An Improved Patch based Multi-View Stereo (PMVS) Algorithm

WANG Lichun*   CHEN Ran   KONG Dehui

Beijing Key Laboratory of Multimedia and Intelligent Software Technology
College of Metropolitan Transportation, Beijing University of Technology
Beijing, China

Abstract—**Multi-view based reconstruction is always focused in computer graphics and many excellent algorithms have been reported these years. According to Middlebury benchmark, PMVS(Patch based Multi-View Stereo) outperforms all the other submitted algorithms [1]. In this paper, we propose an improved PMVS algorithm based on quasi-dense matching to save time cost of the original algorithm. Improved algorithm reduces running time for patch expansion through building a quasi-dense set of initial patches and depresses time complexity of the algorithm. The experiments demonstrate effectiveness of improved algorithm.**

Keywords-Multi-view reconstruction; PMVS; quasi-dense matching

I.   INTRODUCTION

Reconstructing 3D models from multi-view images can reduce artificial workload effectively, so it has a very wide application prospects in the fields of aerial mapping, visual navigation, medical diagnosis, electronic commerce and virtual reality, etc. The goal of multi-view stereo is to reconstruct a complete 3D object model from a collection of images taken from known camera viewpoints. According to the recent survey provided by Seitz [2], many excellent algorithms have achieved better accuracy. The algorithms can be roughly classified into four classes according to the object models employed in the rebuilding process: Voxel-based approaches [3],[4],[5],[6],[7] require knowing a bounding box that contains the scene, and the reconstruction accuracy is limited by resolution of the voxel grid. Deformable polygonal meshes based algorithms [8],[9],[10] demand a good starting point—for example, a visual hull model [11]—to initialize the corresponding optimization process, which limits their applicability. Approaches based on multiple depth maps [12],[13],[14] are more flexible, but require fusing individual depth maps into single 3D model. Finally, patch-based methods [15],[16],[17] represent scene surfaces by collections of small patches (or surfels). Among the four kinds of methods, the first three need prior knowledge (valid depth range, bounding box, or visual hull) to initiate the optimization process, which may not be feasible for outdoor scenes and/or cluttered images. While the patched-based methods do not require any initialization, it is capable of detecting and discarding outliers and obstacles automatically. A quantitative evaluation on the Middlebury benchmark shows that, for four out of the six data sets, the PMVS algorithm [16] outperforms all the other submitted. The PMVS algorithm consumes a bigger time cost while it exhibits the outstanding performance. This paper would improve time performance of the algorithm through exploiting quasi-dense matching method. The paper is organized as follows. Section 2 gives a simple introduction of PMVS algorithm and the method of Quasi-Dense matching briefly. Section 3 describes implementation steps of the improved PMVS algorithm. Experimental results and discussion are given in Section 3. Section 4 concludes the paper.

II.   RELATED WORK

A.   PMVS algorithm

PMVS is an algorithm for calibrated multi-view stereopsis which outputs a dense set of rectangular patches covering the surfaces visible in the input images. The PMVS algorithm associates with each image I_i a regular grid of $\beta \times \beta$ pixels cells $C_i(x,y)$ and attempts to reconstruct at least one patch in every image cell. The algorithm consists of three procedure, simple initial feature matching, patch expansion, and patch filtering:

(1) Feature Matching: Features found by Harris and difference-of-Gaussians (DOG) operators are first matched across multiple pictures, which yields a sparse set of patches associated with salient image regions;

(2) Patch expansion: Spread the initial patches to nearby pixels and obtain a dense set of patches;

(3) Filtering: Visibility constraints are then used to eliminate incorrect matches. The purpose of the initial feature matching step is to generate a sparse set of patches.

The expansion and the filtering steps are iterated 3 times (n=3 in the experiments below) to make patches dense and remove erroneous patches. Patch expansion is mainly part of the PMVS algorithm. The goal of the expansion step is to reconstruct at least one patch in every image cell $C_i(x, y)$. Since the set of initial patches (seed patches) generated by the feature matching step is sparse, thus the expansion step is massive work. Suppose that we can get a relatively dense set of the initial patches, the load of this step would be reduced, and saving of the expansion step will lead to reduction of the running time. This is basic idea of this paper.

B.   Quasi-Dense Matching

Generally, matching points between two views can be got both in the standards sparse approach and dense approach. But sparse matching of interest points is not sufficient for representing scenes or the objects in the scene as it merely reconstructs sparsely distributed 3D points [18],[19]. While the main disadvantages of the best dense stereo methods are

that they only reconstruct smoothed layers of disparities, special configurations in handling multiple views (often in one half-space looking at the other half-space and for images of small baselines), also they are very expensive in terms of time and memory [20]. Here we mainly introduce the Quasi-Dense matching approach proposed by Lhuillier and Quan [21],[22]. The quasi-dense approach gives more robust and accurate geometry estimations than the standard sparse approach, and produces a high density of points that can be used for direct surface reconstruction.

Quasi-dense matches start with the standard sparse matching algorithm between two images to detect the points of interest in each image, then propagate matches from some sparse matches to their neighborhoods. To match these points of interest in two images, Lhuiller and Quan suggested ZNCC (Zero-Mean Normalized Cross Correlation) method, which can get a list of sparse point correspondences. They further make the point correspondences dense by match propagation, first sort the list of point correspondences by decreasing correlation score as seed points. At each step of the propagation, they choose the best corresponding points scored by ZNCC from the current list of seed points. Then, in the immediate spatial neighborhood of the seed points, they look for new potential and add the bests to the current list of seed points according to a combination of local constraints, including correlation, gradient disparity, and confidence. A more detailed description of match propagation and its properties can be found in [21]. Quasi-Dense matching is a very efficient approach, both in time and space, and at each time, only the best match is selected. This drastically limits the possibility of bad matches. For instance, the seed selection step seems very similar to many existing methods [23] for matching points of interest using correlation, but the crucial difference is that we need only take the most reliable ones rather than trying to match a maximum number of them. This is a best-first match propagation strategy which produces denser, but not completely dense point correspondences. The quasi-dense matching approach is robust to outliers lying in the set of initial matches due to the best-first strategy in the propagation.

III. THE IMPROVED PMVS ALGORITHM

In PMVS, the patch expansion step start from the initial patches, and construct no less than one patch in every image cell $C_i(x, y)$. The expansion process spends a lot of time. In order to improve efficiency of expansion, we produce more initial patches. That means we need denser features. For this purpose, the quasi-dense approach is combined into PMVS to generate the initials patches which are denser. The improved PMVS algorithm based on quasi-dense matching is implemented as follows, it consists of five steps (Fig. 1): (1) Initial feature matching: Feature matching to get the sparse matched features; (2) Quasi-dense matching: Combine Quasi-dense matching approach to obtain a denser matched features; (3) Resample: Resample the matched features to be more uniform and stable; (4) Patch expansion; (5) Patch filtering
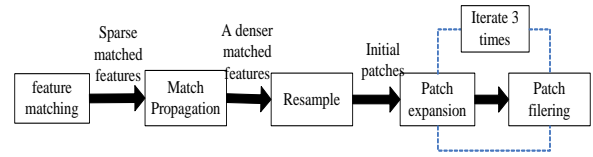

Figure 1.   The flow chart of the improved PMVS algorithm.

In details:

1) Initial Feature Matching: We extract the feature points in each images using SIFT [24] descriptor instead of the Harris and DOG using in the original PMVS. Because SIFT features are invariant to rotation and scaling, which provides robust matching with better stability. The stability of the matching feature points is necessary to the match propagation process in next step, it will enhance the reliability of seed points.

2) Quasi-Dense Matching: In this step we attempt to make the matching points dense by using match propagation. To get more accuracy matching points, The ZNCC is used to reject definitively any match whose ZNCC score is less than 0.8. This gives a list of sparse point correspondences more reliably. We first sort this list of point correspondences using the ZNCC score. These sorted points are called seed points. At each step of the propagation, we choose the best corresponding points scored by ZNCC from the current list of seed points. Then, in the immediate spatial neighborhood of the seed points, we look for new potential matches and add the bests to the current list of seed points according to a combination of local constraints, including correlation, gradient disparity, confidence and uniqueness.

3) Resample: We want to produce uniform distribution of matched points in the images through resampling. The first image plane is initially divided into a regular square grid of 8×8 pixels. For each square region, all point correspondences inside it are used to tentatively fit a plane transformation. The Random Sample Consensus (RANSAC) [25] is used for a robust estimation of the transformation, H. Finally, for each confirmed region correspondence, a pair of corresponding points $u_i \rightarrow H_i u_i$ is created by selecting a representative center point of the patch in the first image, u_i, and its corresponding point $H_i u_i$ in the second image. The corresponding points created this way are called the "quasi-dense correspondences". In practice, we add the quasi-dense correspondences to the list of point correspondences and sort both by ZNCC scores to get the final quasi-dense point correspondences. These resampled corresponding points are more suitable and more reliable for reconstruction thanks to their more uniform distribution in image.

4) Patch Expansion: Patch expansion step start with the initial patches triangulated from quasi-dense point correspondences in the step 3, generate new patches in nearby empty spaces using the properties of the similarity of the unit normal vector and location between the neighboring patches with depth continuity and neighboring constraints.

5) Patch Filtering: To remove erroneous patches, visibility constraints with photometric discrepancy function are used to eliminate incorrect patches lying either in front or behind the observed surface.

IV. COMPARATIVE EXPERIMENT

For testing the algorithm suggested in this paper, we compare it with the PMVS algorithm on the Middlebury Temple (47 views) and Dino (48 views) datasets. The reconstruction results are shown in Fig. 2.

In Fig. 2, the upper row is reconstruction result for the Dino dataset and the lower line is reconstruction result for the Temple dataset. The left column is original model, the middle column is reconstruction result by PMVS and the right column is reconstruction result by using method provided in this paper. The number of final patches and time cost for two algorithms are listed in Table Ⅰ.

The experimental results in Table Ⅰ show that, for Temple datasets the PMVS generates 90081 patches and the improved algorithm generates 89867 patches; for Dino datasets the PMVS generates 99345 patches and the improved algorithm generates 98763 patches. All reconstructed patches equal unit normal vector which means the reconstructed results have equal visual effect. Time cost of improved algorithm and PMVS are also listed in Table Ⅰ.

Table Ⅰ suggests that, improved algorithm has 9.21% time saved and 0.59% patches loss on Dino dataset and 8.85% time saved and 0.24% patches loss on Temple dataset, while compared with the PMVS algorithm. That means the improved algorithm has 9% time cost saved and 0.42% patches loss on average. It also means improved algorithm can save reconstruction time on the condition of having mini effect on the reconstruction result.

All experiments are carried out on a computer equipped with 4GB RAM, 3.30GHz CPU. And the algorithms are implemented in C++.

Figure 2. Reconstruction for the Middlebury Dino and Temple dataset.

TABLE I.   NUMBER OF RECONSTRUCTED PATCHES AND TIME COST

Model	Source Images	Reconstructed Patches		Time Cost(s)	
		PVMS	*Improved*	*PVMS*	*Improved*
Dino	48 views	99345	98763	564	512
Temple	47 views	90081	89867	531	484

V. CONCLUSION

PMVS is one of the excellent Multi-view stereo algorithms, but it handles the datasets with a high time-complexity. To improve this problem, we implement the PMVS algorithm based on the quasi-dense matching. The algorithm shows the significance that the patches expansion performed, considering the influence of the initial patches in the patches expansion steps, we have proposed the quasi-dense matching method to get a quasi-dense set of initial patches. Our experiment demonstrates that this improved method can lessen the running time of expansion step, further reduce the total computation time and improve efficiency of the algorithm.

REFERENCES

[1] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski,"Multi-View Stereo Evaluation," http://vision.middlebury.edu/mview/, 2010

[2] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2006.

[3] O. Faugeras and R. Keriven, "Variational Principles, Surface Evolution, PDE's, Level Set Methods and the Stereo Problem," IEEE Trans. Image Processing, vol. 7, no. 3, pp. 336-344, Mar. 1998.

[4] S. Paris, F. Sillion, and L. Quan, "A Surface Reconstruction Method Using Global Graph Cut Optimization," Proc. Asian Conf. Computer Vision, http://artis.imag.fr/Publications/2004/PSQ04, Jan. 2004.

[5] S. Tran and L. Davis, "3D Surface Reconstruction Using Graph Cuts with Surface Constraints," Proc. European Conf. Computer Vision, 2006.

[6] G. Vogiatzis, P.H. Torr, and R. Cipolla, "Multi-View Stereo Via Volumetric Graph-Cuts," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005.

[7] S. Sinha, P. Mordohai, and M. Pollefeys, "Multi-View Stereo via Graph Cuts on the Dual of an Adaptive Tetrahedral Mesh," Proc. Int'l Conf. Computer Vision, 2007.

[8] C. Herna´ndez Esteban and F. Schmitt, "Silhouette and Stereo Fusion for 3D Object Modeling," Computer Vision and Image Understanding, vol. 96, no. 3, pp. 367-392, 2004.

[9] A. Zaharescu, E. Boyer, and R. Horaud, "Transformesh: A Topology-Adaptive Mesh-Based Approach to Surface Evolution," Proc. Asian Conf. Computer Vision, Y. Yagi, S.B. Kang, I.-S. Kweon, and H. Zha, eds., pp. 166-175, 2007.

[10] Y. Furukawa and J. Ponce, "Carved Visual Hulls for Image-Based Modeling," Int'l J. Computer Vision, vol. 81, pp. 53-67, Mar. 2009.

[11] B. Baumgart, "Geometric Modeling for Computer Vision," PhD dissertation, Stanford Univ., 1974.

[12] M. Goesele, B. Curless, and S.M. Seitz, "Multi-View Stereo Revisited," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 2402-2409, 2006.

[13] C. Strecha, R. Fransens, and L.V. Gool, "Combined Depth and Outlier Estimation in Multi-View Stereo," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 2394-2401, 2006.

[14] D. Bradley, T. Boubekeur, and W. Heidrich, "Accurate Multi-View Reconstruction Using Robust Binocular Stereo and Surface Meshing," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2008.

[15] M. Habbecke and L. Kobbelt, "Iterative Multi-View Plane Fitting," Proc. 11th Fall Workshop Vision, Modeling, and Visualization, 2006. 73−80

[16] Furukawa Y, Ponce J. Accurate, dense, and robust multi-view stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(8): 1362−1376

[17] 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

[18] R.I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge Univ. Press, June 2000.

[19] A. Heyden, "Geometry and Algebra of Multiple Projective Transformations," PhD thesis, Lund Inst. of Technology, 1995.

[20] V. Kolmogorov and R. Zabih, "Multi-Camera Scene Reconstruction via Graph Cuts," Proc. Seventh European Conf. Computer Vision, 2002.

[21] Lhuillier M and Quan L. Match propagation for image-based modeling and rendering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(8):1140–1146, August 2002.

[22] Lhuillier M and Quan L. A quasi-dense approach to surface reconstruction from uncalibrated images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(3):418–433, March 2005.

[23] P.H.S Torr and D.W.Murray,"The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix," Int'l J. Computer Vision, vol. 24, no. 3, pp. 271-300, 1997.

[24] Lowe D.G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004.