Joint Layout Estimation and Global Multi-View Registration for Indoor Reconstruction: Supplementary Material

Jeong-Kyun Lee*[1], Jaewon Yea*[2], Min-Gyu Park[3], and Kuk-Jin Yoon[1]

[1]School of Electrical Engineering and Computer Science, GIST, Gwangju, South Korea
[2]LG Electronics, Incheon, South Korea
[3]Korea Electronics Technology Institute, Seongnam, South Korea
{leejk,kjyoon}@gist.ac.kr, jaewon.yea@lge.com, mpark@keti.re.kr

In this supplementary material, we specify all the parameters mentioned in the paper and show all of the experimental results for qualitative evaluation, using four synthetic datasets of the augmented ICL-NUIM datasets [1] (*Livingroom1*, *Livingroom2*, *Office1*, and *Office2*), and six real-world datasets of the SUN3D datasets [4] (*mit_dorm_next_sj*, *mit_lab_hj*, *mit_76_458*, *harvard_c9*, *harvard_c3*, and *brown_cogsci_4*), respectively. We also show the jointly estimated layout along with the reconstruction result for each dataset.

1. Parameter Setting

In this section, we specify the parameters mentioned in the paper. In Eq. (4), the penalty weight of the Potts model, $\alpha_{\mathbf{p},\mathbf{q}}$, is set to 5, and a point $\mathbf{q}$ belongs to $\mathcal{N}_{\mathbf{p}}$, *i.e.*, a set of neighboring points of the point $\mathbf{p}$, if the distance between $\mathbf{p}$ and $\mathbf{q}$ is within 0.02m. In Eq. (5), γ is set to 0.1. When an occupancy grid map is generated to find the boundary of the room, ∂O, as in Fig. 4, the size of a grid is set to 0.01m $\times$ 0.01m, and each cell of the grid map is regarded as an occupied cell if at least one point is projected into the cell. In Eq. (6), τ_1 and τ_2 are set to 5° and 0.3m, respectively. In Eq. (9), a virtual point $\mathbf{q}$ on the layout plane is generated if the distance between a point $\mathbf{p}$ on the fragment $\mathbf{F}_i$ and the layout plane is within 0.1m. In Eq. (8), the weight parameters are defined as $\lambda_1 = c_1 \frac{n_l}{n_f}$ and $\lambda_2 = c_2 \frac{n_l}{n_p}$, where $c_1 = 1$, $c_2 = 0.3$, $n_l = \sum_i |\mathcal{C}_i|$, $n_f = \sum_{i,j} |\mathcal{C}_{i,j}|$, and $n_p = |\mathcal{I}|$. $|\cdot|$ means the cardinality of a set.

*Denote equal contribution.

2. Experimental Results for Synthetic Datasets

(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

Figure 1. Comparison of reconstruction results for the *Livingroom1* dataset

(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

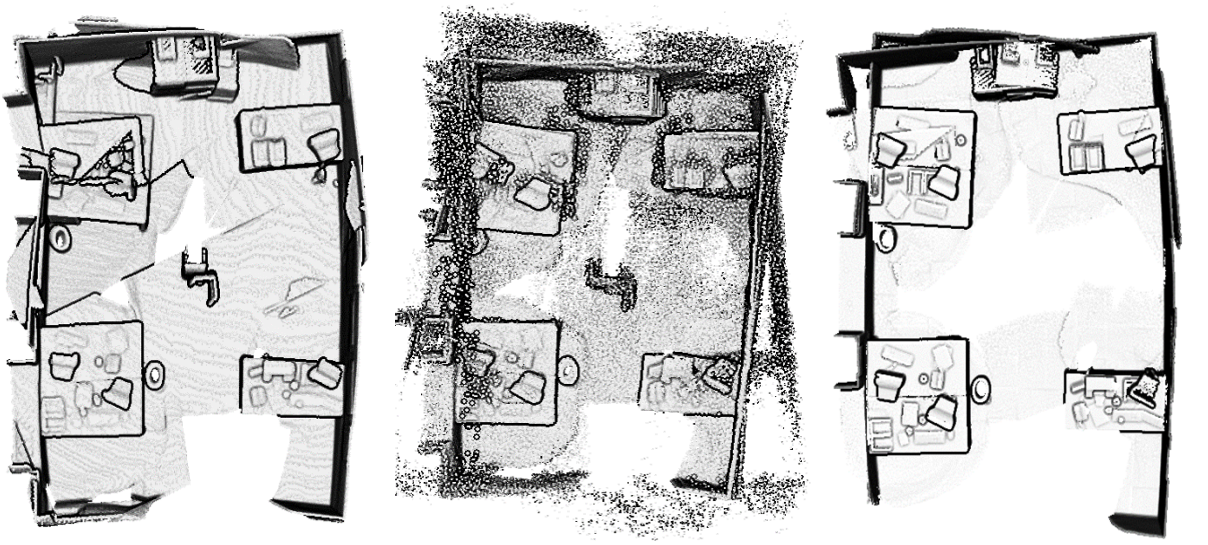
(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

Figure 2. Comparison of reconstruction results for the *Livingroom2* dataset

(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

Figure 3. Comparison of reconstruction results for the *Office1* dataset

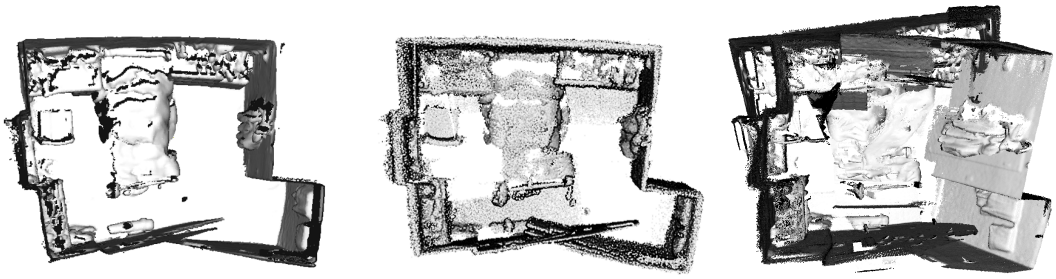


(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

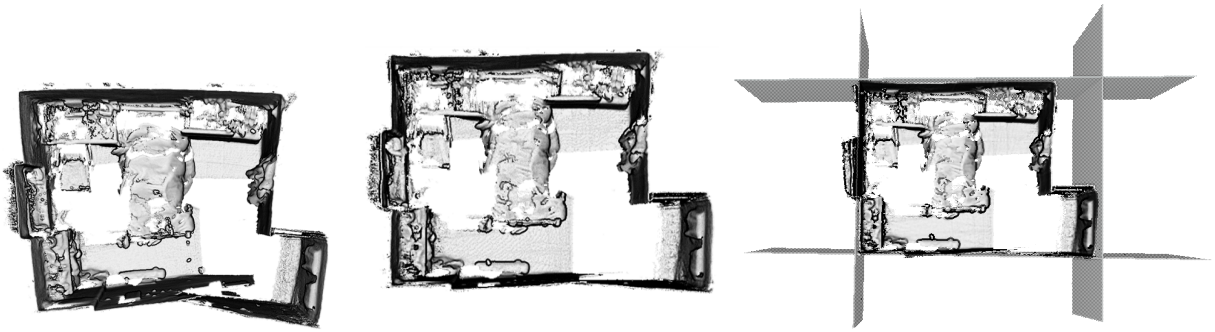
(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

Figure 4. Comparison of reconstruction results for the *Office2* dataset

3. Experimental Results for Real-world Datasets

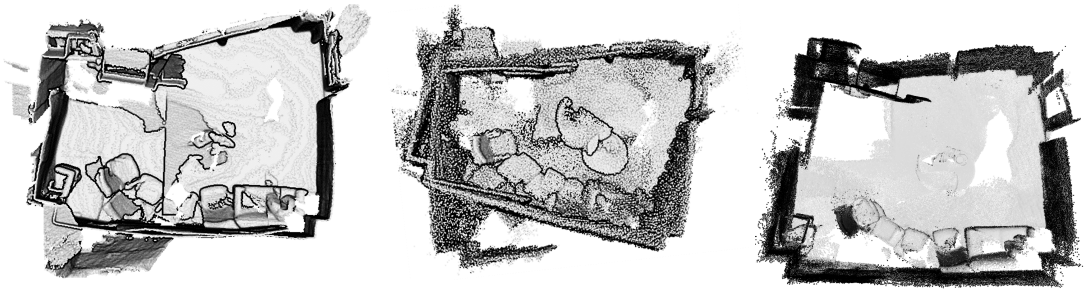


(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]



(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

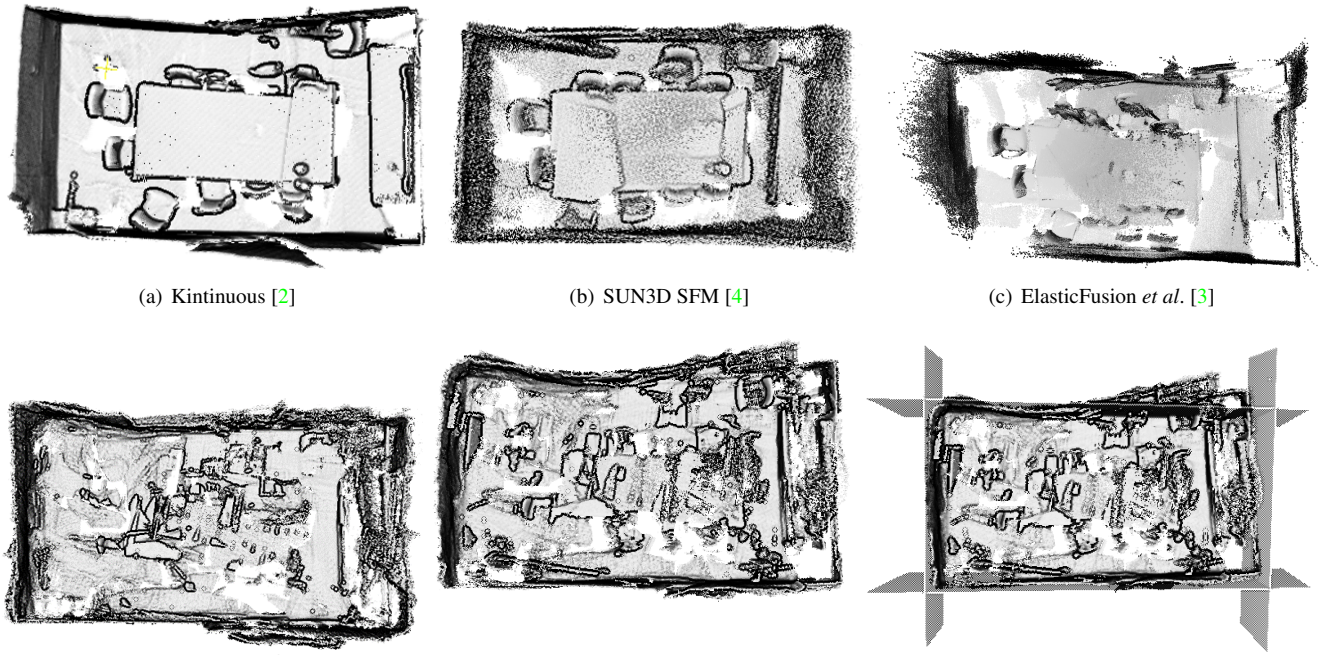
Figure 5. Comparison of reconstruction results for the *mit_dorm_next_sj* dataset



(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

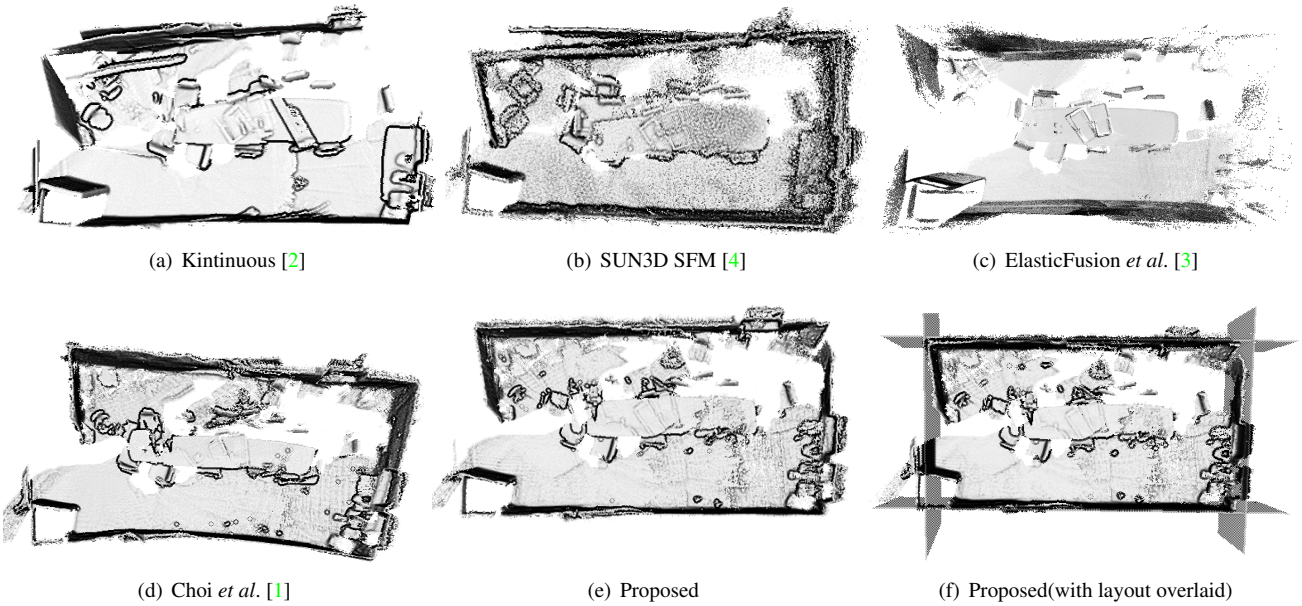
Figure 6. Comparison of reconstruction results for the *mit_lab_hj* dataset



(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

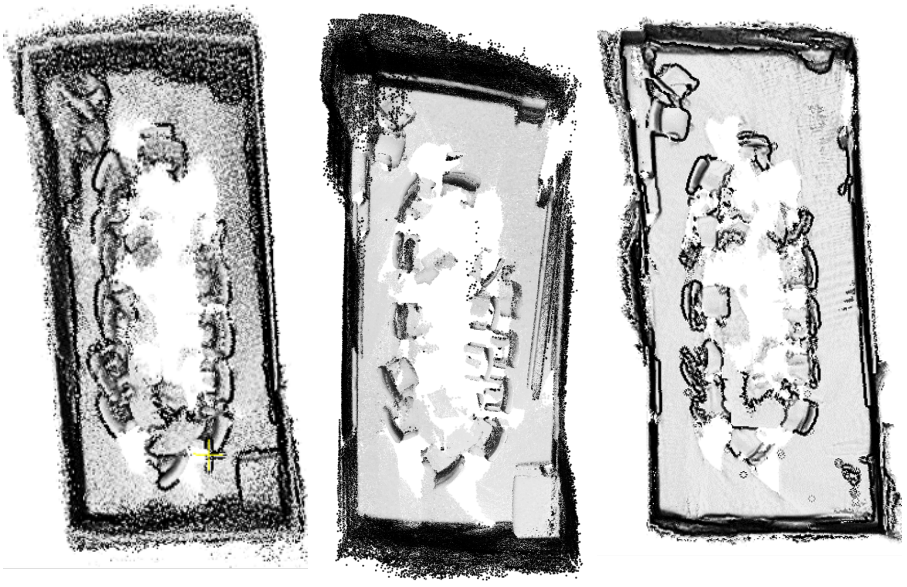
Figure 7. Comparison of reconstruction results for the *mit_76_458* dataset



(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Proposed (f) Proposed(with layout overlaid)

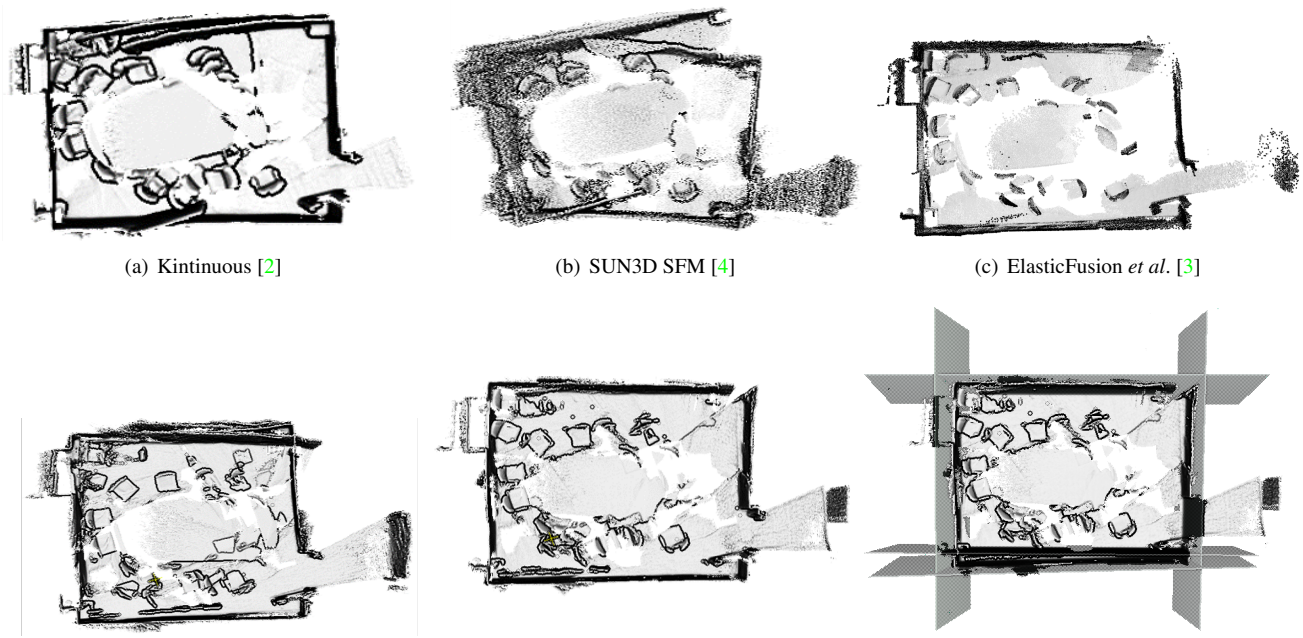
Figure 8. Comparison of reconstruction results for the *harvard_c9* dataset



(a) SUN3D SFM [4] (b) ElasticFusion *et al*. [3] (c) Choi *et al*. [1]

(d) Proposed (e) Proposed(with layout overlaid)

Figure 9. Comparison of reconstruction results for the *harvard_c3* dataset. In this case, Kintinuous [2] does not work for this dataset.



(a) Kintinuous [2] (b) SUN3D SFM [4] (c) ElasticFusion *et al*. [3]

(d) Choi *et al*. [1] (e) Prop osed (f) Proposed(with layout)

Figure 10. Comparison of reconstruction results for the *brown_cogsci_4* dataset

References

[1] S. Choi, Q.-Y. Zhou, and V. Koltun. Robust reconstruction of indoor scenes. In *CVPR*, 2015. 1, 2, 3, 4, 5, 6, 7, 8, 9

[2] T. Whelan, M. Kaess, M. Fallon, H. Johannsson, J. Leonard, and J. McDonald. Kintinuous: Spatially extended KinectFusion. In *RSS Workshop on RGB-D: Advanced Reasoning with Depth Cameras*, 2012. 2, 3, 4, 5, 6, 7, 8, 9

[3] T. Whelan, S. Leutenegger, R. S. Moreno, B. Glocker, and A. Davison. Elasticfusion: Dense slam without a pose graph. In *RSS*, 2015. 2, 3, 4, 5, 6, 7, 8, 9

[4] J. Xiao, A. Owens, and A. Torralba. SUN3D: A database of big spaces reconstructed using sfm and object labels. In *ICCV*, 2013. 1, 2, 3, 4, 5, 6, 7, 8, 9