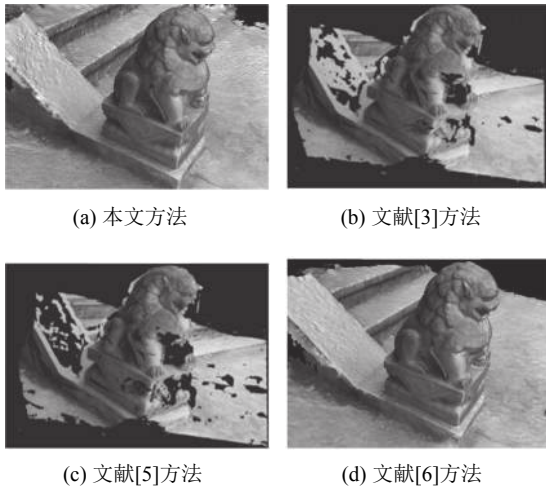


图6 Lion图像集重建结果对比

Fig. 6 Comparison of reconstruction results of each method on the Lion dataset

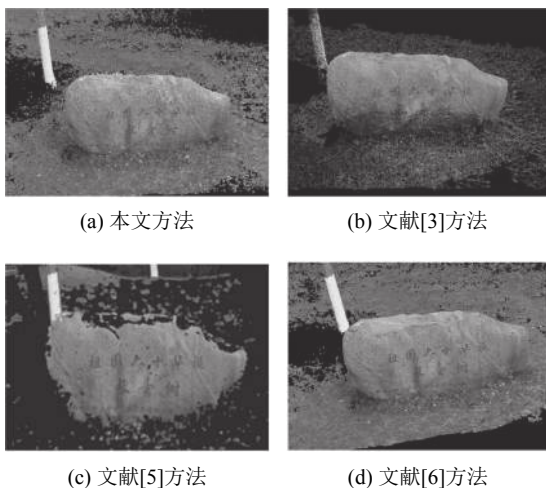


图7 Stone图像集重建结果对比

Fig. 7 Comparison of reconstruction results of each method on the Stone dataset

### 3 结论

针对大场景多视立体3维重建计算量巨大的问题,作者提出了一种提升3维重建效率和精度的图像选择方法。在参考图像的选择上既考虑了计算效率又

考虑了参考图像集所包含场景信息的完整性;在邻近图像的选择上分析尺度评价因子、参考与邻近图像对应相机主光轴夹角及3维点与相机主光轴夹角对重建误差的影响,制定邻近图像选择方法。实验结果表明,本文算法能明显提高重建效率和重建精度。在未来的研究工作中,将继续改进算法以提高多视立体视觉中深度图精度,并进一步提高重建完整性。

### 参考文献:

- [1] Li Ming,Zhang Weilong,Fan Dingyuan.Automatic texture optimization for 3D urban reconstruction[J].*Acta Geodaetica et Cartogra-phica Sinica*,2017,46(3):338-345.[李明,张卫龙,范丁元.城市三维重建中的自动纹理优化方法[J].*测绘学报*,2017,46(3):338-345.]
- [2] Duan Zhixin,Bai Zhihui,Chen Ranli,et al.3D reconstruction of similar material model based on multi-view images[J].*Laser Journal*,2018(8):31.[段志鑫,白志辉,陈冉丽,等.基于多视影像的相似材料模型三维重建方法[J].*激光杂志*,2018(8):31.]
- [3] Shen S,Hu Z.How to select good neighboring images in depth-map merging based 3D modeling[J].*IEEE Transactions on Image Processing*,2014,23(1):308-318.
- [4] Goesele M,Snave N,Curless B,et al.Multi-view stereo for community photo collections[C]//Proceedings of the IEEE 11th International Conference on Computer Vision.*Rio de Janeiro: IEEE*,2007:1-8.
- [5] Bailer C,Finckh M,Lensch H P A.Scale robust multi view stereo[C]//Proceedings of the European Conference on Computer Vision.*Berlin:Springer*,2012:398-411.
- [6] Schönberger J L,Zheng E,Frahm J M,et al.Pixelwise view selection for unstructured multi-view stereo[C]//Proceedings of the European Conference on Computer Vision.*Cham: Springer*,2016:501-518.
- [7] Agarwal S,Snave N,Simon I,et al.Building rome in a day[C]//Proceedings of the IEEE 12th International Conference on Computer Vision.*Kyoto:IEEE*,2009:72-79.
- [8] Liu Yiguang,Yi Shoulin,Wu Pengfei,et al.A novel 3D reconstruction algorithm for large-scale scenes[J].*Journal of Sichuan University(Engineering Science Edition)*,2015,47(6):91-96.[刘怡光,易守林,吴鹏飞,等.一种新的大场景三维重建算法[J].*四川大学学报(工程科学版)*,2015,47(6):91-96.]
- [9] Snave N,Seitz S M,Szeliski R.Modeling the world from internet photo collections[J].*International Journal of Computer Vision*,2008,80(2):189-210.
- [10] Brown M,Lowe D G.Unsupervised 3D object recognition and reconstruction in unordered datasets[C]//Proceedings of the 5th International Conference on 3-D Digital Imaging and Modeling.*Ottawa:IEEE*,2005:56-63.

- [11] Vergauwen M, van Gool L. Web-based 3D reconstruction service[J]. *Machine Vision and Applications*, 2006, 17(6): 411–426.
- [12] Wu Lijun, Zhang Huaxiao. A simplicial solution about 0–1's whole number programming[J]. *Mathematics in Practice and Theory*, 2005, 35(3): 216–219. [吴黎军, 张华孝. 一类0–1整数规划问题的单纯形解法[J]. *数学的实践与认识*, 2005, 35(3): 216–219.]
- [13] Shao W, Hu W, Huang X. A new implicit enumeration method for linear 0–1 programming[C]//Proceedings of the 2008 International Workshop on Modelling, Simulation and Optimization. *Hong Kong: IEEE*, 2008: 298–301.
- [14] Tong Zhe, Zhang Bingjiang, Li Hui. Research on the cutting plane method for solving integer linear programming problems with multiple sets of solutions[J]. *Mathematics in Practice and Theory*, 2017, 47(5): 158–164. [仝哲, 张炳江, 李慧. 关于存在多组最优解的整数线性规划问题的割平面法的研究[J]. *数学的实践与认识*, 2017, 47(5): 158–164.]
- [15] Bleyer M, Rhemann C, Rother C. PatchMatch stereo-stereo matching with slanted support windows[C]//Proceedings of the British Machine Vision Conference. *Dundee: BMVA Press*, 2011, 11: 1–11.
- [16] Lowe D G. Distinctive image features from scale-invariant key points[J]. *International Journal of Computer Vision*, 2004, 60(2): 91–110.
- [17] Strecha C, Hansen W V, Gool L V, et al. On benchmarking camera calibration and multi-view stereo for high resolution imagery[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. *Anchorage: IEEE*, 2008: 1–8.

(编辑 赵 婧)

引用格式: Wang Ying, Xu Zhenyu, Zheng Yunan, et al. An image selection method for efficient and accurate 3D scene reconstruction[J]. *Advanced Engineering Sciences*, 2019, 51(5): 178–184. [王莹, 徐振宇, 郑豫楠, 等. 一种提升3维重建精度和效率的图像选择方法[J]. *工程科学与技术*, 2019, 51(5): 178–184.]