论 文

无人机图像全自动生成大比例尺真正射影像方法

郭复胜*, 高伟, 胡占义

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

* 通信作者. E-mail: fsguo@nlpr.ia.ac.cn

收稿日期: 2013-01-08; 接受日期: 2013-05-10

国家高技术研究发展计划 (863 计划)(批准号: 2013AA12A202)、国家自然科学基金 (批准号: 61203278) 和中国科学院战略性先导科技专项 (批准号: XDA06030300) 资助项目

摘要 随着无人机遥感技术逐步从研究开发阶段发展到实际应用阶段, 采用无人机图像进行大比例尺的真正射影像生成方法的研究具有重要的现实意义. 本文旨在探索将计算机视觉中获得巨大成功的多视三维重建技术应用到对无人机影像处理中, 给出了一种基于运动恢复结构重建算法和多视图立体视觉算法全自动生成大比例真正射影像的方法. 本文首先分析了无人机图像 PMVS 重建点云的特点, 给出一种由基于面片多视图立体视觉稠密点生成数字表面模型的方法, 然后详细介绍了包括正射影像图像坐标映射模型、可见性计算、基于 Markov 随机场能量优化的面片选择和匀光处理等真正射影像生成的关键步骤. 实验结果以及和商业软件的比较表明: 本文给出的方法在野外地形和城市区域均能获取有效的真正射影像结果.

关键词 真正射影像 无人机 多视立体几何 PMVS Markov 随机场

1 引言

数字正射影像图具有表现直观、真实、信息量丰富、易于更新等优点, 因而逐渐成为数字城市空间数据框架的重要组成部分, 并在国土资源调查、城市规划、城市管理和军事应用等方面发挥越来越重要的作用. 与其他移动测图平台相比, 无人机系统具有机动、灵活、低成本等特点, 因此非常适合应用于大比例尺测图^[1]. 随着无人机遥感技术逐步从研究开发阶段发展到实际应用阶段, 采用无人机图像进行大比例尺的真正射影像生成方法的研究具有很重要的现实意义. 数字正射影像纠正是指同时消除地形和地面三维目标投影变形的正射影像处理流程. 传统的正射影像图是中心投影, 又称为伪正射影像图, 而真正射影像图是指垂直投影, 需要利用三维数字表面模型 DSM 进行微分纠正, 消除投影差, 使地形与地物均被纠正到正确的平面位置, 从而在真正意义上确保影像处处为垂直正射投影.

关于真正射影像生成算法理论, 计算机视觉、计算机图形学和摄影测量领域都做过相关的研究. 对无人机图像的真正射影像处理的关键技术环节包括遮挡检测与补偿、影像镶嵌匀光^[2]. 目前主流的正射纠正中遮蔽检测的方法大多属于 Z-buffer 检测方法^[3~8], Habid 等^[3]给出了一种在矢量建筑物模型下, 通过检测扫描方向上投影射线与天底方向夹角变化趋势的基于角度的检测方法. Zhou 等^[5]给出了一种基于三角网的遮蔽检测方法. Rau 等^[7]给出了一种比较地物点到摄影中心距离的 Z-buffer 方法. Amhar 等^[8]结合数字表面模型 (DBM) 和传统数字地形模型 (DTM) 对遮蔽进行检测, 其主要

引用格式: 郭复胜, 高伟, 胡占义. 无人机图像全自动生成大比例尺真正射影像方法. 中国科学: 信息科学, 2013, 43: 1383–1397, doi: 10.1360/112012-591

思路是采用数字地表模型的可见性分析, 以检测遮蔽地区, 并将结果用于引导后续的几何纠正, 以避免重复纠正而产生的影像叠加现象. 这些方法主要存在的问题在于由于 DSM 与影像分辨率的差异和地形陡变等引起遮蔽的伪可见性噪声, 使得无法准确获得真实的可见性. 而对影像匀光, 主要算法有两大类: 基于灰度空间的 Blending 算法^[9,10] 和基于图像梯度的算法^[11,12]. 文献^[13,14] 则在进行匀光前, 通过一种基于 MRF 的能量优化算法先为正射影像面片选择合理的图像来源, 然后通过灰度空间加权的 Blending 算法或通过约束面片边缘梯度变化的泊松方程 (Poisson equation) 进而解决纹理一致性的问题. 本文生成的正射影像在概念上应该属于真正射影像, 因此后文无特别说明, 不再区分真正射影像和正射影像的概念, 指的都是真正射影像.

另外, 随着计算机视觉领域中特征点检测和匹配算法^[15,16]、鲁棒性估计算法^[17,18]、自标定算法^[19~21]、运动恢复结构重建算法 SfM (structure from motion)^[22,23] 以及多视图立体视觉算法 MVS (multi-view stereopsis)^[24,25] 等技术的不断进步和完善, 基于图像的自动三维重建技术有了突破性的发展. 这些成果已经引起了摄影测量领域研究人员的重视, 并成功地应用在了诸如植被环境遥感^[26]、考古学^[27,28] 等方面的研究. 针对无人机图像的快速重建处理^[29~33], Bundler^[22] 结合基于面片多视图立体视觉算法 PMVS (patch-based multi-view stereo)^[25] 是文献中比较常用的算法, 并成功进行了摄像机的标定和点云模型的重建. 如果无人机图像带有诸如 GPS 和 IMU 的辅助信息, 可以参考文献^[34] 的方法对摄像机进行全自动的外参数标定. 有些文献^[1,27] 则是通过一些商业软件^[35,36] 生成正射影像, 但是这些文献对如何生成正射影像的具体算法则少有报道.

本文旨在探索如何将计算机视觉中的全自动三维重建技术应用到由无人机图像生成大比例尺真正射影像的方法. 本文结构如下: 第 2 节给出由无人机图像生成正射影像完整算法流程; 第 3 节给出生成真正射影像的一些预处理过程, 包括由无人机图像重建准稠密点云的方法和数字表面模型生成的方法; 第 4 节详细介绍了真正射影像的生成的关键步骤, 包括图像坐标映射模型、可见性计算、基于 Markov 随机场 MRF (Markov random field) 能量优化的面片选择、匀光处理等; 第 5 节报道了具体的实验与结果分析; 最后, 在第 6 节进行了总结并对需要进一步开展的工作进行了讨论.

2 无人机图像生成正射影像算法流程

给定若干幅无人机图像, 全自动生成真正射影像算法流程图如图 1 所示.

首先利用 SfM 和 MVS 方法全自动生成稠密三维点云, 每一个三维点在图像集合中的可见性关系也一并获取. 然后通过 B 样条拟合方法, 利用重建的三维点云构造规则格网表面模型. 该部分内容将在第 3 节进行详细介绍.

在得到点云的规则表面模型后, 利用三维点云的可见性关系确定规则格网的可见性关系集. 由于每一块规则格网通常可以在多幅图像中可见, 采用 Markov 随机场方法选择并确定最优的一一对应可见性关系. 最后利用该一一一对应可见性关系生成正射影像, 并进行晕色处理, 实现纹理和颜色的一致. 该部分内容将在第 4 节进行介绍.

3 由无人机图像生成数字表面模型

真正射影像生成, 一般需要三维数字表面模型数据的支持. 在测绘领域, 获取大尺度三维数字表

图 1 真正射影像生成算法示意图

Figure 1 The schematic diagram of the true ortho-image generation

面模型是一项昂贵和耗时的工作. 本文的解决方案是由计算机视觉手段从无人机图像全自动重建三维准稠密点云, 然后由点云生成数字表面模型. 本节介绍的是真正射影像的预处理过程, 包括以下两个步骤: 无人机图像的准稠密三维点云重建和由 PMVS 点云生成数字表面模型.

3.1 准稠密三维点云重建

这里无人机图像三维点云重建有两部分核心内容: (1) 输入多幅场景图像, 通过图像匹配和运动恢复结构 SfM 技术, 全自动恢复出相机的拍摄位姿和场景的稀疏几何结构. (2) 通过多视立体视觉 MVS 方法得到场景的准稠密三维点云重建结果. 在视觉领域中广泛应用的增量式重建算法 Bundler 可以完成摄像机投影矩阵的标定, 同时结合 PMVS 算法, 可以得到准稠密的三维点云.

Bundler^[22] 算法首先选择用于重建的初始图像对, 然后通过迭代地增加单幅 (或少量几幅) 图像, 对新的公共匹配点进行三角化和捆绑调整, 并对新加入的图像进行标定. PMVS^[25] 则在 Bundler 标定相机投影矩阵的基础上, 通过特征点的匹配、重建以及空间点的扩散, 最后得到比较稠密、精度相对较高的重建离散点云. 关于 Bundler 和 PMVS 的具体细节, 请见文献 [22, 25].

无人机图像通过上述 Bundler 结合 PMVS 的三维重建, 可以全自动地获取摄像机的投影矩阵和三维点云. PMVS 在输出重建点云的同时, 也给出了该点法向小面片 p 的准可见性 $V(p)$ 和可见性图像集 $V^*(p)$, 具体定义如公式 (1) 和 (2) 所示. 该面片 p 具有中心 $c(p)$, 单位法向量 $n(p)$, 以及参考图像 $R(p)$.

$$V(p) = \{I | n(p) \cdot c(p)O(I) / |c(p)O(I)| > \cos(\tau)\}, \quad (1)$$

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

式中, I 表示图像, $O(I)$ 表示图像光心, $h(\cdot)$ 表示面片 p 在图像 I 和其参考图像上投影的灰度一致性度量, 具体定义为 1 减去归一化互相关值 NCC(normalized cross-correlation).

3.2 由 PMVS 点云生成数字表面模型

PMVS 点云生成数字表面模型的过程主要包括了通过坐标系转换和 B 样条拟合内插等步骤。

为了得到地面的数字表面模型的规则格网信息, 需要首先计算点云所在的地平面或者垂直向上的方向信息, 以建立适应数字表面模型生成的局部坐标系. 假如已经有了可以进行绝对定向的地面控制点或者相机拍照时刻记录的 GPS 的位置信息, 那么可通过相似变换以很简单地确定世界参考坐标系, 则地面可以直接确定. 而在很多情况下, 这些辅助定位信息无法得到 (包括本文研究的无人机图像数据对象). 确定局部坐标系的方法很多, 通过实验验证基于主成分分析 PCA(principal component analysis) 的方法对我们的数据处理简单而且有效, 下面介绍该方法.

将地面近似为一个平面, 当点云基本位于一个平面上时, 从原理上说 PCA 的第一主成分应该对应该地平面. 将 PMVS 重建的三维点云作为样本点构造协方差矩阵, 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. 然后将离散点变换到地平面为 XY 平面, 垂直向上为 Z 轴的坐标系上. 重建点在坐标系变换前的 X 和变换后 X_{PCA} 的关系可以记为

$$X_{PCA} = T \times (X - \bar{X}), \quad (3)$$

其中, T 为样本 X 的协方差矩阵的特征值从大到小排列后对应特征向量构成的变换矩阵, $\bar{X}$ 为离散三维点的均值.

另一方面, PMVS 获得的点虽然很稠密, 但不是规则点. 通过 B 样条拟合, 然后内插可以获得规则格网上点的坐标, 即得到了数字地表模型 DSM. 多层次 B 样条拟合 MBA(multilevel B-spline algorithm)^[37] 使用由粗到精的控制顶点的分级策略, 各级控制顶点间距由上一级控制顶点间距二分得到, 从而产生多层次三维点云拟合逼近函数, 可以很好地兼顾函数光滑性与逼近精度问题. 本文使用了均匀三次 B 样条函数进行了曲面拟合, 每个控制格网的点 P_t 由 P_t 周围的 16 个控制顶点决定, 完成由 PMVS 点生成规则表面模型.

4 真正射影像的生成

在得到点云的规则表面模型后, 首先需要获取格网坐标以及在原始图像间的可见性关系, 然后利用可见性关系将中心投影下获得的图像映射为格网对应的正射投影的图像, 最后选择合适的格网面片, 对最终整个正射图像进行匀色, 消除不同面片图像之间的纹理不一致现象. 在得到规则表面模型情况下, 下面分别介绍生成真正射影像算法的各个关键步骤.

4.1 正射影像坐标模型映射

介绍正射影像生成方法之前, 我们需要介绍正射影像某个面片到原始图像映射过程的坐标变换关系. 从原始图像到正射影像生成过程中, 涉及 4 个坐标系, 分别是原始图像坐标系、重建点云坐标系、地面局部坐标系和正射影像坐标系, 它们的关系如图 2 所示. 我们最终要得到的是正射影像各像素点 x_{rec} 与原始图像点坐标 x_{org} 之间的映射关系 $x_{rec} = \varphi(x_{org})$.

正射影像图像点 x_{rec} 与对应的数字表面模型上点 X_{dsm} 之间存在相似变换的关系, 计算方法如公式 (4) 所示.

$$x_{rec} = \lambda(X_{dsm} - X_{dsm_min}), \quad (4)$$

图 2 真正射影像映射过程的坐标系转换关系

Figure 2 The involved coordinate system transformations during the true ortho-image generation. (a) The involved coordinate system transformations; (b) the coordinate system transformation flow chart of the true ortho-image generation

其中, λ 为数字表面模型与正射影像像素间的比例关系, 该值可通过正射影像的分辨率来确定, $X_{\text{dsm_min}}$ 为数字表面模型的最小值.

而数字表面模型上的点 X_{dsm} 与重建三维点 X 之间的关系可以通过 3.2 小节的公式 (3) 及 B 样条内插确定, 需要说明的是, 其中由 X_{PCA} 计算重建三维点 X 需要由公式 (3) 求取反变换得到.

最后重建三维点 X 与原始图像点 x_{org} 的关系, 可以由投影关系确定 $x_{\text{org}} = \text{Proj} \times X$, 其中投影矩阵 Proj 可以直接由摄像机内外参数的标定结果得到, 这里假定已经通过第 3 节介绍的 SfM 和 PMVS 方法得到, 这里可以认为是已知量.

4.2 规则格网可见性计算

关于规则格网的可见性计算, 传统摄影测量中的一般做法是利用数字表面模型采用 Z-buffer 算法来确定格网的可见性信息. 由于获取的无人机图像并不能保证光心方向与地面严格垂直, 部分图像存在倾斜, 特别在城区获取的图像, 其三维重建的点云与图像对应物体间存在一定的遮挡关系, 因此需要首先准确地确定三维重建的点云在原始图像上的可见性图像集, 然后借助重建点云的可见性来推测规则格网的可见性. 需要指出的是, 规则格网 Gd_i 和正射影像上的一个面片 P_i 只有一一对应关系, 所以正射影像面片 P_i 在原始图像上的可见性关系和对应的规则格网可见性 Gd_i 是一致的.

正像 3.1 小节所述, 得到的 PMVS 点云结果中, 同时给出了可见性图像集, 利用该信息可以辅助完成规则格网可见性计算. 为了计算规则格网 Gd_i 映射后的正射图像面片 P_i 与原始图像 I_k 之间的可见性关系, 令 $V = \text{vis}(P_i)$ 表示正射图像的面片 P_i 在原始图像集中的可见性列表, 其中, $\text{vis}(P_i) = \text{vis}(Gd_i)$. 下面需要区分格网 Gd_i 内是否包含重建的三维点两种情况来分别讨论面片 P_i 的可见性, 这里“格网内包含重建的三维点”的含义是指三维重建点投影到主平面上的平面点落在格网在主平面上的投影形成的二维格网内.

图 3 图像映射过程可见性标记集

Figure 3 The candidate label sets of visible images of the final ortho-image

如果 Gd_i 中包含了一个或多个原始重建的 PMVS 离散三维点 $Pt_k (k = 1, \dots, m)$, 由于每一个重建三维点 Pt_k 都包含一个可见性图像集 $V^*(Pt_k)$, 则 $\text{vis}(P_i) = \bigcup_{k=1}^m V^*(Pt_k)$, 即取所有点在原始图像可见性的并集. 最后验证格网 Gd_i 四个顶点是否均满足投影到可见图像 I_j 的边界内, 如果任意一个投影点超越图像边界, 则将 I_j 从 $\text{vis}(P_i)$ 中去除. 若经过图像越界检查后, $\text{vis}(P_i)$ 为空, 则按 Gd_i 未包含任何重建的离散三维点方法处理.

如果 Gd_i 未包含任何重建的离散三维点, 则需要通过搜索的方法得到平面坐标与格网 Gd_i 中最近的重建三维点 Pt' , 并令 $\text{vis}(P_i) = V^*(Pt')$, 并按照前面同样的方法, 检验格网顶点在该三维点的可见性图像列表的投影是否越界, 以确定该格网最终的可见性列表.

4.3 基于 MRF 能量优化的规则格网面片选择

一个规则格网对应了正射影像的一个面片, 但一个面片 (格网) 可能在多个原始图像中可见, 这样就面临着最终的正射影像面片应该从哪个图像中选择的问题. 选择面片的目标是能让正射影像的各个相邻格网面片纹理尽量一致、均匀 (当然最好是某一区域的面片从同一幅图像中选择出来的), 同时, 选择的纹理最好能保证清晰、丰富. 据此, 我们可以指定一个面片选择的优化函数. 参考文献 [14] 的方法, 我们也采用利用 MRF 场能量优化的方法对面片进行了选择.

首先, 构造网络图 G , 如图 3 所示, 将每个格网 Gd_i 对应于正射影像图的面片 P_i 作为顶点 V , 将与相邻格网对应顶点连接构成一条边 E , 从而构造图 $G(V, E)$, 其中 V 为顶点集, E 为边集. 每个顶点包含了 k 个元素的标记 $\{l_i\}$, 该集合的元素为面片对应的格网在原始图像中可见的图像序号, 即 $l_i \in \{1, 2, \dots, N\}$, N 为原始图像的总数.

为了减少内存需求, 我们首先对每个格网的可见性按照小面片质量进行了排序, 同一格网 Gd_i 对应的不同小面片图像内图像梯度和越大, 图像质量越好. 若可见性个数大于 k , 则取出其前 k 个序号,

如果不足 k , 则以 -1 作为特殊的标记补足.

这样, 我们可以建立 MRF 框架下的目标能量函数, 通过最小化该能量函数, 得到各格网最优的标记号

$$\min_{l_1, l_2, \dots, l_N} \sum_{i=1}^N E_{\text{data}}(l_i) + \lambda \sum_{(i,j) \in M} E_{\text{smooth}}(l_i, l_j), \quad (5)$$

其中数据项定义为

$$E_{\text{data}}(l_i) = \begin{cases} -\int_{P_i} \|\nabla(I_{l_i}(\phi_{l_i}(p)))\|^2 dp, & l_i \neq -1, \\ +\infty, & l_i = -1, \end{cases} \quad (6)$$

式中, I_{l_i} 为第 l_i 幅原始图像, $\phi_{l_i}(p)$ 表示面片 P_i 中像素 p 到第 l_i 幅原始图像的映射结果, 映射模型见 4.1 小节介绍. 数据项的本质是计算原始图像映射到正射影像小面片内的图像梯度值的和. 正像文献 [14] 所指出的那样, 梯度和可以同时衡量格网法线方向与光线射线之间的夹角、图像模糊程度、图像对比度和图像分辨率等信息, 因此, 数据项可以用梯度和来定义. 其中取负值是为了梯度和最大最优.

平滑项定义为

$$E_{\text{smooth}}(l_i, l_j) = \begin{cases} \int_{N(P_i, P_j)} \|I_{l_i}(\phi_{l_i}(p)) - I_{l_j}(\phi_{l_j}(p))\|^2 dp, & l_i \neq l_j, \\ 0, & l_i = l_j, \end{cases} \quad (7)$$

式中, $N(P_i, P_j)$ 为相邻面片 P_i 和 P_j 之间的公共边. 该项计算了不同原始图像映射到规则网格边界的像素差值, 如果该边相邻的两格网映射到同一原始图像, 则平滑项赋值为 0. 该约束旨在使得相邻格网在可见性的图像集中选择图像时, 尽可能地选择相同序号的图像, 或者在边界处变化不大的图像. λ 值为数据项和平滑项的权重值.

由于构造的图 G 中各顶点的标记集包含的元素不一致, 即各个格网对应的图像可见性序号不同, 所以能量函数不满足 Graph Cuts 的 regularity 条件, 求解这样的能量极小值可以用 BP(belief propagation)^[38] 或 TRW-S (sequential tree-reweighted message passing)^[39] 方法, 本文采用了 BP 来求解该优化模型.

4.4 正射影像匀光处理

通过 MRF 能量优化后, 可以在一定程度上减少不同图像映射到正射影像带来的纹理不一致现象, 但是并不能完全消除不同光照带来的拼接痕迹, 特别是在格网的边界地方, 会有明显的拼接缝出现. 对此, 本文直接采用文献 [12] 的 Poission 方法进行了匀光处理. 该方法基于图像梯度场, 在图像空间以图像梯度差值实现了全局匀光, 既保持了目标区域 Ω 中的原有梯度变化, 同时又最小化了目标图像 I 边界 $\partial\Omega$ 的灰度差异^[40]. Poission 方法使用文献 [40] 的目标函数, 如 (8) 式所示.

$$\min \iint_{\Omega} \|\nabla I - \nabla I_s\|^2 dx dy, \quad (8)$$

$$I|_{\partial\Omega} = I^*|_{\partial\Omega},$$

式中, $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$ 为梯度算子, $\|\cdot\|$ 是向量 $\cdot$ 的 L_2 范数, I 为待处理图像, I_s 为源图像, I^* 为目标图像, ∇I_s 为源图像梯度场, $I|_{\partial\Omega}$ 和 $I^*|_{\partial\Omega}$ 分别表示待处理图像与目标图像边界 $\partial\Omega$ 区域的灰度.

图 4 确山地形数据 PMVS 重建结果和其中一幅图像

Figure 4 The PMVS reconstruction point cloud from QueShan terrain images and one of the original images. (a) The camera calibration information; (b) the sparse points of Bundler; (c) reconstructed PMVS points; (d) one representative image

5 实验结果及结果分析

本文对多种低空无人机系统数据进行了实验, 包括灾害测量系统、资源调查系统以及部分环境调查系统, 采用了多种不同小型固定翼无人机的飞行平台, 无人机飞行高度为 200~2000 m. 飞行前, 由航线规划软件可以对拍摄的航向重叠和旁向重叠进行设置, 数据预设的航向重叠和旁向重叠都很高. 为了保证所获取的图像不模糊, 相机的曝光时间在 1/200 s 以下. 我们对多组数据进行了实验, 以验证本文算法的稳定性和有效性. 这里仅给出一组城区数据, 一组地形数据结果以及结果分析, 其他几组数据的情况类似. 同时, 对这两组数据, 我们将结果与微软拼图软件 Microsoft ICE 和同样可以全自动生成正射影像的商业软件 Agisoft^[35] 进行了对比.

5.1 实验结果

5.1.1 野外地形影像数据

地形数据正射影像生成的实验我们采用了河南确山的数据, 该组数据一共获取 212 幅图像, 飞行高度在 500 m 左右, 航向重叠 > 80%, 旁向重叠 > 30%. 获取图像的相机类型为理光 GR DIGITAL2, 像素为 3648×2736, 像元大小为 2.0233 μm, 挂载的镜头是 6 mm 的定焦镜头. 如图 4 所示, 图 4(a) 给出了标定的摄像机轨迹信息, 图 4(b) 是 Bundler 输出的稀疏重建点结果, 图 4(c) 是重建的 PMVS 点云结果, 图 4(d) 是原图像中一幅具有代表性的图像. 在图 4(c) 中, 由 PMVS 重建出来的非规则三维点个数为 961267 个.

按照第 4.1 小节的方法, 我们通过主成分分析对局部坐标系进行了确定, 同时利用 B 样条拟合的手段, 得到规则格网上点云的内插坐标. 在进行 PCA 局部坐标系确定之前, 我们首先对孤立点云进行了简单处理, 只有当点坐标满足公式 (9) 时, 才参与估计主平面计算.

$$|z - \bar{z}| < 3 \text{ var}(z), \quad (9)$$

图 5 对非规则点云进行 PCA 坐标变换以及 B 样条内插后结果

Figure 5 The points results after the PCA transform and B-spline interpolation. (a) The partial detail of irregular points after PCA transformation; (b) the regular points after B-spline fitting and interpolation; (c) the regular grid after B-spline fitting and interpolation; (d) the digital model from points after B-spline fitting and interpolation

图 6 规则格网可见性个数示意图

Figure 6 The visible image number of the DSM grid. (a) The visible image number of the DSM grid; (b) the label image indicating that the regular grid is visual or not in some image; (c) the schematic diagram of the grids which need to find the visible image label by nearest grid; (d) the schematic diagram of the closed grid image and the boundary of the grid image

式中, z 为局部坐标系下的高程坐标, $\bar{z}$ 为 z 坐标均值, $\text{var}(z)$ 为 z 坐标方差.

图 5 为对无规则点云进行 PCA 坐标变换以及 B 样条内插后的结果, 其中, 图 5(a) 是 PCA 坐标变换后的局部点云放大图, 图 5(b) 为 B 样条内插后的规则点云, 图 5(c) 为 B 样条拟合及内插后的规则格网, 图 5(d) 为 B 样条拟合的点云表面模型.

当完成非规则点内插后, 可以按照第 4.2 小节方法计算规则格网的可见性图像. 我们将格网表示成图像的形式, 每个规则格网视为图像的一个像素, 该格网的可见性图像个数视为该像素的灰度值, 这样就构成了图 6(a) 的规则格网可见性图像个数示意图. 同时, 我们将有可见性图像的格网标示为 1,

图 7 匀色后的正射影像及中间过程的局部图

Figure 7 The final ortho-image result and the intermediate results of part of the image. (a) The final ortho-image result; (b) the partial enlarged detail of the ortho-image without patch selection by MRF; (c) the partial enlarged detail of the ortho-image with patch selection by MRF; (d) the partial enlarged detail of the ortho-image with fusion by Poisson

没有可见性图像的格网标示为 0, 形成了图 6(b) 的规则格网是否具有可见图像的二值标记图. 这里需要同时确认封闭格网的边界以及边界内孔洞的可见性图像, 我们对图 6(b) 的结果采用形态学的膨胀操作和 floodfill 填充方法进行了空洞填补, 获得了图 6(c) 需要通过近邻获取可见性图像的格网示意图和图 6(d) 封闭格网及边界.

最后, 按照第 4.4 小节的方法, 可以得到最终生成的匀色后的正射影像图 7(a), 其中, 图 7(b) 显示的是未选择面片的正射影像, 图 7(c) 为基于 MRF 能量优化选择面片的正射影像, 图 7(d) 为经过泊松匀色后的正射影像. 图 7(b), (c) 和 (d) 分别显示的是图 7(a) 中红色框对应的同一局部区域放大的结果. 由肉眼观察可知, 经过 MRF 能量优化方法选择面片后, 相邻面片尽可能的从同一幅图像中进行了选择, 经过泊松匀色后, 正射影像的整体梯度达到了一致, 提高了视觉效果.

5.1.2 城区影像数据

相比于地形数据, 城区影像由于立面遮挡、地面拓扑结构复杂等因素, 城区的真正射影像制作更加困难. 在没有其他任何辅助信息的情况下, 将我们的算法在城区进行了实验, 也取得了比较理想的结果.

城区数据正射影像生成实验我们采用了河南洛阳体育馆上空利用飞艇拍摄的数据, 该组数据共有 207 幅图像, 飞行高度在 500 m 左右, 由于飞艇上没有稳定平台, 所拍摄的图像倾角很大. 获取图像的相机类型为富士 FinePix S9600, 像素为 3488×2616 , 像元大小为 $2.2 \mu\text{m}$, 挂载的镜头是 28 mm 的定焦镜头, 曝光时间小于 $1/300 \text{ s}$.

图 8 为得到的正射影像结果, 其中, 图 8(a) 是一幅原始图像, 由图 8(a) 可以看出原始影像的倾角比较大, 楼房的立面遮挡比较严重. 图 8(b) 为 PMVS 重建的点云, 图 8(c) 为 B 样条拟合内插后的规则格网, 图 8(d) 为最终生成的城区正射影像全图. 正射影像图的结果仅显示了建筑物的顶面, 表明确定由垂直投影得到, 是比较理想的正射影像, 图 8(e) 给出了图 8(d) 中红色框区域对应的局部放大图, 图 8(f) 为贴上正射影像纹理后的三维模型结果.

图 8 城区正射影像生成结果

Figure 8 The final ortho-image result from urban area images. (a) One representative image; (b) reconstructed PMVS points; (c) the regular grid result after B spline fitting and interpolation; (d) the final ortho-image result from urban area images; (e) the partial enlarged detail of the ortho-image result; (f) the 3D model with the texture mapping

5.2 城区真正射影像结果与商业软件的结果比较

为了验证本文算法的有效性, 我们将结果与微软全自动拼图软件 Microsoft ICE 和可以全自动生成真正射影像的商业软件 Agisoft^[35] 进行了对比. 由于地形图像没有明显的突出物特征, 且高程很少突变, 特别是经过泊松匀色后, 边界难以区分, 所以地形区域不适宜作为进行比较的场景. 相反, 城区影像中存在高楼等, 且城市建筑物高程突变较多, 所以利用城区图像可以很好地检验真正射影像生成的算法好坏. 本文对洛阳体院馆影像的结果进行了对比, 图 9 分别给出了洛阳体育馆图像的微软拼图软件全自动拼接的结果和 Agisoft 生成的正射影像结果.

将图 9(b) 和图 9(d) 的结果与图 8(e) 的结果对比, 红色框显示的是同一区域的局部细节结果, 可以发现本文的算法对该组数据具有一定的优越性. 当然, 微软全自动拼图软件 Microsoft ICE 是拼图软件, 并且中间过程并未生成三维点云和三维模型, 所以结果并不能算是真正射影像. 而从 Agisoft 生成的真正射影像中间结果可以推测其生成真正射影像的思路和本文很相似, 由于商业软件保密等原因, 其算法实现的一些细节无法知道. 分析我们算法优于该商业软件可能的原因是 Agisoft 得到的重建的点云结果不如 PMVS 得到的点云理想, 并且图像可见性的选取不如本文的算法. 这些对比实验, 也进一步验证了本文算法的有效性.

5.3 结果分析与讨论

根据文献 [41] 对数字正射影像地图的质量评定理论, 本文对生成的正射影像质量的评价也通过图面 (影像) 质量的评定和数字精度的质量评定两个方面进行了分析. 图面质量的评定主要衡量了影像

图 9 微软全自动拼图软件 ICE 的拼图结果和 Agisoft 生成正射影像结果

Figure 9 The generated mosaic image by Microsoft ICE and the ortho-image by Agisoft. (a) The mosaic image by Microsoft ICE; (b) the partial enlarged detail corresponding to the left red box; (c) the ortho-image result; (d) the partial enlarged detail corresponding to the left red box

的镶嵌线是否密合, 色调是否均匀, 反差是否适中, 影像是否清晰等, 并且此项检查主要依赖于作业人员的目视判断. 本文生成正射影像算法中, 通过 MRF 能量优化进行了规则格网的面片选择, 同时经过晕光处理后, 发现可以较好地消除不同光照带来的拼接痕迹. 数字精度的评价主要衡量平面精度和接边精度, 而平面精度的衡量依赖于加密的控制点或者数字线划图等其他测量结果. 由于我们目前没有获取这些基准数据, 所以还无法对生成的正射影像绝对精度进行定量的衡量. 这里要说明的是, 在本文算法里, 在生成正射影像前, 对图像进行了准稠密的三维重建, 并基于这些三维点云进行了正射影像生成. 我们前期的工作^[34]表明, 这些重建的三维点的精度比较精确, 记录的三维点在图像上的可见性信息也比较准确, 所以我们觉得基于这些三维点云生成的正射影像也可应该具有一定的精度. 当然, 定量精度的评估是一项需要进一步开展的工作.

6 总结与讨论

本文将计算机视觉领域取得巨大成功的 PMVS 点云重建方法应用到无人机图像的真正射影像生成中, 给出了一种全自动生成大比例尺真正射影像的方法, 通过多组无人机数据实验验证了本文方法的有效性. 实验结果表明, 无论是野外地形数据还是城区图像数据, 都能得到比较理想的真正射影像结果. 同时, 我们对洛阳体育馆城区图像生成的真正射影像结果, 与微软的拼图软件生成的拼图结果, 以及由商业软件 Agisoft 得到的正射影响结果进行了比较, 进一步验证了本文算法的有效性. 由于目前还无法获得高精度的地面控制点数据, 因此我们的结果目前只限于通过观察视觉效果来判断正射影像结果的优劣, 还无法对生成的正射影像绝对精度进行定量的衡量. 下一步的工作中, 我们将使用更多数据对本文方法进行测试, 同时, 也希望与测绘部门合作, 能对得到的真正射影像进行绝对精度的定量分析与评估.

参考文献

- 1 Wei F C, Wang H F. Analysis on the generation of the city large-scale digital true orthoimage. ChengShi Jianshe LiLun

- Yan Jiu, 2011, 30: 1-2 [魏富朝, 王洪峰. 城市大比例尺数字真正射影像图制作分析. 城市建设理论研究 (电子版), 2011, 30: 1-2]
- 2 Zhu Q, Yu J. Review of true orthoimage generation methods in urban area. Highl Sci Online, 2011, 4: 1680-1686 [朱庆, 于杰. 城区真正射影像处理方法综述. 中国科技论文在线精品论文, 2011, 4: 1680-1686]
- 3 Habid A, Kim E, Kim C. New methodologies for true orthophoto generation. Photog Eng Remo Sens, 2007, 73: 25-36
- 4 Zhou G, Chen W, Kelmelis J, et al. A comprehensive study on urban true ortho rectification. IEEE Tansac Geosci Remo Sens, 2005, 43: 2138-2147
- 5 Oda K, Lu W, Uchida O, et al. Triangle-based visibility analysis and true orthoimage generation. Int Arch Photog, Remo Sens Spat Inf Sci, 2004, 34: 623-628
- 6 Rau J, Chen N, Chen L. Hidden compensation and shadow enhancement for true orthophoto generation. In: Proceedings of Asian Conference on Remote Sensing, Taipei, 2000. 112-118
- 7 Rau J, Chen N, Chen L. True orthophoto generation of built up areas using multi-view images. Photog Eng Remo Sens, 2002, 68: 581-588
- 8 Amhar F, Josef J, Ries C. The generation of true orthophoto using a 3D building model in conjunction with a conventional DTM. Int J Photog Remo Sens, 1998, 32: 16-22
- 9 Bernardini F, Martin I, Rushmeier H. High quality texture reconstruction from multiple scans. IEEE Trans Visual Comput Graph, 2001, 7: 318-332
- 10 Baumberg A. Blending images for texturing 3D models. In: Proceedings of the British Machine Vision Conference, Cardiff, 2002. 404-413
- 11 Levin A, Zomet A, Peleg S, et al. Seamless image stitching in the gradient domain. In: IEEE European Conference on Computer Vision(ECCV), Prague, 2004. 377-389
- 12 Kazhdan M, Hoppe H. Streaming multigrid for gradient-domain operations on large images. ACM Trans Graph(TOG), 2008, 27: 21
- 13 Victor L, Denis I. Seamless mosaicing of image-based texture maps. In: IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Minneapolis, 2007. 1-6
- 14 Gal R, Wexler Y, Ofek E, et al. Seamless montage for texturing models. Comput Graph Forum, 2010, 29: 479-486
- 15 Lowe D. Distinctive image features from scale-invariant keypoints. Int J Comput Vis(IJCV), 2004, 60: 91-110
- 16 Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features(SURF). Comput Vis Image Underst, 2008, 110: 346-359
- 17 Fischler M, Bolles R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM, 1981, 24: 381-395
- 18 Rousseeuw P, Leroy A. Robust Regression and Outlier Detection. New York: John Wiley & Sons Inc, 1987. 1-313
- 19 Nistér D. An efficient solution to the five-point relative pose problem. IEEE Trans Patt Anal Mach Intel, 2004, 26: 756-770
- 20 Hartley R, Zisserman A. Multiple View Geometry in Computer Vision Second Edition. Cambridge: Cambridge University Press, 2004. 1-672
- 21 Pollefeys M, van Gool L. Visual modeling with a hand-held camera. Int J Comput Vis, 2004, 59: 207-232
- 22 Snavely N, Seitz S, Szeliski R. Photo tourism: exploring photo collections in 3D. ACM Trans Graph(TOG), 2006: 835-846
- 23 Strecha C, Pylvanainen T, Fua P. Dynamic and scalable large scale image reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition(CVPR), San Francisco, 2010. 406-413
- 24 Hiep V, Keriven R, Labatut P, et al. Towards high resolution large-scale multi-view stereo. In: IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Miami, 2009. 1430-1437
- 25 Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis. IEEE Trans Patt Anal Mach Intel, 2010, 32: 1362-1376
- 26 Dandois J, Ellis E. Remote sensing of vegetation structure using computer vision. Remo Sens, 2010, 2: 1157-1176
- 27 Verhoeven G, Doneus M, Briese C, et al. Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. J Arch Sci, 2012, 39: 2060-2070
- 28 Fulvio R, Antonia S, Andrea L, et al. Archaeological site monitoring: UAV photogrammetry can be an answer. In: the XXII Congress of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences(ISPRS), Melbourne, 2012. 583-588

- 29 Darren T, Arko L, Christopher W. An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle(UAV) imagery, based on structure from motion(SfM) Point Clouds. *Remo Sens*, 2012, 4: 1392–1410
- 30 Previtali M, Barazzetti L, Scaioni M. Multi-step and multi-photo matching for accurate 3D reconstruction. In: the XXXVIII Congress of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences(ISPRS), Munich, 2011. 103–108
- 31 Steve H, Arko L. An accuracy assessment of georeferenced point clouds produced via multi-view stereo techniques applied to imagery acquired via unmanned aerial vehicle. In: the XXXIX Congress of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences(ISPRS), 2012. 475–480
- 32 Barazzetti L, Remondino F, Scaioni M, et al. Fully automatic UAV image-based sensor orientation. In: the XXXVIII Congress of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences(ISPRS), Munich, 2011. 1–6
- 33 Erin O, Carl S, David N, et al. Evaluation of image collection requirements for 3D reconstruction using phototourism techniques on sparse overhead data. In: *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, Baltimore, 2012. 8390
- 34 Guo F S, Gao W. Batch reconstruction from UAV images with prior information. *Acta Auto Sin*, 2013, 39: 834–845 [郭复胜, 高伟. 基于辅助信息的无人机图像批处理三维重建方法. *自动化学报*, 2013, 39: 834–845]
- 35 Agisoft LLC, 2011. AgiSoft PhotoScan User Manual. Professional Edition, Version 0.8.0. <http://www.agisoft.ru/pscan/help/en/pscan-pro.pdf>
- 36 LPS. <http://geospatial.intergraph.com/products/LPS/LPS/Details.aspx>
- 37 Oyvind H. Approximation of Scattered Data with Multilevel B-splines. Technical Report STF42 A01011, SINTEF 2001
- 38 Tappen M F, Freeman W T. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In: *Proceedings of the Ninth IEEE International Conference on Computer Vision(ICCV)*, Nice, 2003. 900–907
- 39 Kolmogorov V. Convergent tree-reweighted message passing for energy minimization. *IEEE Trans Patt Anal Mach Intell(PAMI)*, 2006, 28: 1568–1583
- 40 Perez P, Gangnet M, Blake A. Poisson image editing. *ACM Trans Graph*, 2003, 22: 313–318
- 41 Xie H H. Research on digital orthophoto map grade estimation. *Remo Sens Inf*, 2003, 3: 35–40 [谢惠洪. 数字正射影像地图的质量评定. *遥感信息*, 2003, 3: 35–40]

Automatic generation of large scale true ortho-image from UAV images

GUO FuSheng*, GAO Wei & HU ZhanYi

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

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

Abstract With the rapid development of unmanned aerial vehicles (UAV) remote sensing technology, UAV finds its place in various applications where the large-scale orthoimage generation is a popular and important step. This paper is meant to apply hugely successful multi-view 3D reconstruction techniques in recent years in computer vision field to UAV images, and presents a method of automatic generation of large scale true orthoimage based on structure from motion (SfM) and multi-view stereo (MVS) technologies. This paper includes the following key parts: First, an approach is proposed to build a Digital Surface Model (DSM), seemingly a necessary step for any good orthoimage generation, from the points cloud reconstructed by Patch-based Multi-view Stereo (PMVS).

Then the key steps of our orthoimage generation method are presented, such as reference system definition and transformation, the patch visibility computation, Markov Random Field(MRF) model based patch selection and seamless mosaicking. Finally, our orthimage generation method is tested with several sets of UAV images, and compared with some commercial software. The experiments show that our orthimage generation method performs satisfactorily for both terrain and urban area UAV images.

Keywords true orthoimage, UAV, multi-View stereo, PMVS, Markov random field

GUO FuSheng was born in 1982. He received his bachelor degree in 2004 and master degree in 2008 from the PLA Information Engineering University. Currently, he is a Ph.D. candidate at the Institute of Automation, the Chinese Academy of Sciences. His research interest covers 3D reconstruction and photogrammetry.

GAO Wei was born in 1980. He received his Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences at 2008. Currently, he is an associate professor at the Institute of Automation, Chinese Academy of Sciences. His research interest covers computer vision and 3D reconstruction.

HU ZhanYi was born in 1961. He received his B.S. degree in automation from the North China University of Technology in 1985, the Ph.D. degree (Docteur d'Etat) in computer vision from the University of Liege, Belgium, 1993. He is now a professor at the Institute of Automation, Chinese Academy of Sciences.