MULTIRESOLUTION PATCH-BASED DENSE RECONSTRUCTION INTEGRATING MULTIVIEW IMAGES AND LASER POINT CLOUD

Rongchun Zhang [a,b], Xuefeng Yi [c,*], Hao Li [c,*], Lanfa Liu [d,f], Guanming Lu [b], Yuanyuan Chen [e], Xiantao Guo [a]

[a] School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China, 210023 - (rongchunzhang, lugm, xiantaoguo)@njupt.edu.cn
[b] School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, 210023 - rongchunzhang@njupt.edu.cn
[c] School of Earth Sciences and Engineering, Hohai University, Nanjing, China, 211100 - hhuyxf@126.com; lihao@hhu.edu.cn
[d] Hubei Provincial Key Laboratory for Geographical Process Analysis and Simulation, Central China Normal University, Wuhan, China, 430079 - lanfa@ccnu.edu.cn
[e] College of Civil Engineering, Nanjing Forestry University, Nanjing, China, 210037 - cheny@njfu.edu.cn
[f] College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China, 430079 - lanfa@ccnu.edu.cn

Commission II, WG II/2

KEY WORDS: Dense Reconstruction, Point Cloud, Multiview Images, Registration, Integration, PMVS, Octree.

ABSTRACT:

A dense point cloud with rich and realistic texture is generated from multiview images using dense reconstruction algorithms such as Multi View Stereo (MVS). However, its spatial precision depends on the performance of the matching and dense reconstruction algorithms used. Moreover, outliers are usually unavoidable as mismatching of image features. The lidar point cloud lacks texture but performs better spatial precision because it avoids computational errors. This paper proposes a multiresolution patch-based 3D dense reconstruction method based on integrating multiview images and the laser point cloud. A sparse point cloud is firstly generated with multiview images by Structure from Motion (SfM), and then registered with the laser point cloud to establish the mapping relationship between the laser point cloud and multiview images. The laser point cloud is reprojected to multiview images. The corresponding optimal level of the image pyramid is predicted by the distance distribution of projected pixels, which is used as the starting level for patch optimization during dense reconstruction. The laser point cloud is used as stable seed points for patch growth and expansion, and stored by the dynamic octree structure. Subsequently, the corresponding patches are optimized and expanded with the pyramid image to achieve multiscale and multiresolution dense reconstruction. In addition, the octree's spatial index structure facilitates parallel computing with highly efficiency. The experimental results show that the proposed method is superior to the traditional MVS technology in terms of model accuracy and completeness, and have broad application prospects in high-precision 3D modeling of large scenes.

1. INTRODUCTION

Three-dimensional Laser Scanners and digital cameras are top-rated sensor devices in remote sensing mapping, intelligent driving, and smart cities. Multi View Stereo (MVS) (Seitz et al., 2006, Strecha et al., 2008) is an important technology used to generate a dense point cloud through multiview images dense reconstruction. Both Semi-Global Matching (SGM) (Hirschmüller, 2007) and Patch-based Multi-view Stereo (PMVS) (Shen, 2013) are the two popular dense reconstruction methods. The dense point cloud model brings richer texture details and higher resolution, but the model's spatial precision may be affected by various error factors. In comparison with the dense point cloud, lidar point clouds have better spatial precision but lower texture resolution. It is possible to generate a high-quality dense point cloud. Both the precision and texture details can be improved by integrating lidar point cloud and multiview images for dense reconstruction.

The following problems may exist in the dense reconstruction process based on multisource data. The first problem is that the spatial resolution of the point cloud, the viewing angle of the image, and the shape of the target surface all affect the modeling accuracy. For example, the lidar point cloud with higher spatial

resolution can keep the integrity of the dense reconstruction, and the forward-looking perspective for images can keep better co-visibility. The spatial resolution of point clouds collected by different sensor devices will result in different densities of seed points. Because the image acquisition method is flexible and the shooting angles are diverse, the distribution of seed points may vary greatly when the point cloud is reprojected to the image. Object shape differences may also cause laser point clouds to have different projected pixel distributions on different views. An
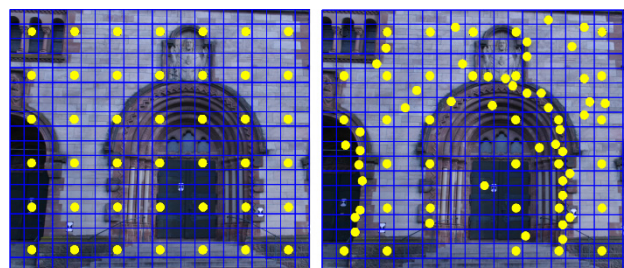

Figure 1. Schematic diagram of the point cloud reprojection on the image with different shooting angles and building shapes. (a) the reprojection for the ideal case; (b) the reprojection for the actual situation.

---

* Corresponding author

example of the above is shown in Figure 1. The yellow points in Figure 1(a) and Figure 1(b) represent the reprojections on the image for the point cloud. The left and right pictures refer to the reprojection distribution under the ideal condition and the practical condition, respectively.

During dense reconstruction, a patch is a rectangular covering the surfaces visible in the input images, which is shown in Figure2. The details of the patch definition and the PMVS method have been explained by Furukawa (Furukawa, 2007).
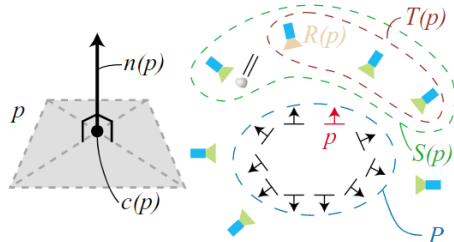
Figure 2. A patch and its associated images (Furukawa, 2007).

The size of the seed patch depends on the image resolution and the spatial resolution of the point cloud. Therefore, it is difficult to choose the optimal growing patch size. Computational efficiency is the second problem. The traditional PMVS algorithm is based on the patch growth and expansion, which cannot be calculated in parallel due to the calculation sequence dependency, resulting in low calculation efficiency. During the modeling process, the data volume of the laser point cloud is generally huge. It makes the PMVS algorithm more demanding on hardware. The computational efficiency is difficult to guarantee.

Locher proposed a progressive prioritized multiview reconstruction method, which can visualize the output point cloud during reconstruction, and the runtime is improved largely. The most contribution of the method is first delivering a dense point cloud using a sparse point cloud generated by SfM as a computational budget in a progressive manner (Locher, 2016). The accuracy of this method depends on the quality of the sparse point cloud. Moreover, the mapping relationship between the sparse point cloud and multiview images has been determined by SfM, while how to achieve dense reconstruction integrating multisource data does not be considered in this method.

Based on the above research and analysis, this paper proposes a multiresolution patch-based 3D dense reconstruction method based on the integration of multiview images and the laser point cloud, which considers the spatial distribution constraints of laser point clouds. In this method, the laser point cloud is used as the seed points, and the octree structure and pyramid image are used for multiscale and multiresolution patch expansion to improve the efficiency and accuracy of dense reconstruction.

The remainder of this paper is organized as follows: Section 2 summarizes the related work; the details of the proposed method are introduced in Section 3; Section 4 arranges the experiments and the corresponding results; followed by conclusions and discussions in Section 5.

2. RELATED WORK

In the previous period, many scholars have studied the extraction of geometric shapes from images, such as textures, shadows, contours, and stereo correspondence. MVS is a method for extracting geometric information using stereo image pairs (Seitz

et al., 2006, Strecha et al., 2008). The quality and the precision of images and the interior and exterior orientation elements of the camera determine the modeling effect of MVS. The development of Structure from Motion (SfM) technology makes the calculation of the interior and exterior orientation elements of the camera easy and simple, and SfM models the geometry of two or more views under strict scene assumptions (Hartley et al., 2000). Carlo Tomasi early presented a technical idea of visual reconstruction algorithms (Carlo et al., 2011). RANSAC (Mach, 1981) allows SfM to robustly estimate the pairwise geometric relationship between two or more views under noise matching. In recent years, SfM and Visual Simultaneous Localization and Mapping (VSLAM) (Karlsson et al., 2005) have been rapidly developed and widely used in the field of urban 3D scene modeling, autonomous driving and indoor navigation, etc. Based on this, MVS algorithms can get better results and are widely used in various industries.

In computer vision, MVS algorithms were initially developed in a laboratory environment (Tsai, 1983, OkutomiM et al., 1993, Faugeras, 1997), where the shooting conditions could be controlled, and the camera could be accurately calibrated. Then, they were used in small outdoor scenes (Strecha et al., 2004, Hornung et al., 2006, Ha Bb Ecke et al., 2007, Sinha et al.,2007, Vogiatzis et al., 2008) and finally, extended to large outdoor scenes (Labatut et al., 2007, Pollefeys et al., 2008, Vu et al., 2009, Furukawa et al., 2010,). Bundler developed by Noah Snavely's (Snavely, 2010) solves the problem of recovering structure from motion (SfM). VisualSFM developed by Changchang Wu (Changchang, 2013) is a GUI application for SfM. MVS software developed by Jancosek (Jancosek et al., 2011) performs well in practical applications. The Multi View Environment (MVE) is a complete end-to-end pipeline implementation for image-based geometry reconstruction developed by TU Darmstadt. Open Multiple View Geometry (OpenMVG) (Moulon et al., 2017) provides customizable tools for sparse reconstruction by SfM in multi-view geometry, such as feature extraction, feature matching, sparse point cloud generation, and so on.

PMVS is an object-based dense matching reconstruction method, while SGM is an image-based dense matching modeling method (Hirschmüller, 2007), and the former has better accuracy while the latter has better efficiency. Moreover, especially for regions with large undulations, PMVS has better performance for aerial images dense reconstruction than SGM. In recent years, MVS modeling methods based on deep learning have achieved better results (Wang et al., 2021, Luo et al., 2019). However, these existing methods all take multiview images as the only data source.

At present, multi-sensor integration and fusion are widely used in various industries, and dense reconstruction methods based on multisource data are still under development (Franceschi M et al., 2015). This paper proposes a PMVS method for the integration of point clouds and images, using laser point clouds with higher precision as the frame and seed points, and using images for patch reconstruction based on point clouds to obtain high-quality 3D models. The main contribution of this work is that using the octree structure to store the laser point cloud. It proceeds multiresolution and multiscale patch expansion and branch according to the octree structure.

3. METHODOLOGY

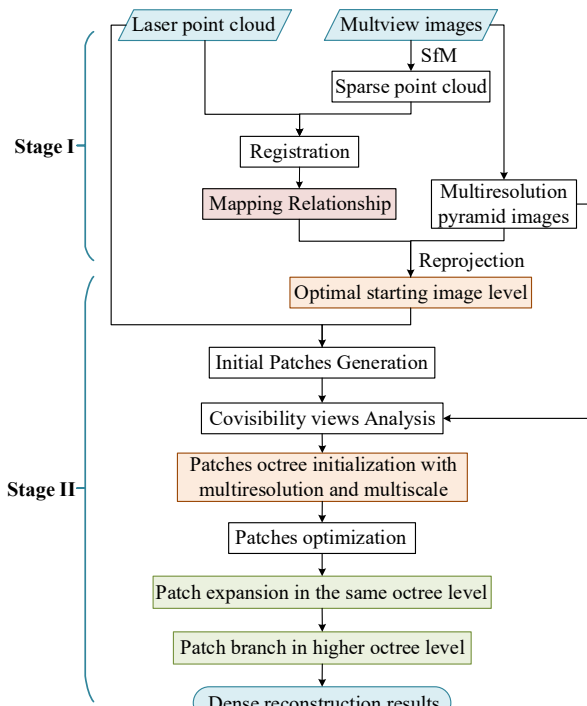The flowchart of the method presented in this paper is shown in Figure 3.

Figure 3. Flowchart of the proposed method.

In the first stage, to obtain the mapping relationship between multiview images and the laser point cloud, an indirect rough-to-fine registration strategy is adopted in this paper, i.e., a sparse point cloud is firstly generated by SfM using multiview images, which is then registered with the laser point cloud to get the correspondence between multiview images and laser point cloud. The buildings have noticeable point features, line features, and planar features. This work uses these features for rough registration. Iterative Closest Points (ICP) refined registration algorithm is proceeded for getting higher accurate results.
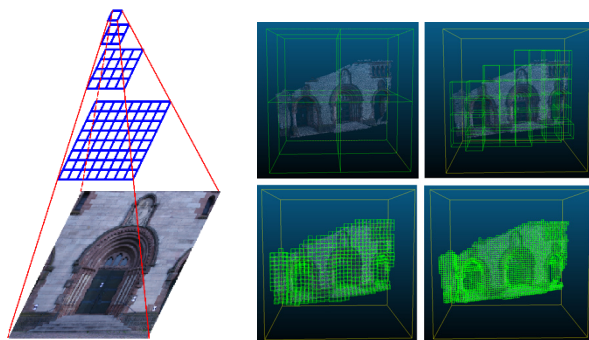

Figure 4. The schematic of pyramid image (left) and octree point cloud (right).

In the second stage, a multiresolution patch-based dense reconstruction is proceeded based on the integration of dynamic octree structure and multiple levels image pyramid, and Figure 4 shows the schematic of pyramid image and octree laser point cloud. The green cubes represent octree cells. Both the pyramid image and laser point cloud structured with octree are used for patch expansion, and the proper correspondence between them needs to be determined first. The definition, expansion, and optimization of patches are crucial steps in such patch-based dense reconstruction techniques, which have a significant effect on the efficiency and accuracy of reconstruction. The point cloud

generated by sparse reconstruction becomes denser as the patch expands, and each patch corresponds to a point in the sparse point cloud. In this paper, the laser point cloud is used as the input of dense reconstruction, and the seed patches for expansion are determined according to the mapping relationship between the laser point cloud and the multiview images. The optimal pyramid image level corresponding to the initial seed patches is predicted by the distance distribution of the projected pixels of the laser point cloud on the multiview images, which is used as the starting image level for multiresolution patch optimization and expansion. The point cloud during dense reconstruction is stored using a dynamic octree structure and dynamically updated with patch optimization and expansion.

3.1 Prediction of the optimal starting level of the image pyramid

Patch expansion and branch should be started from the images with the best resolution matching with the spatial resolution of the laser point cloud. Therefore, a pyramid image is necessary for reconstruction in multiple levels of the octree structure.

In this paper, the optimal starting layer of the pyramid image is determined according to the mapping relationship between the laser point cloud and the multi-view image. Generally, the spatial resolution of the laser point cloud is lower than the image resolution. In this paper, the laser point cloud is reprojected to the pyramid images of multiview images according to the mapping relationship between the laser point cloud and the multi-view image. For laser points which do not be observed in the image, the reprojection is invalid and should be filtered (Azureology, 2022).

According to the distance distribution of the projected pixels of the laser point cloud on the pyramid image, the appropriate pyramid level can be predicted and used as the starting image level for patch optimization and expansion during dense reconstruction. All pyramid levels between the original image and the pyramid starting image level will be used for patch optimization.

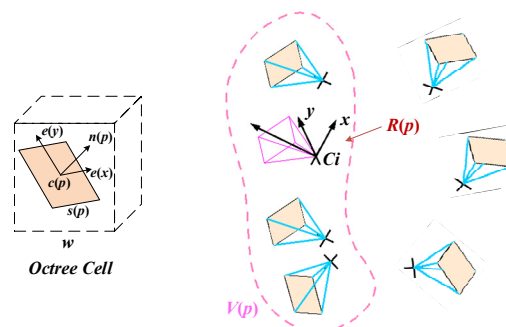### 3.2 Seed patches generation and optimization under laser point cloud reprojection constraints

Figure 5. The schematic of the spatial geometric relationship of patches.

The determination of seed patches initialization is the first step in dense reconstruction. In this paper, the laser point cloud is used as the input, and the corresponding patches are initialized as follows: The number of covisibility views for each point in the laser point cloud can be determined according to the mapping relationship between the laser point cloud and the multiview

images, and those points with more than three covisibility views are taken as candidate seed points and used for patch optimization and expansion, and the remaining points are directly merged with the final results.

Patch optimization determines the reconstruction accuracy. In this paper, Normalized Cross Correlation (NCC) of a patch's projection in the image space $g_I(p)$ is used for patch optimization. The schematic of the spatial geometric relationship between a 3D patch and its corresponding 2D projection patches on the multiview images is shown in Figure 5. Where C_i denotes the position of a 3D point, O_i denotes the camera center.

Let $c(p)$ and $n(p)$ represent the center and the normal vector of the patch, respectively, and they are optimized by maximizing the averaged NCC, i.e., minimizing the $e(p)$ in Equation (1) (Locher, 2016). Here, for each patch, a pixel coordinate system is specified, and the x axis is parallel with the x axis of its corresponding reference image. The initial normal vector $n(p)$ is orthogonal with the coordinate system.

$$e(p) = \frac{1}{|V(p)|-1} \cdot \sum_{I \in V(p) \setminus R(p)} (1 - \langle g_R(p), g_I(p) \rangle) \qquad (1)$$

where $g_R(p)$ denotes the patch's projection in the reference image, $V(p)$ and $R(p)$ represents the visible image set and the reference set of the patch, respectively. The set of $R(p)$ is composed of those patches for which the optical axis and the normal vector are the most similar.
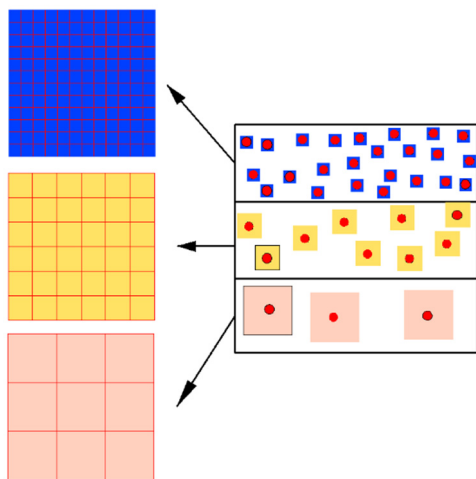
Figure 6. The schematic of the spatial geometric relationship of patches.

The octree levels for patch expansion should match with image resolution to improve the accuracy of patch optimization. In this paper, pyramid images are used to provide 2D projection patches with multiresolution, and the schematic of the correspondence between the image pyramid and octree patch expansion is shown in Figure 6. Considering the patch's scale, the corresponding optimal pyramid level l_I is determined according to Equation (2) (Locher, 2016).

$$l_I = \left\lfloor log_2 \left(\frac{f_{c_i}}{s(p)d_{I_i}} \right) \right\rfloor \qquad (2)$$

where the scale of the patch is denoted by $s(p)$, which is determined according to the relationship between pyramid image resolution and the spatial resolution of the laser point cloud; f_{c_i}

represents the focal length of the image C_i, and d_{I_i} is the corresponding depth. $\lfloor \cdot \rfloor$ indicates rounding integer.

All 2D patches are generated by the candidate 3D seed points projecting in all pyramid image levels between the optimal starting level and the original resolution level. Those multiscale patches through optimization are taken as the final seed patches, and the points of other patches are merged into the final results.

3.3 Seed patches expansion in multiresolution and multiscale space

A dynamic octree structure is used to store point clouds high efficiently in this paper. Both the octree structure and pyramid images are integrated to perform the multiresolution and multiscale patch expansion, which mainly includes patch expansion in the same level and patch branch in the higher level (shown in Figure 7).
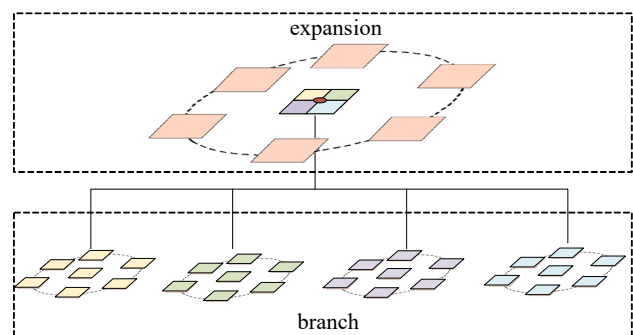
Figure 7. The schematic of the expansion and branch for the patch.

(1) Expansion in the same level: all patches in the same octree level are sampled along a circle with radius R_1 (equal to 0.9 times the width of the cell in the current octree level in this paper) to generate n new patches (n is often set to 6 or 8). If the new patch center locates in another octree cell and the octree cell has not been processed, the patch will be further optimized to determine whether it should be expanded. The expansion stops when there is only one patch in every node of the octree level.

(2) Patch branch in the higher level: when all the patches in the same octree level have been processed, they will be branched into several smaller patches $m \times m$ (this paper takes 5×5). Similar to expansion in the same level, each small patch is expanded along a circle with a smaller radius of R_2 to generate several new patches. If the new patch center shares the same octree node with the old patch center, it will be optimized and branched. The patch branch stops until the point cloud has the same resolution as the original multiview images.

In addition, the laser point cloud is used for patch optimization and expansion, and the octree structure is used as the storage framework, both of which can effectively avoid the wrong expansion of outliers with traditional methods.

4. EXPERIMENTS AND RESULTS

In this paper, both multiview images and laser point cloud data of the Tsinghua School of Tsinghua University are used for experiments (National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, http://vision.ia.ac.cn/data). Multiview images for experiments are listed in Figure 8.

Figure 8. The thumbnail of multiview images.

A sparse point cloud with 6,198 points is first generated by visualSFM (Wu, 2011, Wu et al., 2013, Wu, 2007), which is shown in Figure 9(a). Figure 9(b) shows the laser point cloud with 30,714 points. It can be seen from Figure 9 that the point cloud has been textured by reprojecting on multiview images according to the mapping relationship between the laser point cloud and multiview images.


(a)

(b)
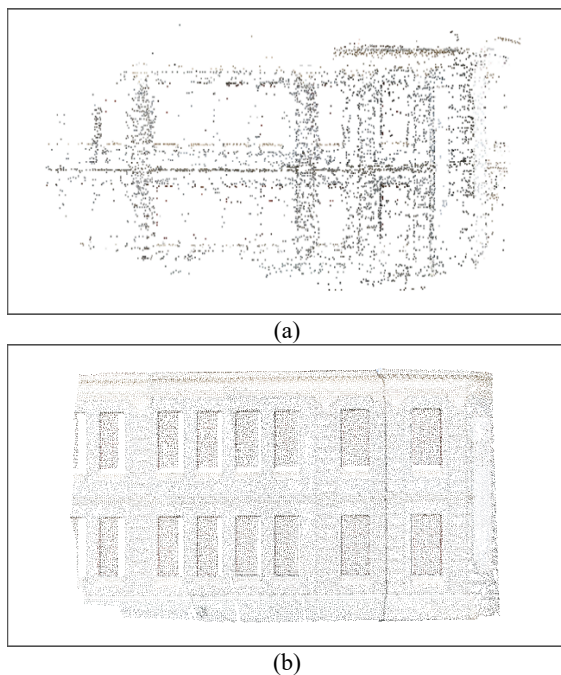
Figure 9. The input point clouds for dense reconstruction: (a) the sparse point cloud from VisualSFM; (b) the laser point cloud.

Figure 10(a) shows the point cloud generated by traditional patch initial optimization (Locher, A. et al., 2016), and Figure 10(b) shows the point cloud generated by initial patch optimization using the proposed method. The number of the two point clouds are 11,224 points and 19,772 points, respectively. It is easy to see that the proposed method preserves more seed points as it considers multiscale patches. While the traditional method filters more useful laser points during initial patch optimization. Here, the points with less than three covisibility views are not included in Figure 10, which will be merged into the final results.


(a)

(b)
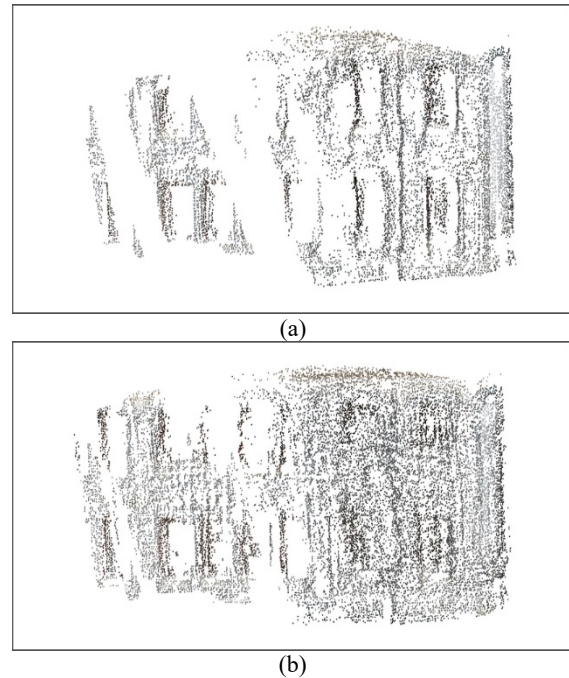
Figure 10. The point clouds generated by initial patch optimization: (a) the traditional method proposed by Locher (Locher, 2016); (b) the proposed method proposed in this work.

(a)

(b)
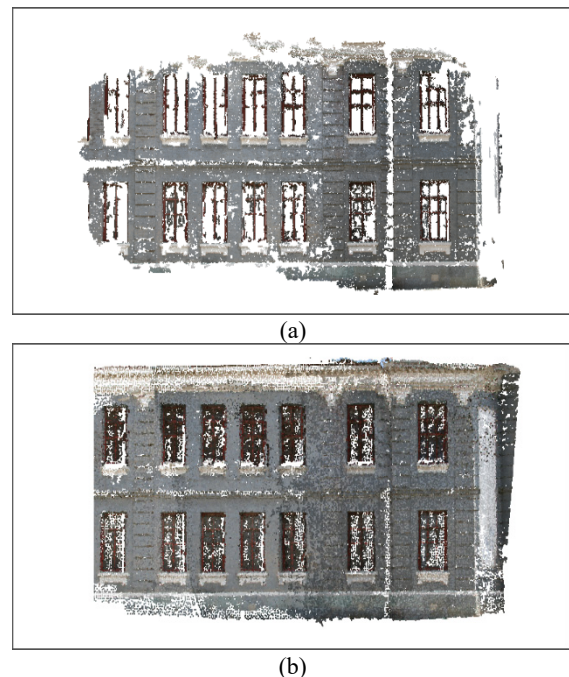
Figure 11. The comparison of the dense reconstruction results: (a) visualSFM+ PMVS; (b) the proposed method.

The dense reconstruction results with the traditional method (VisualSFM + PMVS) and the proposed method are shown in Figures 11(a) and (b), respectively. In Figure 11(a), there are totally 566,256 points generated by the traditional method. 888,451 points are generated using the proposed method, which is shown in Figure 11(b).

A denser laser point cloud is used to evaluate the accuracy and completeness of the proposed method. Figure 12 shows the denser laser point cloud with 206,551 points.
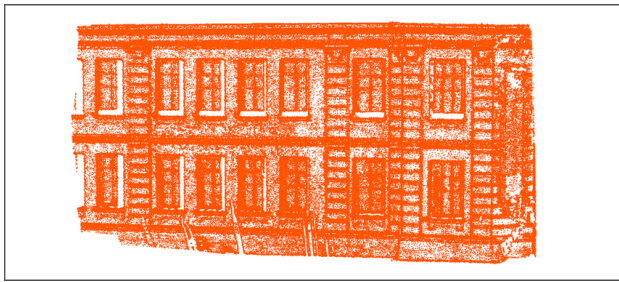
Figure 12. The denser laser point cloud for evaluation.

Absolute distances between the laser point cloud and the generated dense point clouds are compared and analyzed, which are shown in Figure 13. Figure 13(a) shows the comparing results of the final dense point cloud expanding by the proposed method. To make a fair comparison, it doesn't consider the initial laser point cloud merged here. It can be concluded that 84.92% and 75.22% of the point-to-point distances for the proposed method and the traditional method are less than 0.03 meters, respectively.

(a)

(b)
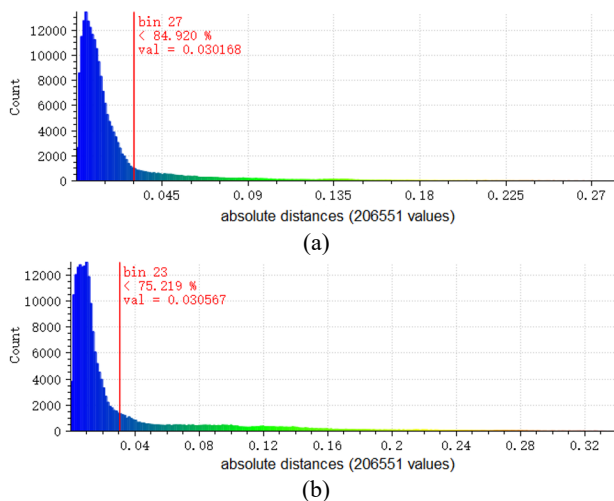
Figure 13. The absolute distances comparison between the original laser point cloud and the dense point clouds respectively generated by the traditional method proposed by Locher and by this work: (a) comparing to the dense point cloud from the proposed method (not including the initial laser point cloud); (b) comparing to the dense point cloud from the traditional method.

The comparison between the traditional method and the proposed method clearly shows that the latter model performs better in terms of completeness and details.

5. CONCLUSIONS

This paper proposes a multiresolution and multiscale patch-based 3D dense reconstruction method using multiview images and considering the spatial distribution of laser point clouds. The proposed method establishes the correspondence between laser point clouds and multiview images through an indirect registration pipeline. The laser point cloud is used as stable seed points, and stored with an octree structure, projections of which on different image pyramid levels constitutes multiscale optimal patches. These patches are optimized and expanded to generate dense points cloud.

The method proposed in this paper has the following advantages: (1) The method selects the laser point cloud with better integrity and higher precision as the seed point, and combines the pyramid image to generate multiscale patches. The completeness and sophistication of the model are continuously improved as patches are optimized and extended. (2) The method predicts the optimal starting layer of the pyramid image, and selects multiscale seed patches to participate in optimization and expansion. Therefore, more laser seed points can be reserved to obtain more patches for expansion. (3) Different octree levels correspond to different spatial resolutions of point clouds. Combining octree levels with image pyramids for dense reconstruction can realize the hierarchical parallel expansion of point clouds, which simultaneously considers both the integrity and local details of the model. (4) For huge laser point clouds, computational efficiency is an essential issue for intensive reconstruction based on multisource data. The method adopts the octree space index, which can realize parallel computing, i.e., patch expansion in different octree nodes is performed simultaneously. Therefore, the computing efficiency improves significantly.

The comparison experiment with the traditional image-based PMVS method shows that the number of initial seed points generated by the proposed method has been significantly improved because the multi-layer scale is considered, and the point cloud model generated by the proposed method has better integrity and higher precision. The method proposed in this paper has practical value in fine 3D modeling of large scenes.

REFERENCES

Azureology, 2022. [Source code]. https://github.com/azureology /kitti-velo2cam

Carlo, Tomasi, 2011. Visual reconstruction: technical perspective. *Communications of the ACM.*

Faugeras, O., 1997. Level set methods and the stereo problem. *Springer, Berlin, Heidelberg.*
Franceschi M , Martinelli M , Gislimberti L , et al., 2015. Integration of 3D modeling, aerial LiDAR and photogrammetry to study a synsedimentary structure in the Early Jurassic Calcari Grigi (Southern Alps, Italy). *European Journal of Remote Sensing*, 48:527-539.

Furukawa, Y., Ponce, J., 2010. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 32(8), 1362-1376.

H Hirschmüller., 2007. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30.*

Ha Bb Ecke, M., Kobbelt, L., 2007. A surface-growing approach to multi-view stereo reconstruction. *IEEE.* IEEE.

Hartley, R., Zisserman, A., 2000. Multiple view geometry in computer vision. *Cambridge University Press.* Cambridge University Press.

Hornung, A., Kobbelt, L., 2006. Hierarchical Volumetric Multi-view Stereo Reconstruction of Manifold Surfaces based on Dual

Graph Embedding. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition.* IEEE.

Jancosek, M., Pajdla, T., 2011. Multi-view reconstruction preserving weakly-supported surfaces. *CVPR 2011.* IEEE.

Karlsson N , Bernardo E D , Ostrowski J , et al, 2005. The vSLAM Algorithm for Robust Localization and Mapping. *International Conference on Robotics and Automation Barcelon. IEEE.*

Labatut, P., Pons, J. P., Keriven, R., 2007. Efficient Multi-View Reconstruction of Large-Scale Scenes using Interest Points, Delaunay Triangulation and Graph Cuts. *IEEE International Conference on Computer Vision.* IEEE.

Locher, A., Perdoch, M., Gool, L.V., 2016. Progressive Prioritized Multi-view Stereo. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE.

Luo, K., Guan, T., Ju, L., Huang, H., Luo, Y., 2019. P-mvsnet: Learning patch-wise matching confidence aggregation for multi-view stereo. *In Proceedings of the IEEE/CVF International Conference on Computer Vision,* 10452-10461.

Mach, C., 1981. Random Sample Consensus: a paradigm for model fitting with application to image analysis and automated cartography.

Multi-view environment. *http://www.gris.informatik. tudarmstadt.de/projects/multiview-environment.*

Moulon, P., Monasse, P., Perrot, R., Marlet, R.,2017. OpenMVG: Open Multiple View Geometry. *International Workshop on Reproducible Research in Pattern Recognition.* Springer, Cham. https://github.com/openMVG/ openMVG.

National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, http://vision.ia.ac.cn/data.

OkutomiM., KanadeT., 1993. A multiple-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

Pollefeys, M., Nister, D., Frahm, J. -M., Akbarzadeh, A., Mordohai, P.*, Clipp, B., Engels, C., Gallup, D., Kim, S. -J., Merrell, P., Salmi, C., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewenius, H., Yang, R., Welch, G., Towles, H., 2008. Detailed Real-Time Urban 3D Reconstruction from Video. *International Journal of Computer Vision,* 78(2-3), 143-167.

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. *In 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)* Vol. 1, 519-528. IEEE.

Shen S, 2013. Accurate Multiple View 3D Reconstruction Using Patch-Based Stereo for Large-Scale Scenes. *IEEE Transactions on Image Processing*, 22(5).

Sinha, S. N., Mordohai, P., Pollefeys, M., 2007. Multi-View Stereo via Graph Cuts on the Dual of an Adaptive Tetrahedral Mesh. *IEEE International Conference on Computer Vision.* IEEE.

Snavely, N., 2010. Bundler: Structure from Motion (SfM) for Unordered Image Collections.

Strecha, C., Fransens, R., Gool, L., 2004. Wide-baseline stereo from multiple views: A probabilistic account. *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on.* IEEE.

Strecha, C., Hansen, W. V., Gool, L. V., Fua, P., Thoennessen,U., 2008. On benchmarking camera calibration and multi-view stereo for high resolution imagery. *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

Tsai, R., 1983. Multiframe image point matching and 3-d surface reconstruction. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 5(2), 159-74.

TU Darmstadt, 2012. [Source code]. https://github.com/simonfu-hrmann/mve/wiki.

Vogiatzis, G., Esteban, C. H., Torr, P., Cipolla, R., 2008. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 29(12), 2241-2246.

Vu, H. H., Keriven, R., Labatut, P., Pons, J. P., 2009. Towards high-resolution large-scale multi-view stereo. *IEEE Conference on Computer Vision & Pattern Recognition.* IEEE.

Wang, F., Galliani, S., Vogel, C., Speciale, P., Pollefeys, M., 2021. PatchmatchNet: Learned Multi-View Patchmatch Stereo. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* , 14194-14203.

Wu, C., 2013. Towards Linear-Time Incremental Structure from Motion. *3DV-Conference, 2013 International Conference on.*

Wu C., 2011 "VisualSFM: A Visual Structure from Motion System", http://ccwu.me/vsfm/.

Wu C., Sameer A., Brian C., Steven M. Seitz, 2011. "Multicore Bundle Adjustment", *CVPR 2011.*

Wu C., 2007. "SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT)", http://cs.unc.edu/~ccwu/siftgpu.