

利用空间几何信息的改进 PMVS 算法

史利民¹ 郭复胜¹ 胡占义¹

摘要 基于多视图图像的立体重建是计算机视觉的核心问题之一. 由 Furukawa 提出的 PMVS 算法是目前为止表现最好的多视立体重建算法之一. 但该算法仍存在一些不足. 一方面, PMVS 不能保证重建表面的几何形状与重建面片法向保持好的一致性, 特别是在一些特定拍摄角度下, 如大场景重建经常碰到的俯仰拍摄, 情况尤为严重. 另外, PMVS 算法时间和空间复杂度高, 特别是在利用高分辨率图像重建时, 往往要付出巨大的时间和空间代价. 针对这些不足, 本文提出了一种基于空间几何信息的面片调整和分辨率分层扩散重建的改进策略, 一方面提高了重建精度和表面的光滑性, 另一方面, 在尽量保持场景细节的同时, 提高了重建效率. 文中的实验证实了改进策略的有效性和实用性.

关键词 多视重建, 三维重建, 法线, 主曲率

DOI 10.3724/SP.J.1004.2011.00560

An Improved PMVS through Scene Geometric Information

SHI Li-Min¹ GUO Fu-Sheng¹ HU Zhan-Yi¹

Abstract Multi-view based reconstruction is one of the central problems in computer vision. In recent years, many excellent algorithms have been reported, such as the PMVS by Furukawa. However, we observed that the following two aspects of the PMVS need further improvements. The first one is that the obtained normal of reconstructed point is not well consistent with its local geometry, and the problem becomes more pronounced under certain image capturing configuration such as downward-shooting or upward-shooting, a frequent practice in large scene reconstruction. The second is its inherent space and time complexity. Especially with high resolution images, its space and time loads become unaffordable. In this work, two remedies are proposed. We propose a patch adjusting trick through the scene geometric information to enhance the patch's normal estimation, and a multi-resolution expanding tactic to well balance the computational cost and the reconstruction accuracy. The experiments demonstrate the effectiveness and practicability of our improved algorithm.

Key words Multi-view reconstruction, 3D reconstruction, normal, principal curvatures

基于多视图的场景重建一直以来都是计算机视觉领域的研究热点, 在诸如测量、虚拟现实、电影娱乐以及文化遗产保护等领域都有广泛的应用. 多视立体重建是利用在不同位置拍摄的关于某一场景的一系列图像, 来恢复场景的三维形状. 在众多研究人员的不懈努力下, 近年来, 涌现出一系列比较实用、鲁棒的重建算法. 在一些测试网站上^[1-2] 的结果显示很多算法都具有比较高的重建精度和完整度, 某些结果甚至可以和激光扫描的数据相媲美. 目前主要的重建算法可以大致分为如下四类: 1) 基于体素的方法^[3-8]. 这类方法一般先建立一个包含待重建物体的立方体包围盒, 将其分割为小的立方体栅格. 每一个小立方体称为一个体素. 其中 Space carving 和 Level set 是这类方法的典型代表. Space carving 通过标记每个体素是空的或是非空的, 并

删除标记为空的体素, 最后剩余的非空体素构成了重建物体模型. Level set 方法通过最小化一个建立在体素上的能量函数, 一般用 Bounding box 作为初始表面, 用偏微分方程驱动初始表面逐步收缩到重建物体表面. 该类方法重建的精度取决于立方体栅格的划分细度, 即体素的大小. 随着剖分细度的增加, 计算量以空间划分细度的三次方迅速增长. 2) 基于多视深度图的方法^[9-14]. 这类算法先针对每个视角的图像重建出深度图, 然后对不同视角之间加以一致性约束而得到完整的深度表面, 或者将深度图融合作为后处理来得到场景表面. 3) 基于多边形网格变形的的方法^[15-17]. 这类方法一般从一个初始的网格表面开始 (如 Visual hull), 通过最小化能量函数或者利用灰度一致性、可见性以及几何一致性等约束驱动初始网格迭代变形, 使网格收敛到物体表面. 不同于体素方法, 它的精度不受空间离散化细度的影响. 4) 基于多视匹配的点 (或面片) 重建方法^[18-21]. 该类算法首先通过多视特征匹配重建出稀疏的空间点 (面片), 然后通过扩散得到物体表面的稠密或准稠密点云 (面片). 另外有些方法集成了上面的某几种方法, 比如通过深度图融合或者重建空间点云首先建立初始的网格表面, 然后采用变

收稿日期 2010-07-22 录用日期 2010-12-27
Manuscript received July 22, 2010; accepted December 27, 2010
国家高技术研究发展计划 (863 计划) (2009AA012103), 国家自然科学基金 (60835003) 资助

Supported by National High Technology Research and Development Program of China (863 Program) (2009AA012103) and National Natural Science Foundation of China (60835003)

1. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190
1. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

形迭代进一步优化网格^[21-22]. 在这四类方法中, 前述三类算法都需要一些初始信息, 如初始 Bounding box 或初始表面、初始视差范围等. 这一要求使得它们比较适用于单个物体的重建, 而不适用于复杂大场景重建, 因为在这种情况下, 所需的初始信息不易估计. 第四类算法则无需任何初始信息. 在这类方法中, 由 Furukawa 等提出的 PMVS (Patch based multi-view stereopsis)^[20] 无论在重建精度还是完整性上, 都具有非常突出的表现, 可以说是这类算法的典型代表, 同时也代表了当前多视重建所达到的水平. 在 Seitz^[1] 等提供的测试网站上, 其各项指标都几乎是最高. 该算法由于不需要任何初始信息如 Bounding-box 或初始场景表面、视差范围等, 所以对于场景重建, 特别是室外大场景重建, 有更广泛的应用价值, 在很多重建应用中都取得了很好的效果^[23-24]. 尽管如此, 我们在实际使用中发现该算法仍有可改进之处. 本文的工作即是在原始的 PMVS 算法基础上, 针对其不足提出的一些改进. 本文的结构安排如下: 第 1 节是对 PMVS 算法的简单介绍, 并分析其存在的不足之处. 第 2 节和第 3 节是针对 PMVS 算法的不足提出的改进算法. 其中第 2 节给出基于法向校正的面片调整算法, 第 3 节给出多分辨率扩散重建的算法框架. 第 4 节是利用本文算法对真实场景的重建实验以及与 PMVS 重建结果的比较. 第 5 节和第 6 节对本文算法进行了进一步讨论, 并给出一些简单结论.

1 PMVS 算法简介

由于本文的工作是对 PMVS 算法的改进, 本节先对 PMVS 算法做一个大致的介绍. PMVS 是一种基于多视匹配的重建算法. 它的基本思路是首先在所有图像上提取 Harris 和 DoG (Difference of Gaussian) 特征点, 然后利用特征匹配、重建, 得到种子点. 最后用种子点向周围扩散得到稠密的空间有向点云或面片. 这里有向是指具有法线方向. 设计严密、实现精细是 PMVS 算法之所以成功的重要原因. 为了方便描述该算法流程和后面的改进算法, 下面首先按照文献 [20], 给出一些必要的定义.

给定已标定图像集合 $Im = \{I_i | i = 1, 2, \dots, n\}$, $O(I_i)$ 为图像 I_i 所对应相机的光心. PMVS 首先将每幅图像 I_i 划分成图像块 $C_i(x, y)$, x, y 为图像块所在的行列索引. 每个图像块大小为 $\beta \times \beta$. 算法最终目标是尽量保证每幅图像的每个图像块内都能重建出空间面片. 对每个重建空间面片 p 有如下几个属性:

- $c(p)$: p 的中心;
- $n(p)$: p 的单位法线;
- $V(p)$: p 的准可见图像集;

$V^*(p)$: p 的可见图像集;

$C(P)$: p 相邻图像块集;

$R(p)$: p 的参考图像.

其中, $V(p)$, $V^*(p)$, $C(P)$ 和 $R(p)$ 的定义分别为

$$V(p) = \left\{ I_i | I_i \in Im, n(p) \cdot \frac{\overrightarrow{c(p)O(I_i)}}{|c(p)O(I_i)|} > \cos(\tau) \right\} \quad (1)$$

$$V^*(p) = \{I | I \in V(p), h(p, I, R(p)) \leq \alpha\} \quad (2)$$

$$C(p) = \{C_i(x', y') | p \in Q_i(x, y), |x - x'| + |y - y'| = 1\}$$

$$R(p) = \min_{I \in V^*(p)} \sum_{J \in V^*(p) \setminus I} h(p, I, J) \quad (3)$$

其中, $h(p, I, J)$ 是面片 p 在图像 I 和 J 上投影的灰度一致性度量, 取为 1 减去其 NCC (Normalized cross-correlation) 值. τ 取为 $\pi/3$.

定义面片 p 在集合 $V(p)$ 和 $V^*(p)$ 上的灰度一致性度量函数

$$g(p) = \frac{1}{|V(p) \setminus R(p)|} \sum_{I \in V(p) \setminus R(p)} h(p, I, R(p))$$

$$g^*(p) = \frac{1}{|V^*(p) \setminus R(p)|} \sum_{I \in V^*(p) \setminus R(p)} h(p, I, R(p)) \quad (4)$$

每个面片的大小定义为其在参考图像上投影的最小正方形包围盒是 $\mu \times \mu$, μ 等于 5 或 7, 本文选择 5. 每个重建面片的最少可见图像数量记为 γ . 对任意图像 $I_i \in Im$ 中任一图像块 $C_i(x, y)$ 有两个对应的集合 $Q_i(x, y)$ 和 $Q_i^*(x, y)$ 用来存储该图像块对应的可见空间面片.

$$Q_i(x, y) = \{p | I_i \in V(p), proj_i(p) \in C_i(x, y)\}$$

$$Q_i^*(x, y) = \{p | I_i \in V^*(p), proj_i(p) \in C_i(x, y)\}$$

其中, $proj_i(p)$ 是 p 在图像 I_i 上的投影.

PMVS 算法主要分为三步: 特征点匹配、扩散、过滤, 其中扩散是其核心. 整个算法流程如图 1 所示. 特征点匹配是通过在所有图像上提取 Harris 和 DoG 特征点, 用满足极几何约束的潜在匹配点对三角化重建出潜在的空间面片的过程. 这些由特征点重建出的空间面片称之为种子面片. 扩散从种子面片开始, 利用相邻面片具有相似的法向和位置的特性, 逐步扩散重建出其周围的空间面片. 为了及时地发现和剔除外点 (即扩散重建错误), 扩散结束后, 要进行过滤处理, 将灰度一致性、几何一致性比较弱的面片剔除. PMVS 迭代执行三次扩散和过滤操作.

通过这种方式可以鲁棒地扩散出准稠密的空间面片. 算法的具体操作参见文献 [20].

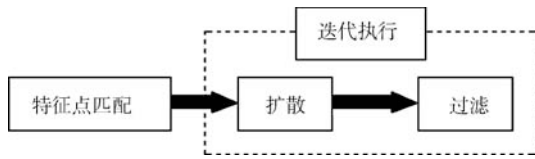


图 1 PMVS 算法流程

Fig. 1 The flow chart of PMVS

尽管诸多应用表明 PMVS 性能优良, 但我们发现, PMVS 中每个面片的法向和位置作为整体优化, 这种优化策略尽管简化了优化过程, 却可能造成重建的面片法向与空间几何形状不一致 (即重建面片的法向存在较大误差). 当法向初值不好时, 这种情况尤为明显. 由于该算法中每个种子面片的法线初值选择为参考图像对应的光心和面片中心的连线, 如果场景图像是仰拍或俯拍得到的 (在大场景重建中这种情况是很常见的), 选择任何一个图像作为参考图像确定出的法向初值很可能都与真实的法向偏离很远, 此时就很容易发生上述问题.

另一方面, 从上面的介绍可以看出, PMVS 算法的时间和空间复杂度都比较大. 每扩散重建一个面片, 就要进行一次关于法向和中心位置的优化. 当图像较多, 特别是图像分辨率很大时, 时间复杂度和空间复杂度将非常大. 虽然可以选择较低的分辨率等级和利用多线程并行处理, 但低的分辨率会造成重建点云的稀疏化和对细节刻画能力以及精度的降低. 而多线程的功效还有赖于计算机的配置, 对于一般配置的机器, 仍然起不到太大的改观.

针对这两大问题, 本文提出了相应的改进策略. 我们知道分布在物体局部表面的稠密或准稠密点云在一定程度上能够提供这一局部区域的几何信息, 比如法线、曲率等. 本文正是以此为出发点, 利用重建出的点云来估计局部几何特征, 进而利用这些几何特征来改进 PMVS 算法的不足.

2 基于法向校正的面片调整

本节及第 3 节将利用扩散重建出的点云来估计重建物体或场景的局部几何特征, 对 PMVS 算法本身存在的上述两个问题做相应的改进. 为了下面描述的方便, 我们有时将重建面片称作点. 面片作为点考虑时, 将其看作位于面片中心且具有法线方向的有向点.

针对 PMVS 算法容易产生面片法向与几何形状不一致, 进而影响到重建精度的问题, 我们在扩散重建的过程中实时地进行法向校正和重新优化. 在每次扩散和滤波处理之后, 利用每个重建点的局部

邻域, 重新估计并修正该面片的法线, 并以新的法线作为初值, 对面片做进一步优化. 改进后的算法流程如图 2 所示.

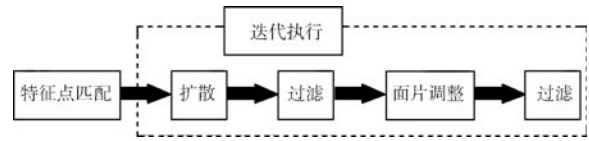


图 2 加入法向校正的 PMVS 算法流程

Fig. 2 The flow chart of PMVS with normal rectification

我们在原始 PMVS 算法的基础上, 在每次迭代过程中加入了面片调整和再次过滤处理. 过滤的算法同前, 面片调整具体操作如下:

设 P 是本次迭代第一次过滤后得到的面片集, $P' \subset P$ 是本次迭代新扩散重建的面片集. $\forall p \in P'$, 在 P 中找到其 k 近邻. 令 P'' 是 P' 中所有面片的 k 近邻组成的集合, 我们要进一步对集合 $B = P'' \cup P'$ 中的面片进行如下调整.

步骤 1. 依次选择 $p \in B$, 执行下面的操作, 直至遍历完 B 中所有的元素.

步骤 2. 选择 p 在 P 中的 k 近邻并计算 k 近邻到 p 的平均距离 d_0 . 如果 $d_0 < d_p$ 且 p 及其 k 近邻中至少包含一个 P' 中的点, 则执行下面的步骤, 否则返回 1.

步骤 3. 用 p 及其 k 近邻作为样本点构造协方差矩阵, 以该协方差矩阵的最小特征值所对应的特征向量作为 z 轴, 其他两个特征值所对应的特征向量作为 x 和 y 轴, 点 p 沿着 z 轴在 xy 平面上的投影点作为原点, 建立局部坐标系. 将 p 及其 k 近邻变换到局部坐标系, 进行二次曲面拟合, 并将 p 沿局部坐标系 z 轴在拟合二次曲面上的投影点处的单位法线转换到原始坐标系, 记作 $n'(p)$.

步骤 4. 确定 $n'(p)$ 的方向: $n'(p) = \text{sign}(n(p)) \cdot n'(p)$.

步骤 5. 如果 $\arccos(n(p) \cdot n'(p)) > \pi/36$, 则执行步骤 6~9, 否则返回步骤 1.

步骤 6. 新建一个面片 d , $c(d) \leftarrow c(p)$, $n(d) \leftarrow n'(p)$, 利用式 (1)~(3) 确定 $V^*(d)$ 和 $R(d)$.

步骤 7. 通过最小化式 (4), 优化 $c(d)$, $n(d)$.

步骤 8. 用式 (1) 增加可见图像到 $V(d)$, 用式 (2) 更新 $V^*(d)$, 用式 (3) 更新 $R(d)$.

步骤 9. 如果 $|V^*(d)| > \gamma$, 在 P 以及 p 所对应的 $Q_i(x, y)$ 和 $Q_i^*(x, y)$ 中删除 p , 将 d 加入 P 及 d 所对应的 $Q_i(x, y)$ 和 $Q_i^*(x, y)$.

完成上述操作, 用更新完的 P 进行过滤及下一次迭代. 上述步骤 2 中 d_p 是距离阈值, 用来限制近邻的所在区域. 如果 p 到其 k 近邻的平均距离比较大, 说明该处面片稀疏, 用 k 近邻估计局部几何性质

不稳定, 故此时不改变 p . 我们选择 d_p 等于面片的长与宽平均值的 $k/4$ 倍. 每次迭代中, 由于新扩散出来的点只会影响其周围邻域, 所以每次只重新调整部分新扩散得到的面片及其近邻. 从后面的实验我们看到, 一方面, 由于实时地利用了局部区域信息重新估计面片法向, 并进一步优化, 使面片法向能够与局部几何形状趋于一致, 在一定程度上提高了重建精度和重建物体表面的光滑性. 另一方面, 由于扩散得到的面片法向和位置更加准确, 有利于扩散重建出更多的面片, 并减少扩散错误.

3 多分辨率分层 PMVS

目前基于图像的三维重建中, 人们把目光主要集中在重建精度的问题上, 而鲜有人考虑密度自适应三维重建算法. 不同于激光扫描, 基于图像的三维重建每一个空间面片(或点)都是从图像计算得到的. 尽管运用高分辨率图像容易得到高精度的重建结果, 但随着图像分辨率的增大, 计算量、存储量的需求都会相应增大, 特别是对于 PMVS 这样的时间、空间复杂度较大的重建算法来说, 问题更加突出. 因此在很多情况下, 我们需要在重建效果和效率之间做出适当的权衡. 如何在尽量保持重建精度、细节的同时提高重建效率正是本节所考虑的问题.

Furukawa 的 PMVS 对用户提供了一个接口, 可以由用户根据需要选择分辨率等级, Level 0 级是原始图像, Level L 是下采样图像, 其长、宽是原始图像长、宽的 $1/2^L$. 选择一个合适的分辨率等级, 使重建结果在时间和空间复杂度上得到一个好的平衡并不是一件容易的事情. 事实上, 重建一个场景或物体时, 并不是处处都需要重建出非常稠密的点云. 我们仅需要在物体结构复杂的区域重建比较稠密的点云, 在结构比较平坦的地方少量的点云足以描述其形状. 基于此种考虑, 本文提出一种多分辨率分层重建的策略. 首先选择在一个较低的分分辨率等级上进行 PMVS 重建, 然后根据重建的结果, 计算每个空间点(面片的中心)的主曲率, 如果其主曲率绝对值的最大值大于某个阈值(说明物体在该点附近至少在一个方向上变化较大), 则在该点的局部区域内在高分辨等级上再次进行 PMVS 处理, 如此反复直到得到满意的结果. 运用这种策略, 算法可以自动地寻找物体的细节部位, 并针对这些部位, 在更高分辨率等级上扩散重建出更稠密和更高精度的点云. 这样, 一方面保证了场景细节的重建, 另一方面有效地提高了重建效率. 下面给出算法的基本框架:

步骤 1. 选择合适的 Level L 进行 PMVS 扩散重建, 得到有向点云 P , 令 $\tilde{P} = P$.

步骤 2. 若 $L = 0$ 则终止程序, 否则 $\forall p \in P$, 在 \tilde{P} 中寻找其 k 近邻并计算 k 近邻到 p 的平均距

离 d_0 . 如果 $d_0 > d_p$, 记其两个主曲率均为 T . 否则利用 k 近邻拟合三次曲面 S (方法同于第 2 节二次曲面拟合). 设 p' 是 p 在 S 上的投影点(同第 2 节), 计算 S 上 p' 处的主曲率 T_1, T_2 , 并将其近似为 p 点的主曲率. 主曲率计算参见文献 [25–26].

步骤 3. 计算每个点主曲率绝对值的最大值, 将主曲率绝对值最大值大于等于阈值 T 的点及到其距离小于 d_p 的近邻构成点集 Q .

步骤 4. 若 Q 是空集, 终止. 否则 $L = \lfloor L/2 \rfloor$, 其中 $\lfloor \cdot \rfloor$ 为下取整. 将所有图像分辨率提升到 Level L 级, 对 Q 中每个点 p , 通过最小化式 (4) 在 Level L 级图像上再次优化其位置和法向, 并根据优化结果重新确定其对应的 $Q_i(x, y), Q_i^*(x, y), V(p), V^*(p)$ 以及 $R(p)$. 然后在其参考图像投影点周围 $W \times W$ 像素的区域作 mask.

步骤 5. 将 Q 加入扩散种子队列, 在所有图像的 mask 域内进行 PMVS 扩散重建, 即保证新扩散的点至少在一幅图像的 mask 域内可见. 新扩散得到的有向点云记作 P' .

步骤 6. $\tilde{P} = \tilde{P} \cup P', P = P', T = T + e$, 返回步骤 2.

\tilde{P} 为最终得到的点云, e 为曲率调节参数, 用来选择在不同分辨率下的种子点. 这里使用三次曲面是为了增加点云对曲面变化的敏感性. 尽管我们的算法可以从一个很低的分分辨率等级开始, 逐层向上, 直到 Level 0. 但对于纹理不太丰富的场景, 太低的分分辨率等级容易产生很多的错误点, 甚至会影响到物体整体的形状结构. 所以在具体操作中应该根据场景的不同, 选择适当的初始分辨率. 另外, 上面我们只是给出了一个算法框架, 根据不同的应用需求用户可以灵活操作, 比如可以根据需求适当选择终止分分辨率等级, 如本文部分实验选择 Level 1.

4 实验

本文实验部分主要利用真实图像, 对原始的 PMVS 算法和本文的改进算法的重建结果进行比较分析. 所有的图像用文献 [27] 的方法确定相机投影矩阵. 实验主要分两部分, 第一部分是对原始的 PMVS 算法和加入法向校正的 PMVS 算法重建结果的比较, 第二部分是单分辨率图像重建和多分辨率分层重建算法的比较.

实验 1 用原始 PMVS 算法和第 3 节算法分别对两组真实图像(悬空寺和云冈石窟, 均包含 10 幅图像)进行重建, 其结果如图 3 和图 4. 在图 3 和图 4 中, (a) 为重建图像, 分辨率为 3008×2000 ; (b) 为原始 PMVS 算法重建的结果; (c) 为加入法向校正后得到的重建结果. (b) 和 (c) 均显示的是带有法线方向的面片.

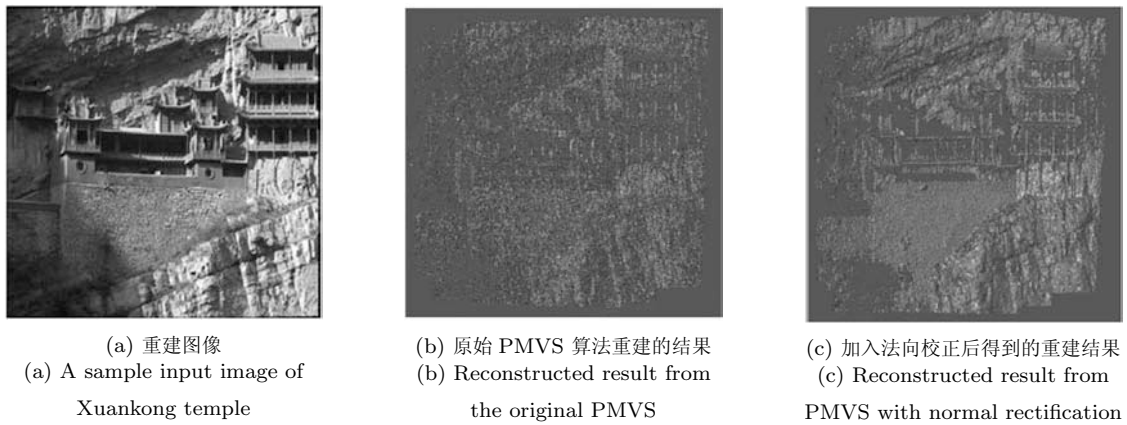


图 3 利用原始 PMVS 算法和加入法向校正的 PMVS 算法重建悬空寺的结果

Fig. 3 Reconstructed results of the Xuankong temple by PMVS with and without normal rectification

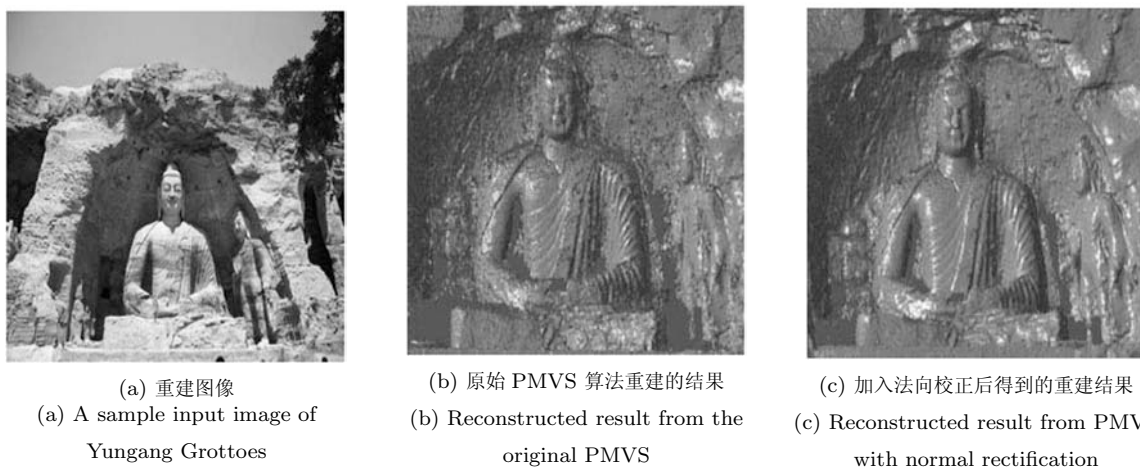


图 4 利用原始 PMVS 算法和加入法向校正的 PMVS 算法重建云冈石窟大佛的结果

Fig. 4 Reconstructed results of the Buddha of Yungang Grottoes by PMVS with and without normal rectification

由于悬空寺拍摄角度不好选择,基本上只能仰拍,如此导致初始法向和真实法向偏离很远,故很多重建点的法线难以优化到真实的方向上.从图 3(b)可以看出,所有点云法向比较散乱.特别是对于墙面,本应该具有比较一致的法向,但 PMVS 的结果却并非如此.而加入法向校正的扩散重建的结果中,点云法线指向基本上是符合真实情况的.另外,正确的法向有利于提高空间点的定位精度,从而有利于扩散出更多的点.在对悬空寺重建实验中,利用面片调整的扩散重建最终得到 250 157 个有向点,而原始的 PMVS 算法得到 235 023 个有向点.

云冈石窟的拍摄角度比较好,接近于正视拍摄,故 PMVS 算法中初始法向较好,优化得到法线多数也比较接近真实值.尽管如此,一些局部位置仍然存在错误,如面部、衣襟等部位.利用法向校正的扩散重建结果中,点云的法向与几何保持了更好的一致性,看上去表面更加光滑.类似于悬空寺的重建实验,通过面片调整的算法比原始算法多扩散出近 10% 的面片.

根据 PMVS 算法的原理,好的法向有助于更精确地确定空间点的位置,所以改进后的算法,在重建精度上应该有所提高.我们通过利用网站^[2]上提供的数据来验证本文的结论.分别用原始的 PMVS 算法和加入法向校正处理的 PMVS 算法对 Fountain 数据的 11 幅图像进行扩散重建 (Level 1).图 5 显示了场景的 Ground truth 和本文法向校正 PMVS 算法的重建结果.通过利用 Geomagic Studio 8.0 软件将重建表面与真实模型相比较,我们得到了重建误差,如表 1 所示(其中正负分别表示正法向和负法向的误差).尽管此时不存在拍摄角度差的问题,但经过面片调整后的算法在多重重建出近 10% 的面片的基础上(原始 PMVS 算法: 408 097,本文算法: 456 018),精度上仍有一定提高.从表 1 看到,平均误差减小得并不十分明显,正法向和负法向分别减小 15% 和 7%,本文认为这一方面是由于 PMVS 算法本身精度已经比较高,且相机标定在很大程度上也影响着最终的重建精度,所以要在精度上有大的提高,单纯在空间几何上调整是不太容易做到的.

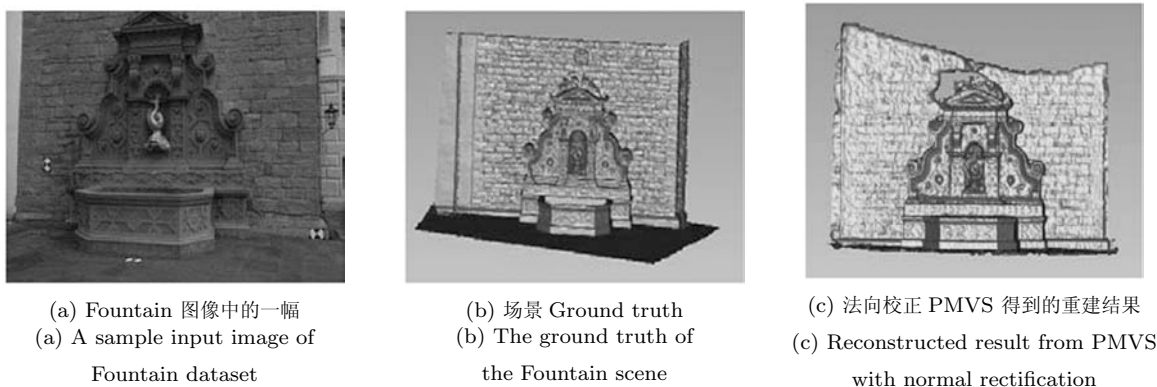


图 5 利用加入法向校正的 PMVS 算法重建 Fountain 的结果
 Fig. 5 Reconstruction results of Fountain dataset by PMVS with normal rectification

表 1 原始 PMVS 算法和加入法向校正的 PMVS 算法对 Fountain 数据的重建精度比较

Table 1 A comparison of accuracies between the results by PMVS with and without normal rectification respectively (on Fountain dataset)

算法	最大误差 (cm)	平均误差 (cm)	标准差 (cm)
法向校正 + PMVS	14.9872/-15.1577	0.4631/-0.8796	1.0981
PMVS	27.2275/-27.2221	0.5453/-0.9456	1.5163

但我们看到标准差以及最大误差有明显的减小, 说

明通过法向校正, 可以很好地滤掉那些误差较大的面片, 使所有重建面片尽可能地逼近于真值.

另外需要强调的是, 以上三组实验数据都表明, 加入法向校正的算法比原始算法能多扩散出 10% 左右的点, 说明改进后的算法在扩散性能也有较大的提高.

实验 2 是利用多分辨率分层扩散和在单一分辨率图像上重建结果的比较. 图 6 和图 7 是对云冈石窟大佛重建结果的比较, 图像分辨率为 3008×2000 . 首先在 Level 2 (分辨率为 752×500) 上进行扩散重

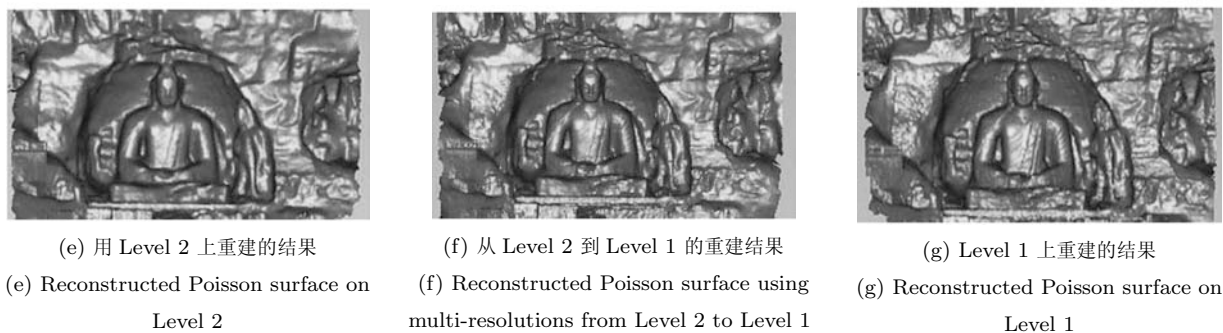
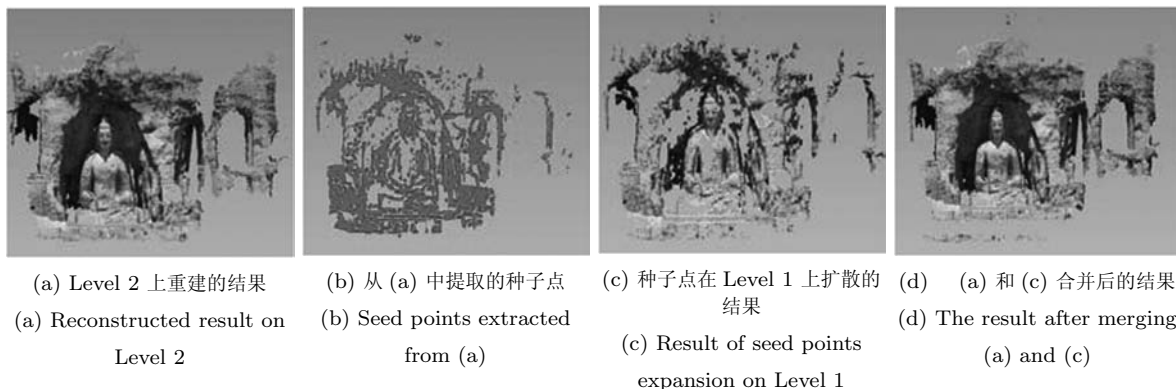
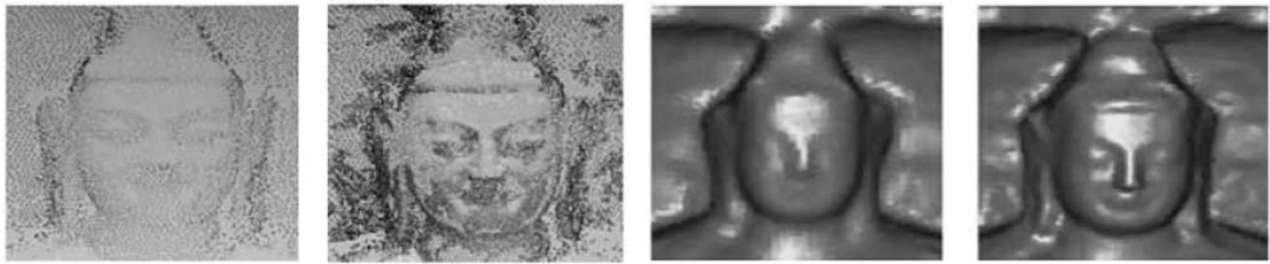


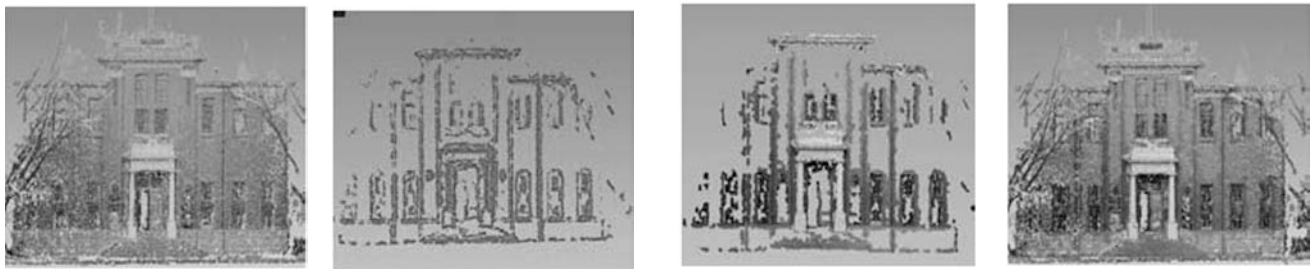
图 6 云冈石窟大佛多分辨率分层重建结果以及和单分辨率重建结果的比较

Fig. 6 Comparison of the results of the Buddha of Yungang Grottoes reconstructed by PMVS and our multi-resolution PMVS algorithm



(a) 图 6(a) 中结果的局部特写 (a) Local detail of result in Fig. 6(a)
 (b) 图 6(d) 中结果的局部特写 (b) Local detail of result in Fig. 6(d)
 (c) 图 6(e) 中结果的局部特写 (c) Local detail of result in Fig. 6(e)
 (d) 图 6(f) 中结果的局部特写 (d) Local detail of result in Fig. 6(f)

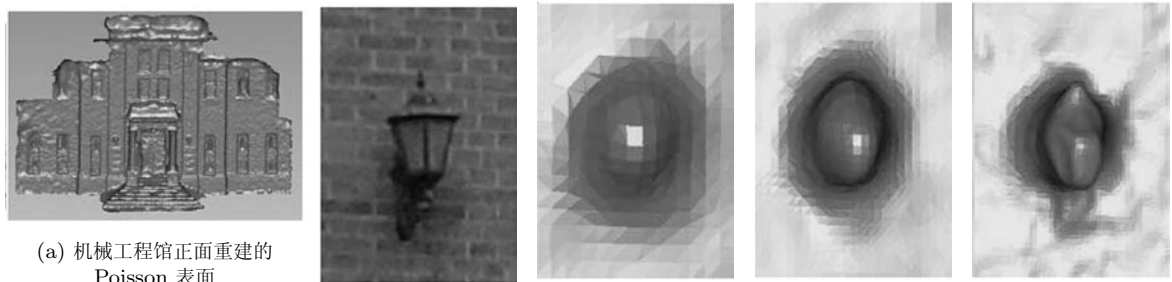
图 7 图 6 中重建结果的局部特写
 Fig. 7 Local detail of results in Fig. 6



(a) Level 2 上重建的结果 (a) Reconstructed result on Level 2
 (b) 从 (a) 中提取的种子点 (b) Seed points extracted from (a)
 (c) 种子点在 Level 1 上扩散的结果 (c) Result of seed points expansion on Level 1
 (d) (a) 和 (c) 合并后的结果 (d) The result after merging (a) and (c)

图 8 清华大学机械工程馆正面多分辨率分层重建结果

Fig. 8 The mechanical engineering building in Tsinghua University reconstructed by our multi-resolution PMVS



(a) 机械工程馆正面重建的 Poisson 表面 (a) Reconstructed Poisson surface of the mechanical engineering building
 (b) 机械工程馆局部 (b) The local detail of the mechanical engineering building
 (c) 对应于 (b) 中区域在 Level 2 上 PMVS 的重建结果 (c) Reconstructed result of local region in (b) using PMVS on Level 2
 (d) 对应于 (b) 中区域从 Level 2 到 Level 1 多分辨率 PMVS 的重建结果 (d) Reconstructed result of local region in (b) using multi-resolution PMVS from Level 2 to Level 1
 (e) 对应于 (b) 中区域从 Level 2 到 Level 0 多分辨率 PMVS 的重建结果 (e) Reconstructed result of local region in (b) using multi-resolution PMVS from Level 2 to Level 0

图 9 机械工程馆正面及局部重建结果

Fig. 9 The front side and local detail of the mechanical engineering building reconstructed by multi-resolution PMVS

建, 得到 178 753 个点 (图 6(a)), 通过选择曲率阈值 $T = 5$, 得到 68 460 个种子点 (图 6(b)). 利用这些种子点在 Level 1 上再次扩散, 重建得到 261 664

个点 (图 6(c)). 将两次扩散重建的点云合并得到最终的重建点云 (图 6(d)). 图 6(e) ~ 6(g) 分别是在 Level 2 上和从 Level 2 到 Level 1 采用多分辨率

扩散重建以及在 Level 1 上重建得到的点云后, 再用 Poisson 表面重建算法^[28] ($depth = 12$) 得到的表面模型. 图 7 是图 6 重建结果的局部特写. 表 2 列出了这三种情况下扩散重建的时间. 通过比较可以看出采用多分辨率分层重建在细节刻画能力上比 Level 2 有了明显提高, 接近于 Level 1, 而运算时间却比 Level 1 少很多.

很显然这种基于曲率的多分辨率扩散策略更适用于高曲率部分仅占很小比例的场景, 如建筑物. 图 8 是对清华大学机械工程馆的重建结果. 本实验利用 10 幅分辨率为 3008×2000 的图像. 首先在 Level 2 上得到扩散重建点云包含 62 153 个点 (图 8 (a)), 然后选择曲率阈值 5 得到 8 676 个种子点 (图 8 (b)). 在 Level 1 上对种子点进一步扩散, 重建出 67 897 个点 (图 8 (c)), 图 8 (d) 是合并两次重建点云的结果. 表 2 给出了不同情况的重建时间. 可以看出从 Level 2 到 Level 1 分层扩散的结果, 相比较 Level 2 上的结果, 棱角处点云更加密集. 且由于场景中高曲率的部分比例很低, 所以运算效率和直接在 Level 1 上重建相比, 有了更加明显的提高. 图 9 是重建的 Poisson 表面以及不同终止分辨率下重建的局部特写. 随着终止分辨率的提高, 细节特征更加明显.

本文实验环境: Intel 双核 (单核主频 2.33 GHz) 计算机, 2 G 内存, Windows 32 操作系统. 算法利用 Visual C++ 实现. 本文所有实验中均取 $\beta = 2$, $k = 100$, $W = 15$.

表 2 实验 2 中多分辨率 PMVS 与原始 PMVS 运行时间表
Table 2 Time costs of PMVS and our multi-resolution PMVS in Experiment 2

场景	PMVS		多分辨率 PMVS
	Level 2	Level 1	Level 2 到 Level 1
云冈石窟大佛	516 s	3 327 s	1 117 s
机械工程馆	178 s	1 583 s	373 s

5 进一步讨论

本文中二次、三次曲面拟合以及法线和曲率计算都调用了 CGAL 算法库^[26] 实现, 采用 KD 树寻找近邻, 这部分的运算时间与整个 PMVS 的运行时间相比很小, 如几十万个点的法线调整在本文实验使用的计算机上使用单线程也可在十几秒内完成. 面片调整主要的运算时间花费在面片的再次优化上. 由于每次迭代要调整部分新扩散出来的面片及其附近的面片, 所以加入面片调整的 PMVS 算法比原始的 PMVS 算法运行时间上会有所提高.

PMVS 利用图像分辨率和图像上均匀采样相结合来控制重建密度. 在分辨率为 3008×2000 的图像序列, 如果直接在原始分辨率图像上采用较小的采样间隔, 一个中型的场景, 比如机械工程馆这样的

建筑物, 会产生几 G 的数据量, 要花费几十个小时, 而同样的场景利用激光扫描, 即便在扫描间隔比较小的情况下, 也只有几百兆的数据, 可想而知前者的生成数据中一定包含了大量的冗余点. 本文采用的多分辨率由粗到细的扩散重建策略, 在每个分辨率等级上都用较小的采样间隔以减少扩散错误, 分辨率由低到高逐层得到更加细致的细节特征, 在很大程度上避免了冗余点的生成, 节省了内存空间和运算时间. 而这种根据重建物体的几何特征密度自适应重建是激光扫描难以做到的. 尽管本文的重建实验仅使用了少量的图像, 但已经能反映这种多分辨率扩散策略在兼顾重建效率和质量方面的优势, 相信对大规模数量图像的重建, 其效果会更加明显. 下一步我们将计划在大场景重建中进一步验证这一策略的有效性.

曲率阈值 T 的选取可以采用两种方式, 一种是根据用户对场景的把握直接给定, 一般选择 $5 \sim 10$. 另一种是通过控制种子点数量自动确定, 如先将点云按主曲率绝对值最大值由大到小排序, 按照用户确定的种子点大致比例确定出曲率阈值.

6 总结

尽管 PMVS 算法是多视重建算法中表现最好的算法之一, 但在实际应用中仍存在一些不足. 如扩散过程中, 由于面片的位置和法向捆绑优化, 不能保证面片法向与重建表面的几何形状保持一致, 进而影响到重建精度. 特别是在由于拍摄角度不好等因素造成的法向初值比较差的情况下, 这种问题尤为严重. 另外该算法较大的时间空间复杂度, 使其对于高分辨率的图像和大场景重建效率很低. 本文针对这些不足, 提出了一种基于物体空间几何结构的面片调整和多分辨率密度自适应重建算法, 对原始的 PMVS 做了适当改进. 实验结果表明, 这两处改进保证了扩散重建的有向面片具有更好的法线、位置精度和更好的空间分布, 在保证一定细节描述能力的同时有效提高了高分辨率图像下场景重建的效率.

References

- Seitz S, Curless B, Diebel J, Scharstein D, Szeliski R. Multi-view stereo evaluation [Online], available: <http://vision.middlebury.edu/mview/>, June 10, 2010
- Strecha C. Multi-view stereo evaluation web page [Online], available: <http://cvlab.epfl.ch/~strecha/multiview/>, June 10, 2010
- Paris S, Sillion F X, Quan L. A surface reconstruction method using global graph cut optimization. *International Journal of Computer Vision*, 2006, **66**(2): 141–161
- Pons J P, Keriven R, Faugeras O D. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, 2007, **72**(2): 179–193

- 5 Tran S, Davis L S. 3D surface reconstruction using graph cuts with surface constraints. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 219–231
- 6 Hornung A, Kobbelt L. Hierarchical volumetric multi-view stereo reconstruction of manifold surfaces based on dual graph embedding. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2006. 503–510
- 7 Kutulakos K N, Seitz S M. A theory of shape by space carving. *International Journal of Computer Vision*, 2000, **38**(3): 199–218
- 8 Seitz S M, Dyer C R. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 1999, **35**(2): 151–173
- 9 Strecha C, Fransens R, Van G L. Combined depth and outlier estimation in multi-view stereo. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2006. 2394–2401
- 10 Bradley D, Boubekeur T, Heidrich W. Accurate multi-view reconstruction using robust binocular stereo and surface meshing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA: IEEE, 2008. 1–8
- 11 Kolmogorov V, Zabih R. Multi-camera scene reconstruction via graph cuts. In: Proceedings of the 7th European Conference on Computer Vision. London, UK: Springer, 2002. 82–96
- 12 Gargallo P, Sturm P. Bayesian 3D modeling from images using multiple depth maps. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2005. 885–891
- 13 Goesele M, Curless B, Seitz S. Multi-view stereo revisited. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 2402–2409
- 14 Campbell N D F, Vogiatzis G, Hernandez C, Cipolla R. Using multiple hypotheses to improve depth-maps for multi-view stereo. In: Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer, 2008. 766–779
- 15 Zaharescu A, Boyer E, Horaud R. Transformesh: a topology-adaptive mesh-based approach to surface evolution. In: Proceedings of the 8th Asian Conference on Computer Vision. Tokyo, Japan: Springer, 2007. 166–175
- 16 Furukawa Y, Ponce J. Carved visual hulls for image-based modeling. *International Journal of Computer Vision*, 2009, **81**(1): 53–67
- 17 Esteban C H, Schmitt F. Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding*, 2004, **96**(3): 367–392
- 18 Lhuillier M, Quan L. A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(3): 418–433
- 19 Habbecke M, Kobbelt L. Iterative multi-view plane fitting. In: Proceedings of the 11th International Fall Workshop on Vision, Modeling, and Visualization. Aachen, Germany: Akademische Verlagsgesellschaft Aka GmbH, 2006. 73–80
- 20 Furukawa Y, Ponce J. Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(8): 1362–1376
- 21 Hiep V H, Keriven R, Labatut P, Pons J P. Towards high-resolution large-scale multi-view stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 1430–1437
- 22 Tylecek R, Sara R. Refinement of surface mesh for accurate multi-view reconstruction. *The International Journal of Virtual Reality*, 2010, **9**(1): 45–54
- 23 Furukawa Y, Curless B, Seitz S M, Szeliski R. Manhattan-world stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 1422–1429
- 24 Furukawa Y, Curless B, Seitz S M, Szeliski R. Towards internet-scale multi-view stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 1434–1441
- 25 Chen Wei-Huan. *Differential Geometry*. Beijing: Beijing University Press, 2006
(陈维桓. 微分几何. 北京: 北京大学出版社, 2006)
- 26 CGAL manual [Online], available: <http://www.cgal.org/Manual/>, June 10, 2010
- 27 Snavely N, Seitz S M, Szeliski R. Photo tourism: exploring photo collections in 3D. *ACM Transactions on Graphics*, 2006, **25**(3): 835–846
- 28 Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction. In: Proceedings of the 4th Eurographics Symposium on Geometry Processing. Cagliari, Italy: Eurographics Association, 2006. 61–70



史利民 中国科学院自动化研究所博士研究生. 2000 年获山西大学计算数学专业学士学位. 2005 年获大连理工大学计算数学专业硕士学位. 主要研究方向为图像处理和三维重建. 本文通信作者.

E-mail: slm_sx@126.com

(**SHI Li-Min** Ph. D. candidate at the Institute of Automation, Chinese

Academy of Sciences. He received his bachelor degree from Shanxi University in 2000 and his master degree from Dalian University of Technology in 2005, respectively. His research interest covers image processing and 3D reconstruction. Corresponding author of this paper.)



郭复胜 中国科学院自动化研究所博士研究生. 分别于 2004 年和 2008 年获解放军信息工程大学摄影测量与遥感专业学士学位和硕士学位. 主要研究方向为三维重建和摄影测量.

E-mail: fsguo@nlpr.ia.ac.cn

(**GUO Fu-Sheng** Ph. D. candidate at the Institute of Automation, Chinese

Academy of Sciences. He received his bachelor degree in 2004 and master degree in 2008 from the PLA Information Engineering University. His research interest covers 3D reconstruction and photogrammetry.)



胡占义 中国科学院自动化研究所研究员. 主要研究方向为摄像机标定、三维重建和视觉机器人导航.

E-mail: huzy@nlpr.ia.ac.cn

(**HU Zhan-Yi** Professor at the Institute of Automation, Chinese Academy of Sciences. His research interest covers camera calibration, 3D reconstruction, and vision guided robot navigation.)