Web Scale Photo Hash Clustering on A Single Machine

Xuewen Yang

June 4 2018

Abstract

In this paper the author present an approach to indoor scene reconstruction from RGB-D video. The key idea is to combine geometric registration of scene fragments with robust global optimization based on line processes. Geometric registration is error-prone due to sensor noise, which leads to aliasing of geometric detail and inability to disambiguate different surfaces in the scene. The presented optimization approach disables erroneous geometric alignments even when they significantly outnumber correct ones. Experimental results demonstrate that the presented approach substantially increases the accuracy of reconstructed scene models.

1. Introduction

Prior work on scene reconstruction with consumer depth cameras recognized the importance of global registration [4] [2]. Nevertheless, no prior system appears to be sufficiently reliable to support automatic reconstruction of complete indoor scenes at a quality level appropriate for particularly demanding applications. This is evidenced by the recent effort of Xiao *et al.* to reconstruct a large number of indoor scenes. Due to the unreliability of automatic scene reconstruction pipelines, the authors resorted to manual labeling to establish correspondences between different views. In this work,the author present a fully automatic scene reconstruction pipeline that matches the reconstruction quality obtained with manual assistance by Xiao *et al.* and significantly exceeds the accuracy of prior automatic approaches to indoor reconstruction.An example reconstruction produced by this approach is shown in Figure 1.Their approach resolves inconsistencies and identifies correct alignments using global optimization based on line processes.Line processes are closely related to robust estimation [1].The advantage of the line process formulation is that the optimization objective retains a least-squares form and can be optimized by a standard high-performance least-squares solver.The author show that this framework is extremely effective in dealing with pairwise registration er-

rors.Their implementation automatically prunes false pairwise alignments even when they significantly outnumber correct ones. Extensive experiments demonstrate that their approach substantially increases reconstruction accuracy.

Their work contains a number of supporting contributions of independent interest. First, they provide infrastructure for rigorous evaluation of scene reconstruction accuracy, augmenting the ICL-NUIM dataset [3] with challenging camera trajectories and a realistic noise model. Second, we perform a thorough quantitative evaluation of surface registration algorithms in the context of scene reconstruction; our results indicate that well-known algorithms perform surprisingly poorly and that algorithms introduced in the last few years are outperformed by older approaches. Third, in addition to accuracy measurements on synthetic scenes we describe an experimental procedure for quantitative evaluation of reconstruction quality on real-world scenes in the absence of ground-truth data.

2. Conclusion

the author presented an approach to scene reconstruction from RGB-D video. The key idea is to combine geometric registration with global optimization based on line processes.The optimization makes the pipeline robust to erroneous geometric alignments, which are unavoidable due to aliasing in the input. Experimental results demonstrate that the presented approach significantly improves the fidelity of indoor scene models produced from consumer-grade video.The presented pipeline is not foolproof. First, if the input video does not contain loop closures that indicate global geometric relations, odometry drift can accumulate and distort the reconstructed model. Real-time feedback that guides the operator to close loops would help.

References

[1] M. J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *IJCV*, 1996. 1

[2] F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard. 3-D mapping with an RGB-D camera. *IEEE Transactions on Robotics*, 2017. 1
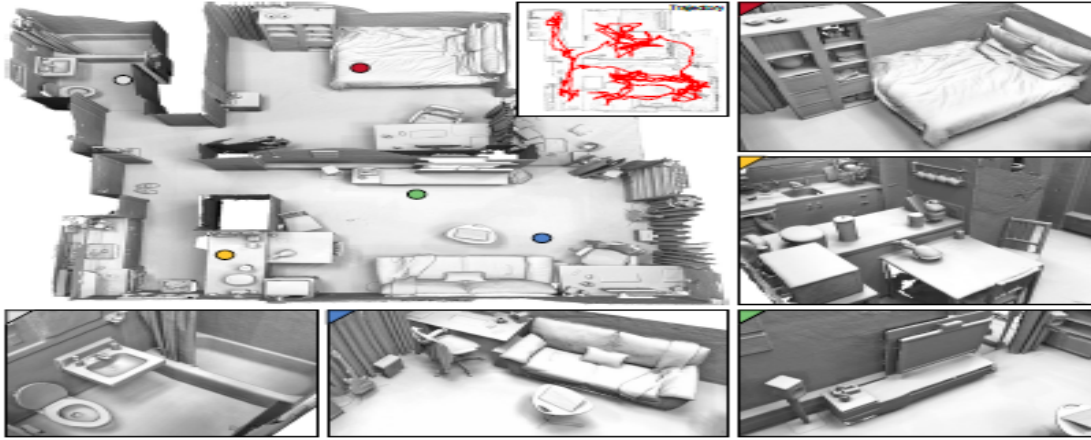
Figure 1. A complete apartment reconstructed by the presented approach. The estimated camera trajectory is 151.6 meters long, folded into a diameter of 8.3 meters.

[3] A. Handa, T. Whelan, J. Mcdonald, and A. J. Davison. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM. In *ICRA*, 2014. 1

[4] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments. *IJRR*, 2014. 1