

Interactive Coherence-Based Façade Modeling Additional Material

Przemyslaw Musialski^{1,2,4}

Michael Wimmer¹

Peter Wonka^{3,4}

¹TU Vienna

²VRVis

³KAUST

⁴ASU

Appendix A: Comparison to Regular Grids

Fig. 1 shows that our system can easily model the failure case described by Müller et al. [MZWvG07] (left), where a global tiling cannot be found because different floors have different numbers of windows. Our automatic facade split, on the other hand, will put such floors in distinct clusters, and as a result they can be split into columns separately (middle). Nevertheless, all windows can be grouped together, and further subdivision can happen in synchronization (right).

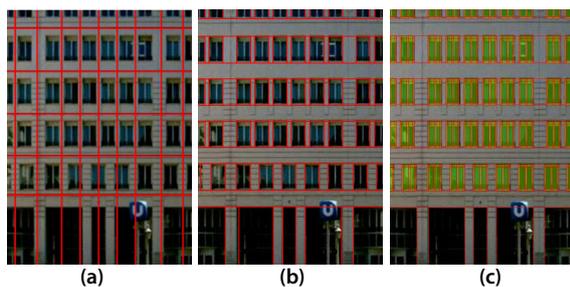


Figure 1: Failure case from Mueller et al. [MZWvG07] (a), courtesy of Mueller et al. (b) Result of our automatic facade split. (c) Result of synchronized subdivision of windows.

Appendix B: User Study



Figure 2: Facades used in the user study.

Appendix C: Precision and Recall Test

We compare the performance of the *automatic part* of our algorithm, the clustering-segmentation-based automatic split, against the edge-statistics method that forms the basis for the approaches of Lee and Navatia [LN04] and also Xiao et al. [XFT*08]. As opposed to evaluating only accuracy (i.e., percentage of elements correctly matched, also called recall), we do a precision and recall test, which also shows the precision (i.e., percentage of true positives produced by the algorithm). We conduct the test on 42 façade images, where we have removed highly unstructured parts, like the ground floors or roofs (see additional material). Precision and recall are measured with respect to a ground truth, which we generated by letting a user split the façade manually into wall and windows using several global split lines.

For the edge-statistics method, we compute canny edges in each axis direction and then count the number of edges per row or column. We draw a split line if the number of positive edges exceeds a given threshold, which is the varying variable in the shown precision-recall chart. We vary the threshold between 0% and 80% of image height or width and use non-maximum suppression. For our automatic split method, we perform the clustering and segmentation on a fixed number of clusters and vary the parameter lambda in the dynamic programming equation, which weights the smoothness term of the dynamic programming function, between 0 and 5. For each façade and each method we compute 100 samples that are equidistantly distributed in the respective parameter range. Figure 3 shows the average precision-recall response of both methods over the 42 façades with respect to the manually generated ground truth.

It can be seen that the results of our clustering-segmentation algorithm approach the ground truth. For example, when detecting 90% of the ground truth edges, the algorithm produces only 20% false positives for 3 or 4 clusters. The edge-statistics method, on the other hand, seems much harder to tune, and leads to a large number of false positives in all cases. This creates problems for the manual modeling phase,

as we shall see in the next test. Note, however, that this is not a direct comparison to Xiao et al. , as no template matching is performed, unlike in the next test.

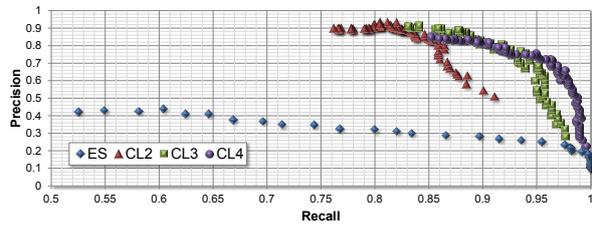


Figure 3: Precision-recall test of our façade-split method. We compare the edge-statistics method (blue diamonds) with our method run with different number of clusters (2 clusters: red triangles, 3 clusters: green squares and 4 clusters: violet circles).

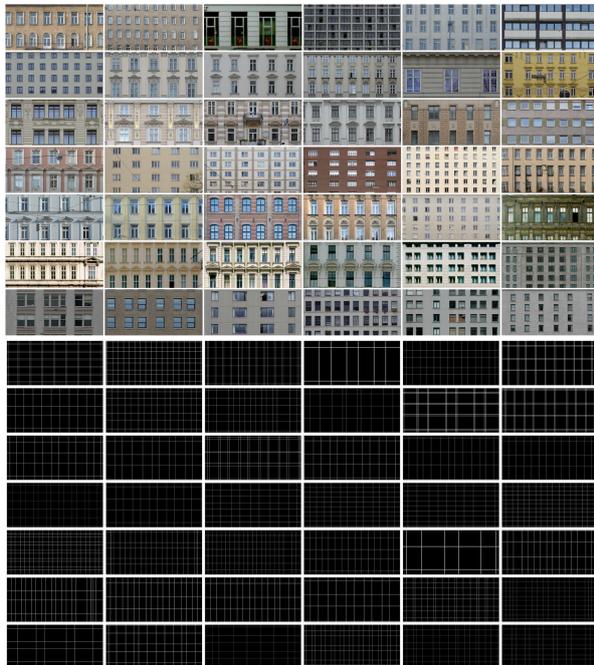


Figure 4: Top: Subset of the facade images taken to perform the precision-recall test in Section C. Bottom: the user-generated ground truth split-lines. Compare Fig. 3.

References

[LN04] LEE S. C., NEVATIA R.: Extraction and integration of window in a 3D building model from ground view images. In *2004 IEEE Computer Vision and Pattern Recognition, 2004. CVPR 2004.* (2004), vol. 2, IEEE, pp. 113–120. 1

[MZWvG07] MÜLLER P., ZENG G., WONKA P., VAN GOOL L.: Image-based procedural modeling of façades. *ACM Transactions on Graphics* 26, 3 (July 2007), 85. 1

[SHFH11] SHEN C.-H., HUANG S.-S., FU H., HU S.-M.: Adaptive partitioning of urban façades. *ACM Transactions on Graphics* 30, 6 (Dec. 2011), 1. 3, 4

[WFP10] WU C., FRAHM J.-M., POLLEFEYS M.: Detecting large repetitive structures with salient boundaries. In *Proceedings ECCV 2010* (2010), vol. 6312, pp. 142–155. 3, 4

[XFT*08] XIAO J., FANG T., TAN P., ZHAO P., OFEK E., QUAN L.: Image-based façade modeling. *ACM Transactions on Graphics* 27, 5 (Dec. 2008), 1. 1

Appendix D: Additional Modeling Results



Figure 5: Additional results: From left to right: results of [WFP10], results of [SHFH11]. Our results: elements grouped by cluster-id at a coarse level after first view auto-split clicks, our detailed results including material groups, and finally a 3d rendering. S denotes the number of shapes and T the modeling time.

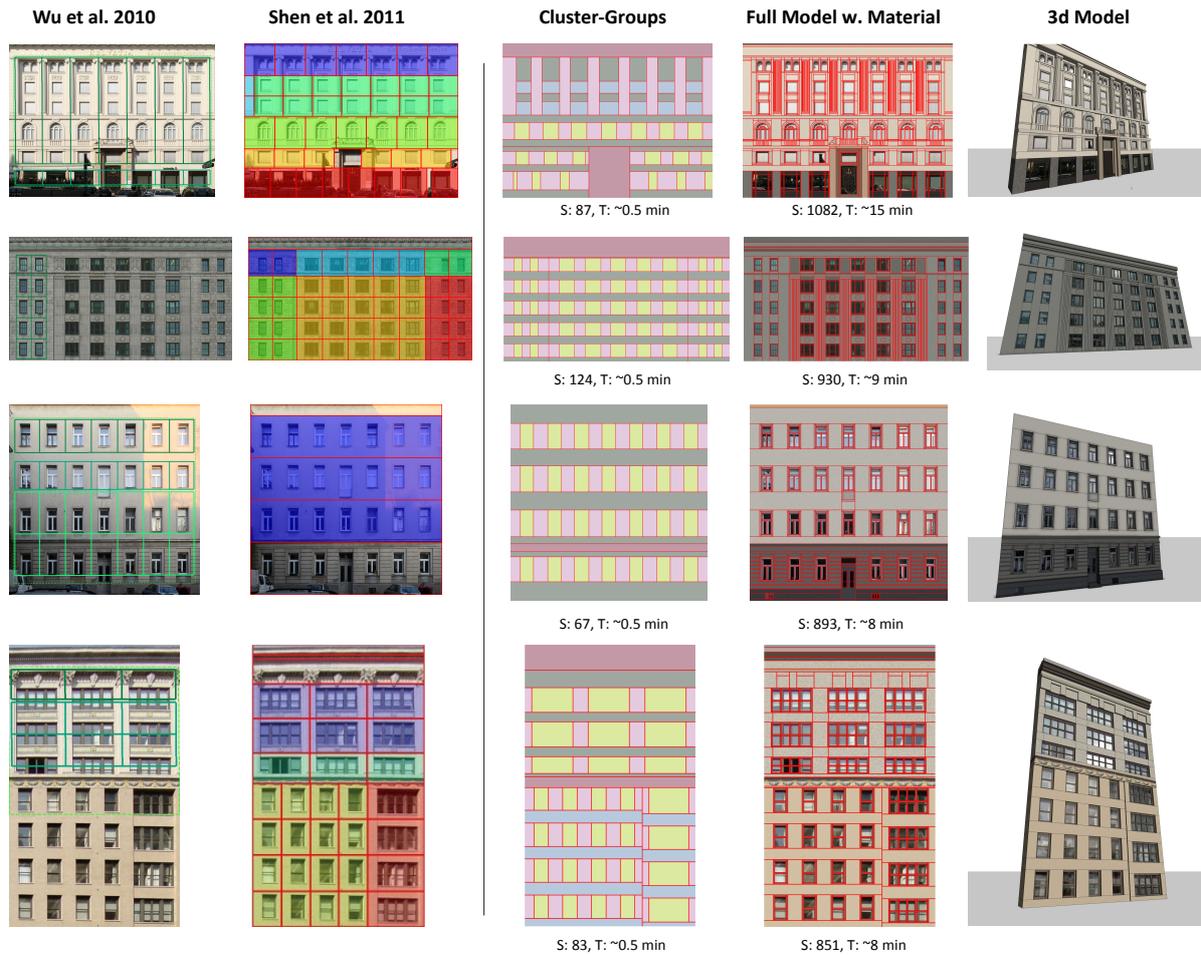


Figure 6: Additional results: From left to right: results of [WFP10], results of [SHFH11]. Our results: elements grouped by cluster-id at a coarse level after first view auto-split clicks, our detailed results including material groups, and finally a 3d rendering. S denotes the number of shapes and T the modeling time.

Appendix E: Removing of Global Illumination Gradients

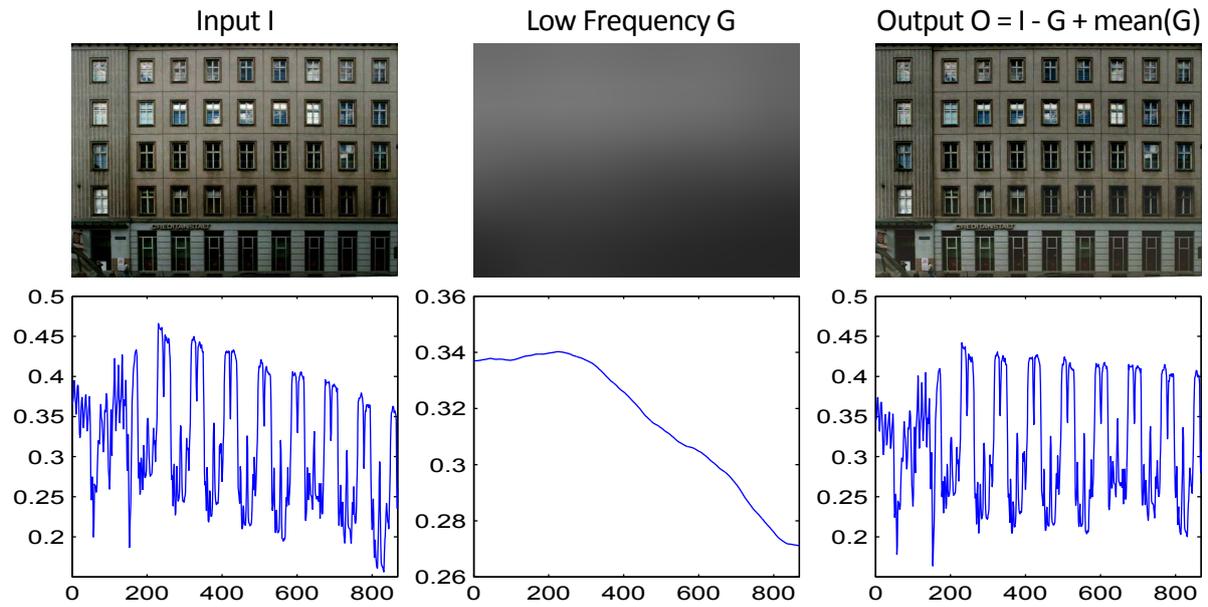


Figure 7: For each channel of the input image I , we determine its low-frequency image G by filtering I with a Gaussian with large σ . We use a kernel of the size of $\frac{w}{2} \times \frac{h}{2}$ and $\sigma = \frac{\max(w,h)}{10}$. Then we subtract the low-frequency signature from the original.