Superpixel Meshes for Fast Edge-Preserving Surface Reconstruction (Supplementary material)

András Bódis-Szomorú[1] Hayko Riemenschneider[1] Luc Van Gool[1,2]
[1]Computer Vision Lab, ETH Zurich Switzerland [2]PSI-VISICS, KU Leuven, Belgium
{bodis,hayko,vangool}@vision.ee.ethz.ch

Abstract

Multi-View-Stereo (MVS) methods aim for the highest detail possible, however, such detail is often not required. In this work, we propose a novel surface reconstruction method based on image edges, superpixels and second-order smoothness constraints, producing meshes comparable to classic MVS surfaces in quality but orders of magnitudes faster. Our method performs per-view dense depth optimization directly over sparse 3D Ground Control Points (GCPs), hence, removing the need for view pairing, image rectification, and stereo depth estimation, and allowing for full per-image parallelization. We use Structure-from-Motion (SfM) points as GCPs, but the method is not specific to these, e.g. LiDAR or RGB-D can also be used. The resulting meshes are compact and inherently edge-aligned with image gradients, enabling good-quality lightweight per-face flat renderings. Our experiments demonstrate on a variety of 3D datasets the superiority in speed and competitive surface quality.

1. Overview

These pages provide the following supplementary qualitative results:

- *Watertight models vs. slanted-plane piecewise-models*: Figure 1 compares our approach qualitatively to two other single-view depth estimation approaches to demonstrate conceptual differences: (1) least-squares superpixel depth estimation with fronto-parallel assumption and SfM points as soft-constraints, (2) slanted-plane iterative global optimization method motivated by slanted-plane stereo methods. It optimizes four plane parameters per superpixel (expensive due to a large parameter space) while enforcing pairwise smoothness. Unlike these, our proposed method is naturally watertight, and penalizes curvature instead of depth differences.
- *Discontinuities*: Figure 2 demonstrates, in more detail, our approach for discontinuity removal.
- *More single-view models*: Figures 3, 4 show more single-view surface models from the different small-to-medium-scale datasets listed in Table 1 of the paper, in comparison to the SfM+DT[1] baseline.
- *Large-scale close-ups*: Figure 5 shows facade-level close-ups for the two large-scale models listed in Table 1, and shown in Figure 9 of the paper. These models are built by simply concatenating single-view models to show their consistency and applicability for large-scale surface reconstruction.

[1]2D Delaunay-triangulation over 2D SfM point observations in a view, followed by lifting the triangulation into 3D via the known sparse SfM depths.

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

2. Qualitative results

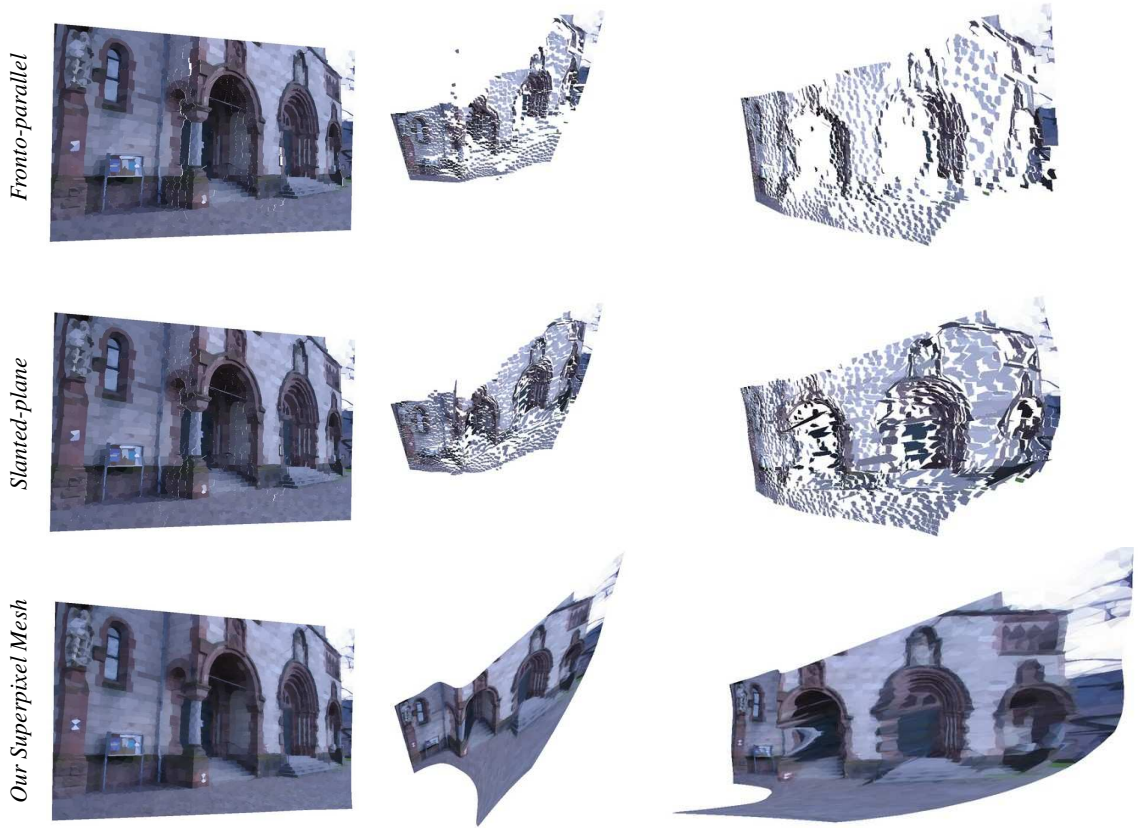

Figure 1. Qualitative comparison between three conceptually different single-view depth estimation approaches (motivated by slanted-plane stereo methods, though). Different methods are shown in different rows: (1) superpixel depth estimation via fronto-parallel plane assumption (corresponds to SLIC+Fronto+Soft in the paper's terminology), (2) global iterative slanted-plane optimization enforcing smoothness (after around 500 iterations and 40 minutes in Matlab), (3) complete image reconstructed using our approach (without clustering, α-shapes and discontinuity removal). All three methods are applied over the same (SLIC) superpixelization. Columns correspond to different viewpoints of the same mesh. The 1st column shows the mesh approximately from the original viewpoint. Our method penalizes curvature: please note the planarity of our model (3rd row) around the edges where no SfM data is available. Unlike the piecewise models, our method is naturally watertight.

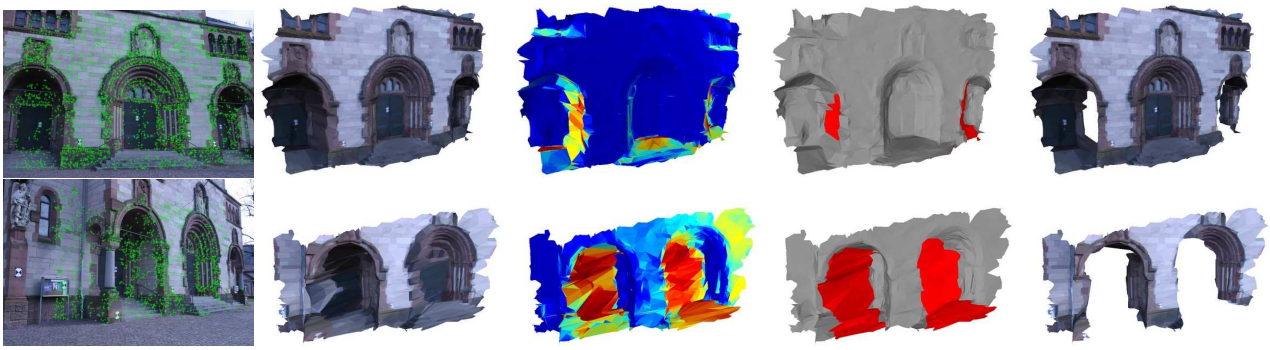

Figure 2. Discontinuity removal demonstrated on two images of HerzJesu. Columns: (1) input image and SfM points, (2) our initial reconstruction, (3) color-coded observation angles of mesh faces (yellow-reddish marks faces seen in sharp angles), (4) graph-cuts segmentation of discontinous regions, (5) final model after refitting.

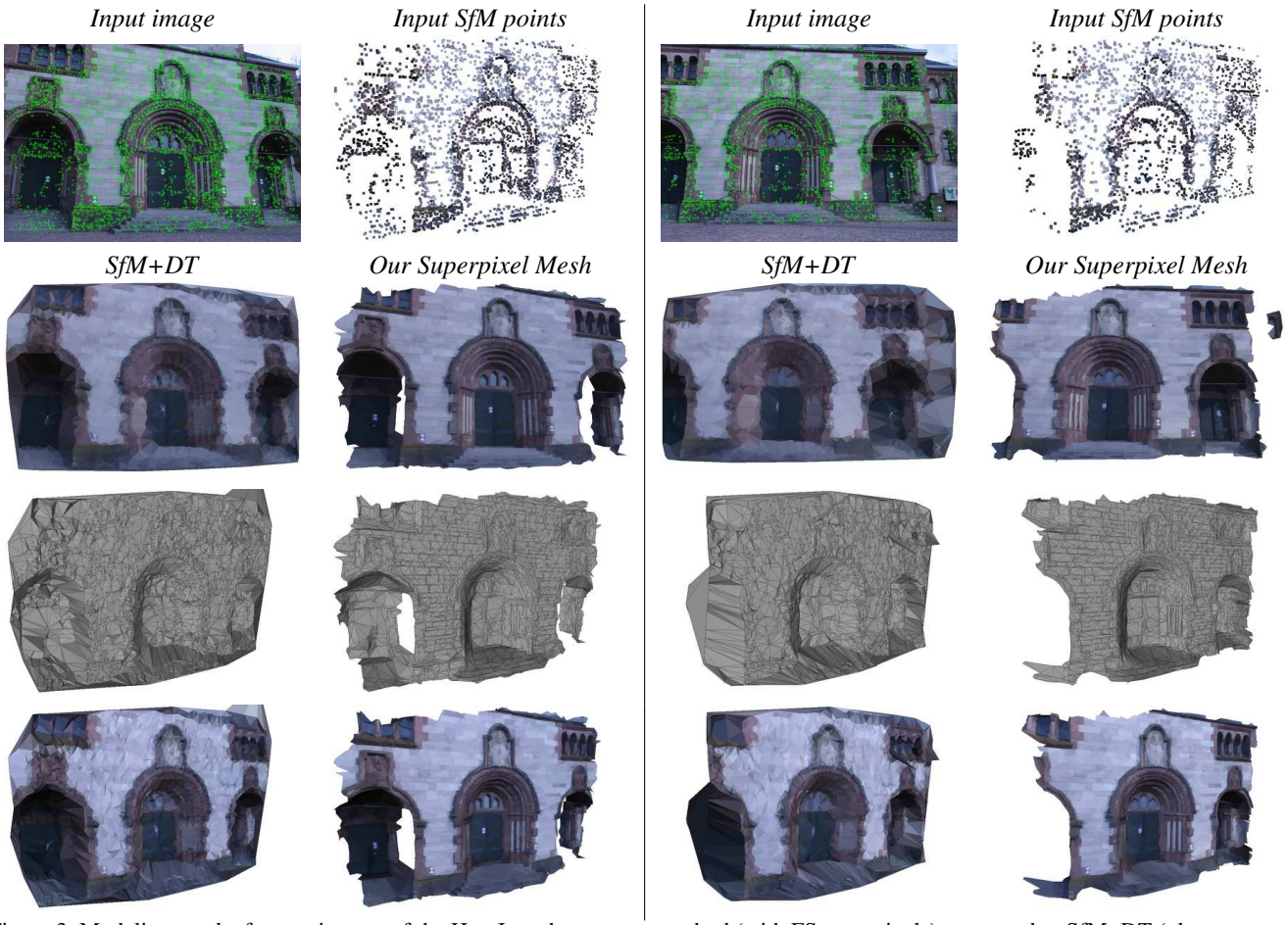

Figure 3. Modeling results for two images of the HerzJesu dataset: our method (with FS superpixels) compared to SfM+DT (please zoom for the details). The left two columns shows results forand the right two columns show results for two different views. 1st row: input image and the subset of SfM points visible in the view. 2nd row: rasterizations of mean-colors per 2D triangle. The rasterized triangles of SfM+DT clearly cross actual geometric edges. 3rd and 4th rows: resulted 3D models without coloring and with per-face flat (textureless) coloring. Our method produces meshes that are smoother and with edges aligned to the actual 3D edges. Also note the effect of 2D α-shapes on the arc in the 4th column. A 2D convex hull would undesirably close the arc.

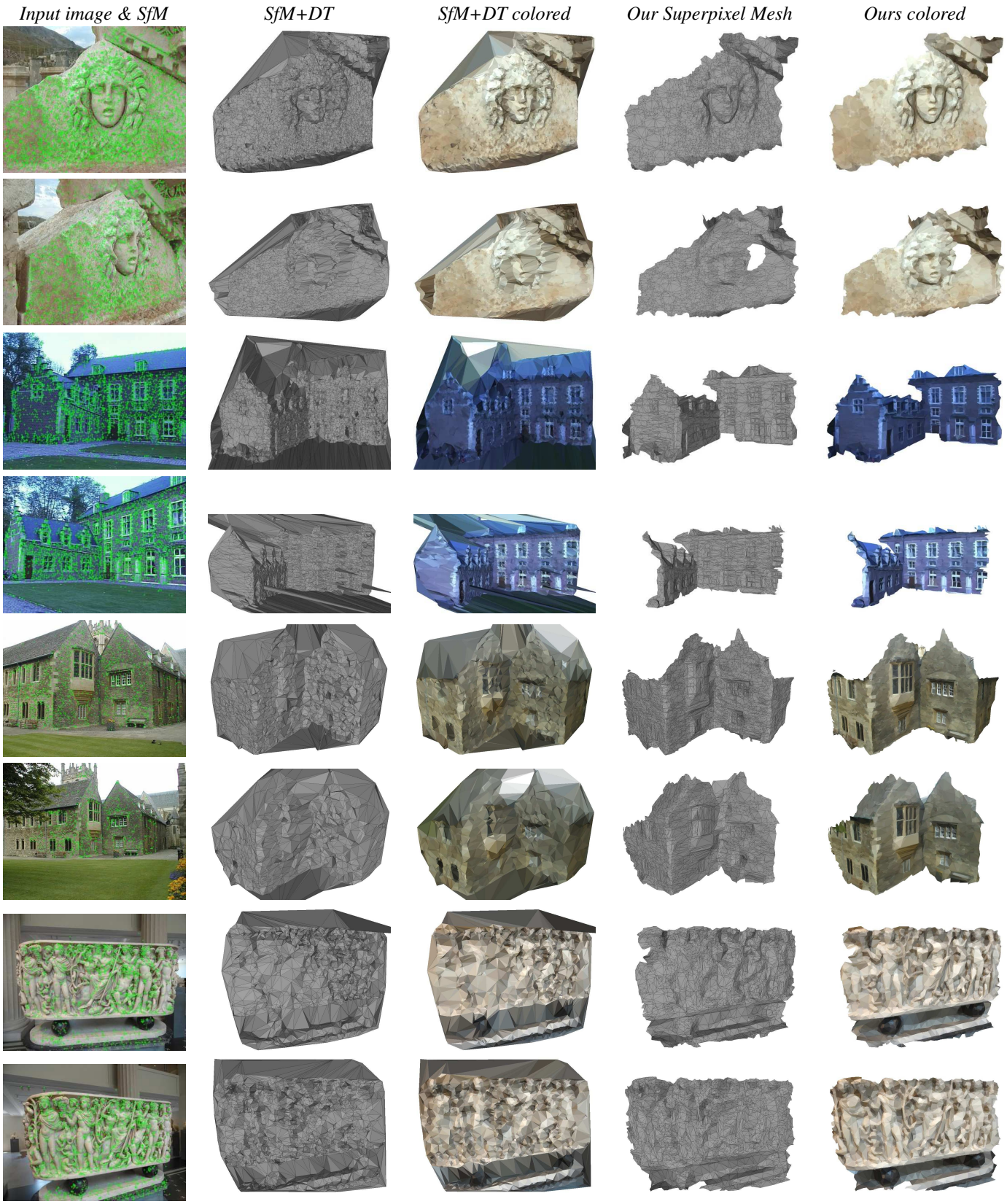

Figure 4. Results for different small and medium-size datasets compared to the SfM+DT baseline. Datasets from top to bottom: Medusa, LeuvenCastle, MertonVI, Dionysos. Columns: (1) input image with visible SfM points overlaid, (2-3) baseline SfM+DT mesh reconstructed from a single view, (4-5) our results using FS superpixels (snapshot from the exact same viewpoint). Please zoom for details.

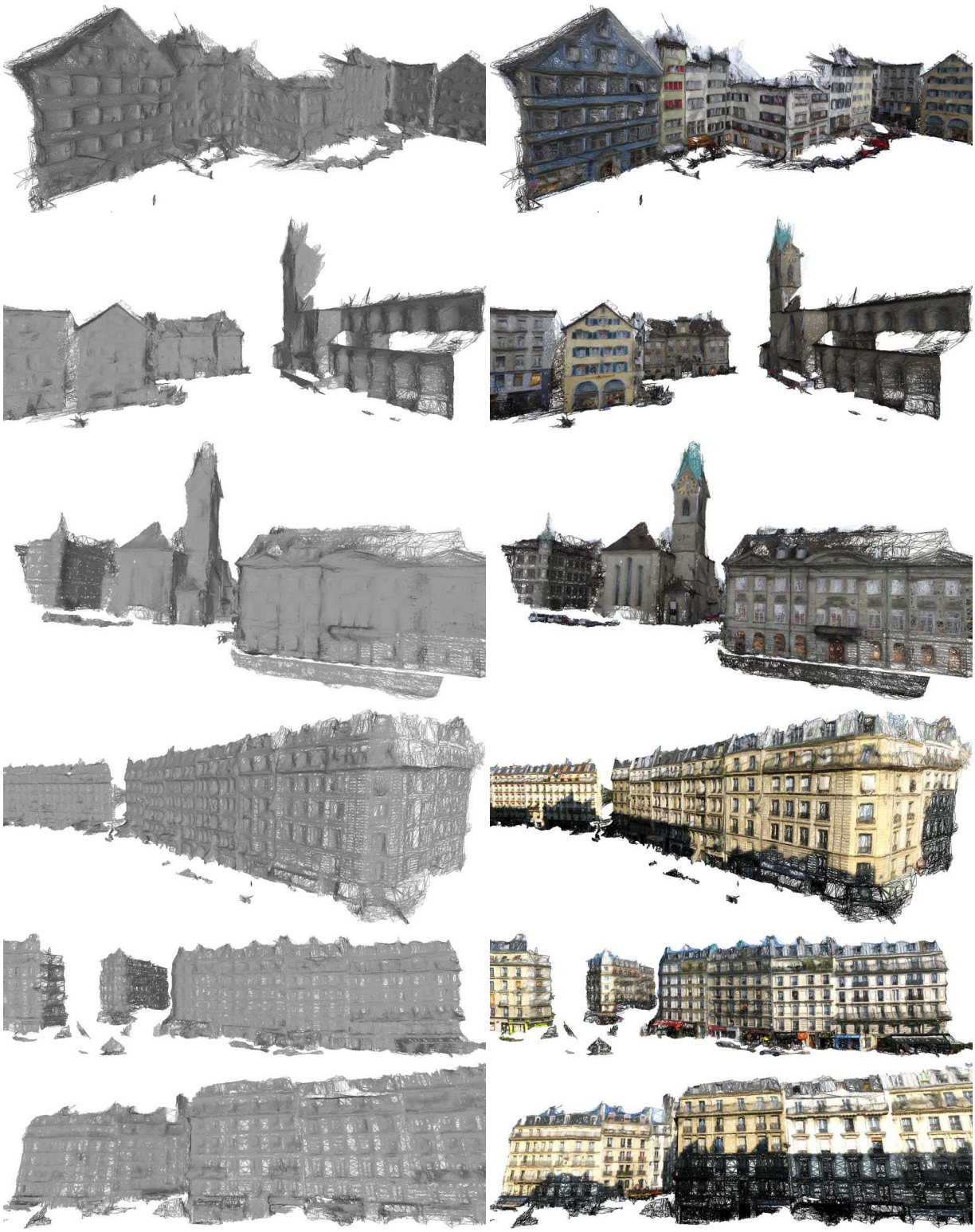

Figure 5. Some close-up views of the outputs of our method for the large-scale Street Z and Street P datasets (the paper contains snapshots of the whole datasets in Figure 9). All meshes reconstructed from individual views are displayed/rendered jointly (concatenated) without a merging step. Please zoom for more detail. Some window-level close-ups are shown in Figure 6 of the paper.