Vision, Modeling, and Visualization (2017)
M. Hullin, R. Klein, T. Schultz, A. Yao (Eds.)

Dense and Scalable Reconstruction from Unstructured Videos with Occlusions (Supplementary Material)

Jian Wei[1,2] and Benjamin Resch[1] and Hendrik P. A. Lensch[1]

[1]Computer Graphics, Tübingen University, 72076 Tübingen, Germany
[2]Communication Engineering, Jilin University, 130012 Changchun, China

This document gives additional details and results of our approach that owing to the space limitation did not fit in the main paper. Section 1 first describes how we handle the visibility conflicts more efficiently, and then Section 2 introduces how associated information of the obtained point cloud is updated in our reconstruction. After these, Section 3 aims to demonstrate the effectiveness of our view selection and clustering for real-world scenes. Moreover, please see the supplementary videos for better visualization of the reconstructed large-scale scenes.

1. Visibility Conflict Invalidation

As mentioned in Section 3 (see Step 2) of the main paper, we extend the conflict invalidation strategy of [WRL16] to support growing of occlusion errors over the multi-view 3D points rather than within single depth maps. This is done by exploiting the points' neighborhood connections. Because detecting the seed invalid areas requires the depth and uncertainty values of the edges in different views, we need to maintain the edge's depth and uncertainty maps of all processed views. However, as the proposed approach only processes a minimum set of reference views, our memory requirements are still modest.

In particular, at the *per-cluster* level, let $\boldsymbol{D}^e$ and $\boldsymbol{U}^e$ denote the edge's depth and uncertainty maps of the views which produce the point cloud $\mathcal{P}$ so far. Assume a point set $\mathcal{P}_n$ created from a new view whose edge's depth and uncertainty maps are D_n^e and U_n^e, respectively. We solve the inconsistency between $\mathcal{P}$ and $\mathcal{P}_n$ by refining $\mathcal{P}$ using D_n^e and U_n^e, followed by conversely refining $\mathcal{P}_n$ using $\boldsymbol{D}^e$ and $\boldsymbol{U}^e$.

To refine $\mathcal{P}$, we first judge the invalidities of its interpolated points using the local detection method of [WRL16] by back-projecting $\mathcal{P}$ to D_n^e, and store the sparse invalidation errors in $\mathcal{E}$ ($|\mathcal{E}| = |\mathcal{P}|$). To grow the invalid areas, two uncertainty values are compared for each non-edge point: We store the uncertainties of all points in $\mathcal{P}$ in a dense set $\mathcal{U}$ as their *expected* uncertainties. We diffuse $\mathcal{E} + \mathcal{U}$ using the points' neighborhood information obtaining another dense uncertainty set $\mathcal{C}$ as the *approximated* uncertainties.

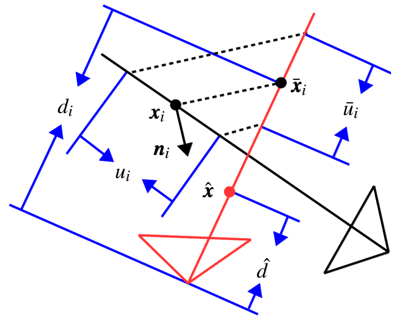

Figure 1: *Uncertainty merging by projecting the 3D position $\boldsymbol{x}_i$ and uncertainty u_i of each point that influences its new position $\hat{\boldsymbol{x}}$ onto the corresponding visual ray (red line).*

If the value of a non-edge point in $\mathcal{C}$ is 1.2 times larger than the corresponding value in $\mathcal{U}$, this point is invalid.

$\mathcal{P}_n$ is refined similarly. The difference is that the invalidation errors in $\mathcal{E}$ are obtained by back-projecting $\mathcal{P}_n$ to all views of $\boldsymbol{D}^e$.

As described in the main paper, invalidation of visibility conflicts is also performed at the *multi-cluster* level, *i.e.*, among multiple point sets each obtained from one view cluster.

2. Associated Information Updating

Once a point gets a new position from the merging step, its other associated attributes (see the point representation in Section 3 of the main paper) are updated accordingly. As in [FG14], the normal and color are calculated using a second, simpler implicit function (see that paper for details). The neighborhood information is updated by merging the connection information of surviving neighboring points. The point scale is recalculated on the optimized surface.

To update the point uncertainty, we utilize the scheme of [CWPS11] which is devoted to merging multiple Gaussian components. Since the uncertainty is defined along the visual ray, we make a modification to support cross-ray merging. For the point $\boldsymbol{P}$

BuildingL ChurchL

Figure 2: *Views selected by our method for reconstructing BuildingL and ChurchL. Different colors of the boundaries represent different view clusters.*

at a new position $\hat{\boldsymbol{x}}$, we project the position and uncertainty of each point $\boldsymbol{P}_i \in \mathcal{P}_f(\hat{\boldsymbol{x}})$ (see Eq. 2 in the main paper) along the tangent plane of the surface at $\boldsymbol{P}_i$ onto the visual ray of $\boldsymbol{P}$, as depicted in Fig. 1. Let the projected position and uncertainty be $\bar{\boldsymbol{x}}_i$ and $\bar{u}_i$, respectively. By denoting the depths of $\bar{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{x}}$ along the ray as d_i and $\hat{d}$, the merged uncertainty is defined by

$$\hat{u} = \sqrt{\frac{\sum_{\boldsymbol{P}_i} w(\boldsymbol{x}_i)(\bar{u}_i^2 + (d_i - \hat{d})^2)}{\sum_{\boldsymbol{P}_i} w(\boldsymbol{x}_i)}}. \tag{1}$$

Hereby, the obtained new value combines multiple input uncertainties and also respects the surface distances. Afterwards we diffuse the updated uncertainties over the merged surface for smoothness.

3. Additional Results

Figure 2 shows the views used and the view clusters constructed by our content-aware method for recovering BuildingL and ChurchL. The figure demonstrates that the proposed method can select the most useful views for the large-scale reconstruction, among which the views observing the same scene geometry contribute to a cluster thus allowing locally consistent and concise reconstruction.

Please see the supplementary videos for comparing the reconstructed large scenes (BathroomL, BuildingL, and ChurchL) in the 3D space.

References

[CWPS11] CROUSE D. F., WILLETT P., PATTIPATI K., SVENSSON L.: A look at gaussian mixture reduction algorithms. In *International Conference on Information Fusion* (2011). 1

[FG14] FUHRMANN S., GOESELE M.: Floating scale surface reconstruction. *ACM Trans. Graph. 33*, 4 (2014), 46. 1

[WRL16] WEI J., RESCH B., LENSCH H. P. A.: Dense and occlusion-robust multi-view stereo for unstructured videos. In *CRV* (2016). 1