Supplementary Material: Joint Object Pose Estimation and Shape Reconstruction in Urban Street Scenes Using 3D Shape Priors

Francis Engelmann, Jörg Stückler, and Bastian Leibe

Computer Vision Group, Visual Computing Institute, RWTH Aachen University

Qualitative Results In this supplementary material we show additional qualitative results from our disparity and pose evaluation.

Videos The attached video shows libELAS reconstructions compared to our improved reconstructions. Note how our approach reduces sensor noise and yields a smooth surface even for occluded parts of vehicles.

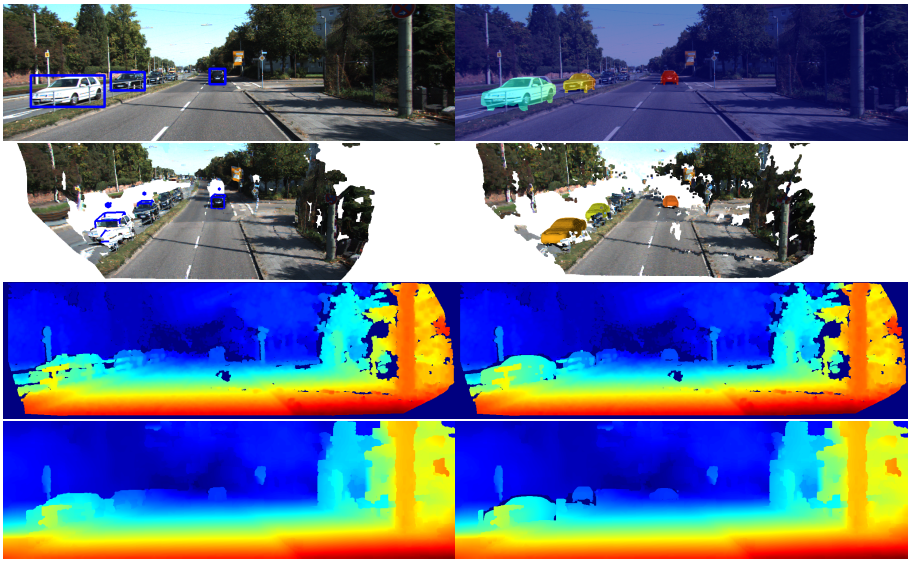

Fig. 1. Qualitative Results. From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

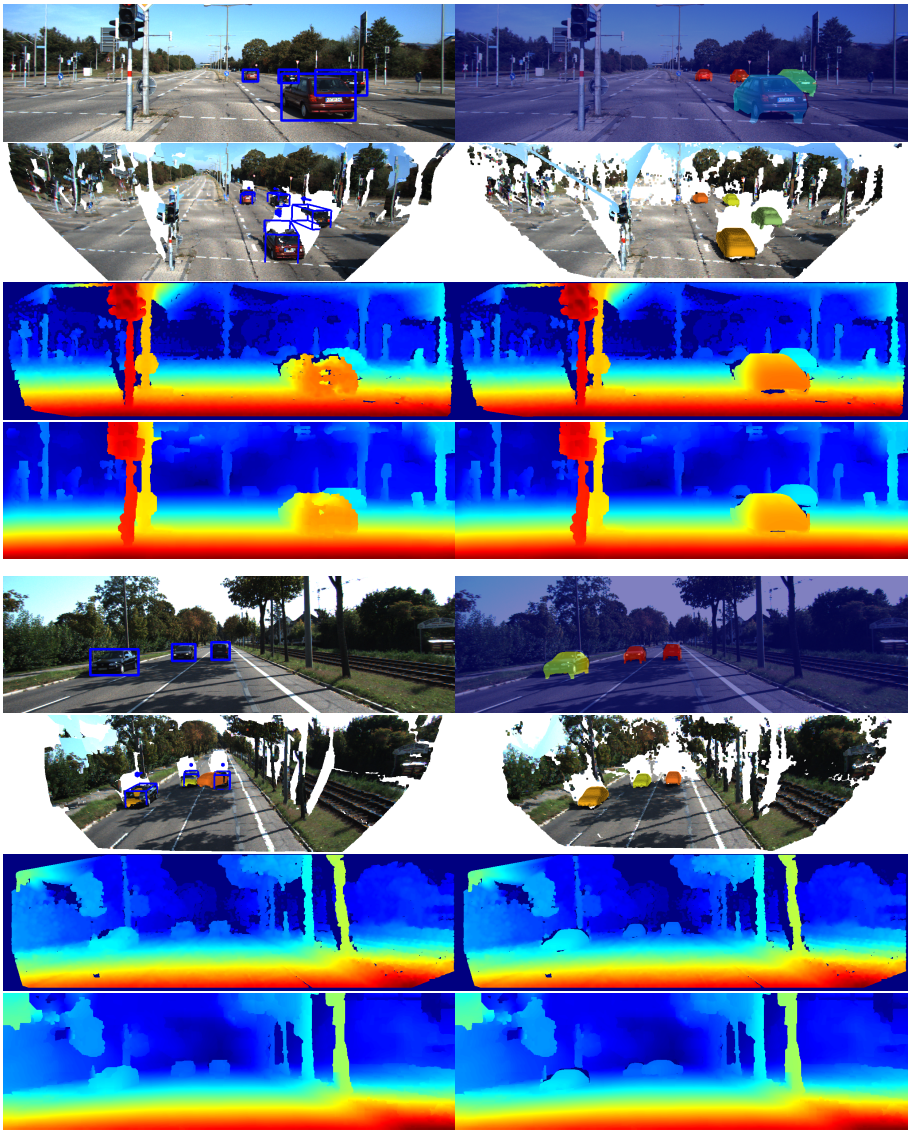

Fig. 2. Qualitative Results From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

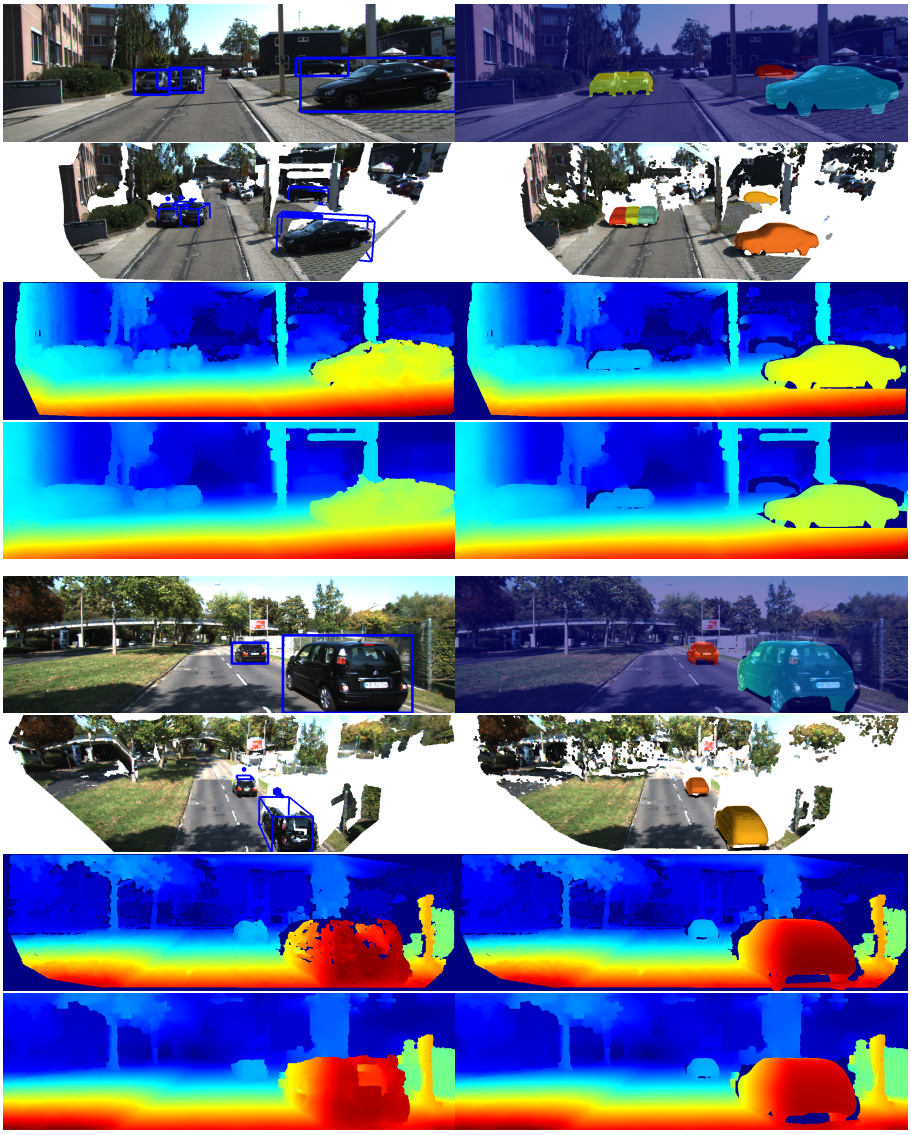

Fig. 3. Qualitative Results From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

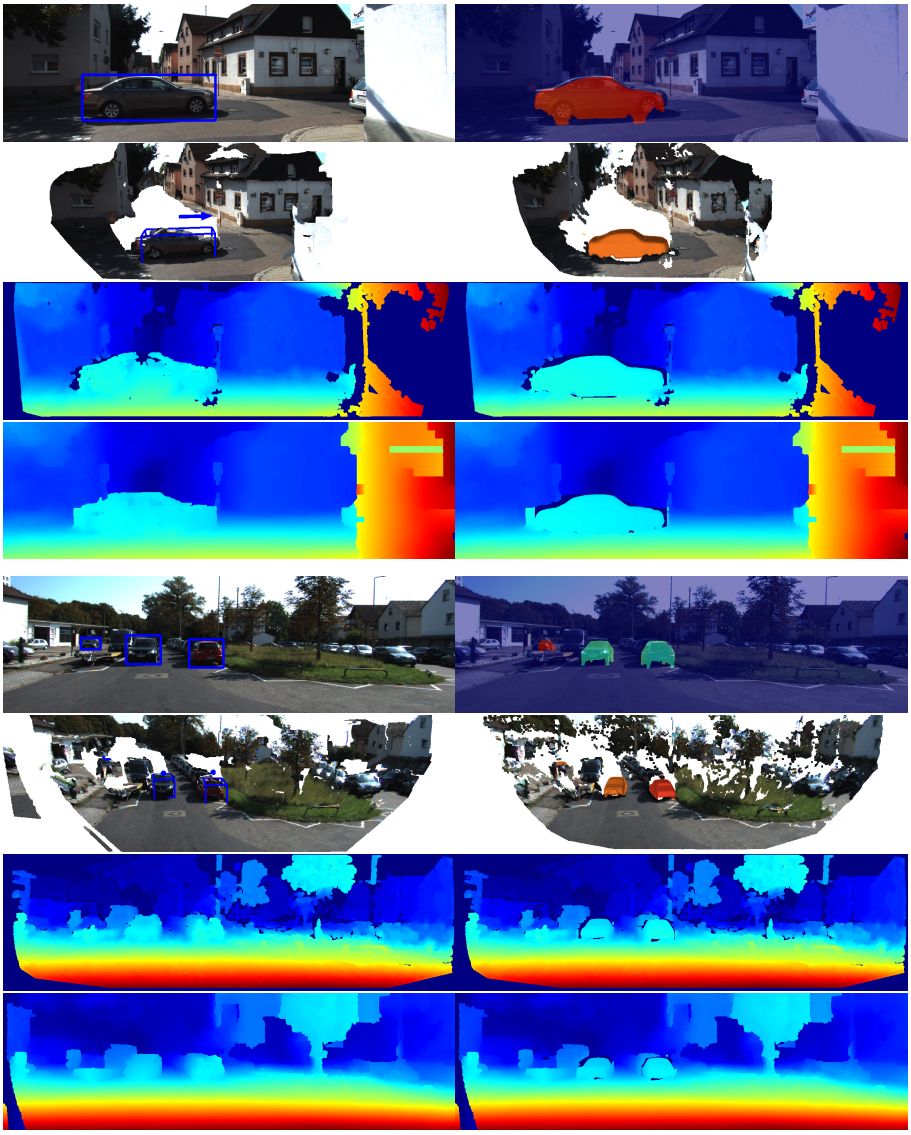

Fig. 4. Qualitative Results From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

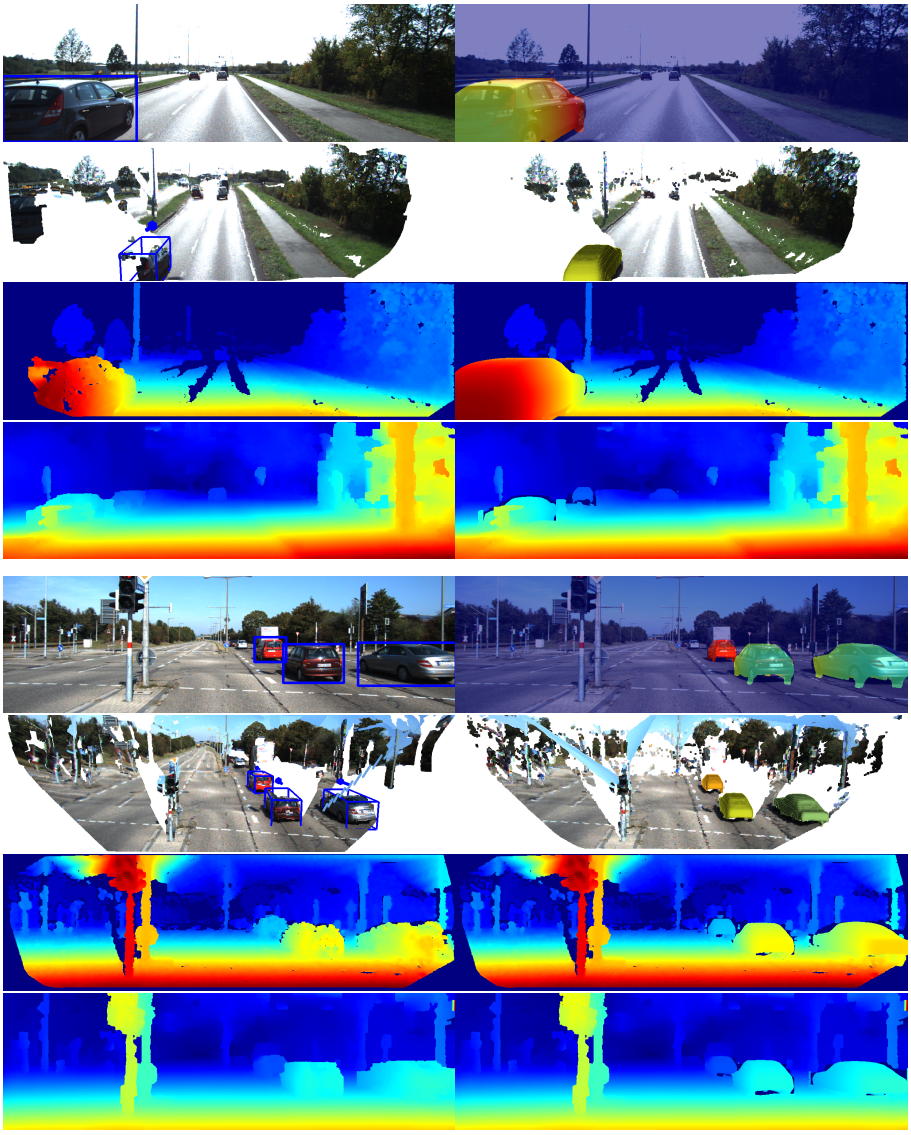

Fig. 5. Qualitative Results From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

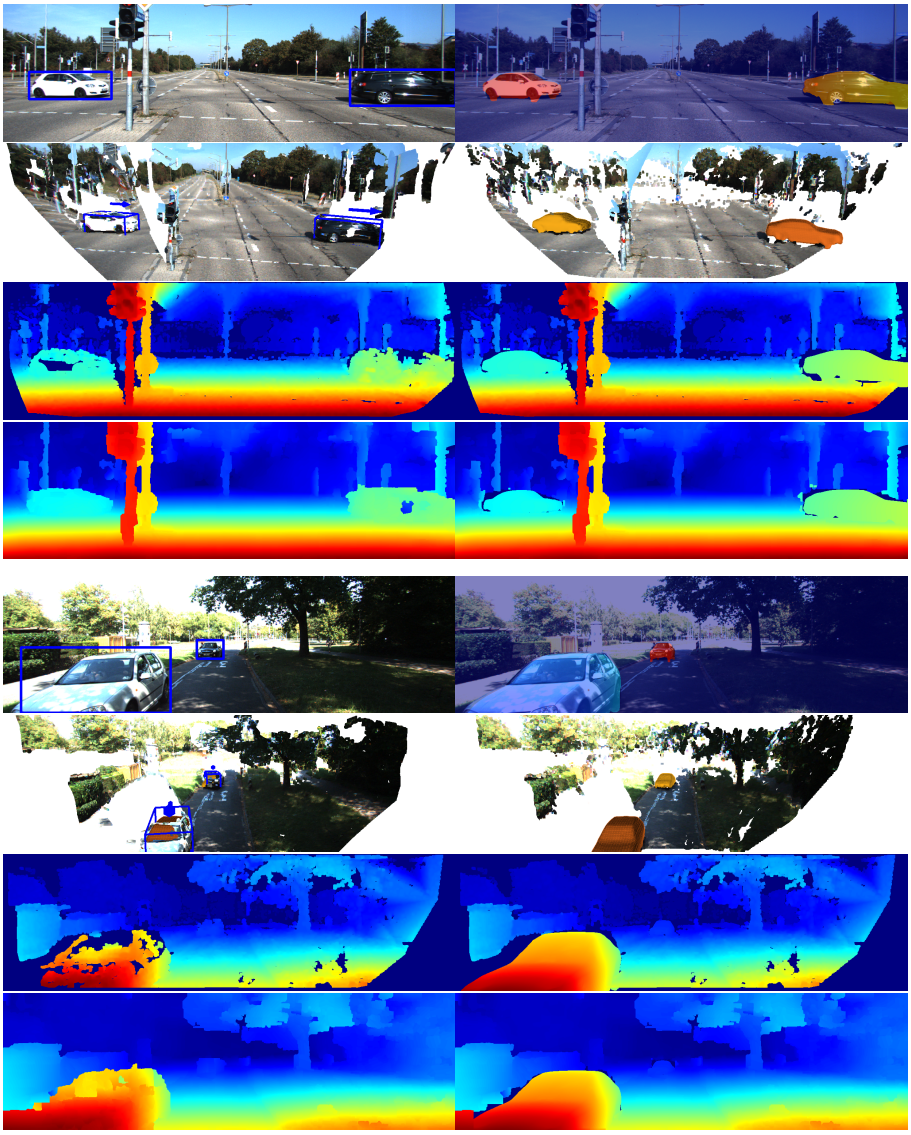

Fig. 6. Qualitative Results From top to bottom, every set of four rows shows: The input image with 2D 3DOP detections and back-projected inferred shapes (first row). The input image with 3D 3DOP detections and the 3D view of inferred shapes (second row). The libELAS disparity map with applied normal filtering as described in the paper and our improved disparity map (third row). The SPS-Stereo disparity map and our improved disparity map (fourth row).

Fig. 7. Example depth reconstruction and segmentation results for SPS-Stereo depth input. Disparity encoded from small (red) to large (blue) values. Right column highlights similar problems as for libELAS depth input, see paper for detailed discussion.

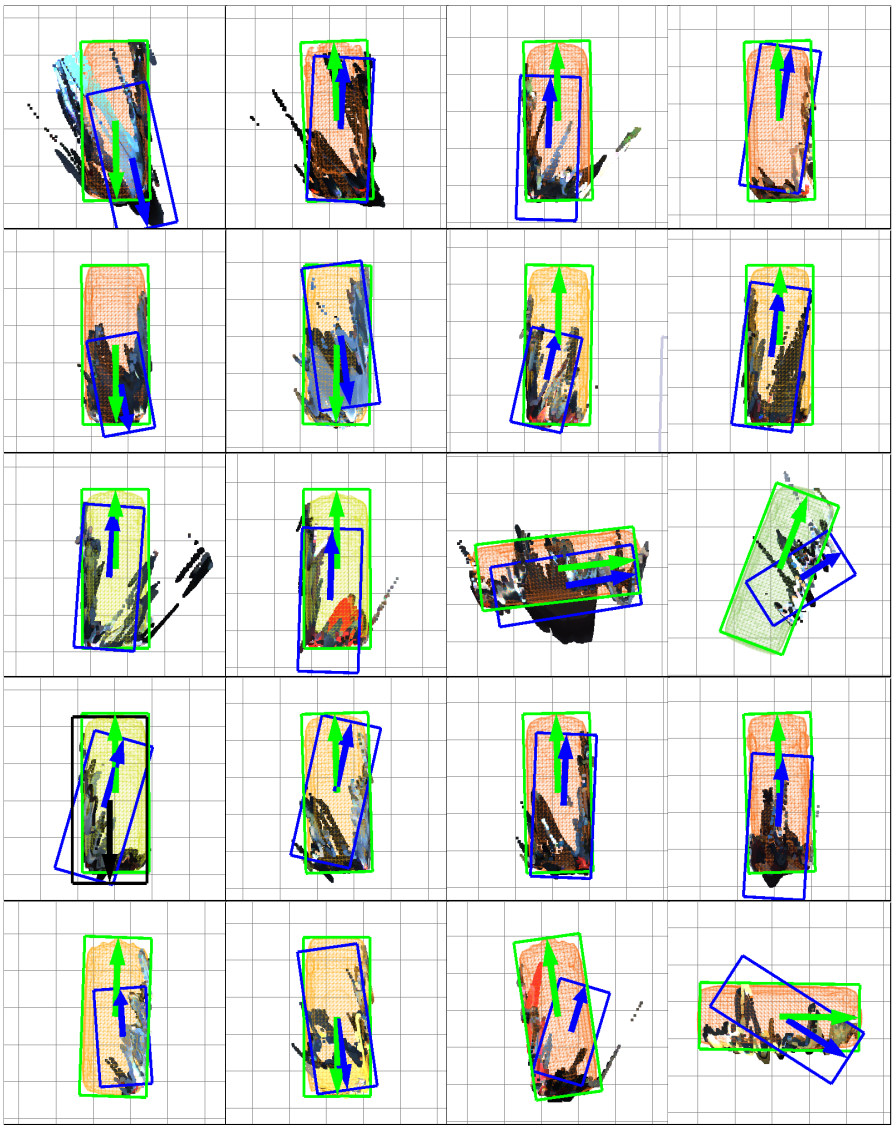

Fig. 8. Pose refinement From a bird's-eye view, we plot the 3D points assigned to the detected car. The blue bounding box corresponds to the 3DOP detection. The green bounding box represents our refined pose. The wireframe mesh represents our fitted shape. The mesh color has no significance in these plots.