

Category-Specific Object Reconstruction from a Single Image

Abhishek Kar*, Shubham Tulsiani*, João Carreira and Jitendra Malik
University of California, Berkeley - Berkeley, CA 94720
{akar, shubhtuls, carreira, malik}@eecs.berkeley.edu

1. Additional Results

We present additional reconstruction results on PASCAL VOC using both the fully automatic and fully annotated settings. As described in the main manuscript, the fully automatic setting involves reconstructing from automatically detected and segmented instances given the full image (using the technique of [1] in our case). In the fully annotated setting, we present results for the methodology described in section 4.1 where we inspect the expressiveness of our models by reconstructing the whole annotated dataset.

Figure 1 shows results using the fully annotated setting. Note that the reconstructions qualitatively look better owing to poses obtained from ground truth keypoints and clean segmentations. The reconstructed models also tend to fit the silhouettes better. This setting is directly comparable to the class-based reconstruction algorithm presented in [2]. The category “boat” is almost always bad because the large intraclass variation results in very noisy estimated cameras in the first place that leads to bad basis shape models being learnt. In some cases like “trains”, the visual hull initialization of our basis shape models (section 2.2) results in an open front as a result of the isosurface extraction failing. Given this bad initial topology, our algorithm does the best it can to explain the silhouettes in the dataset. Figure 2 presents more example reconstructions in the fully automatic setting.

Videos showing more of our results can be found at <http://goo.gl/lmALxQ>. The videos show full 3D reconstructions on PASCAL VOC using both settings above and also texture-mapped depth maps with upto 30° out of plane rotations.

References

- [1] B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik. Simultaneous detection and segmentation. In *European Conference on Computer Vision (ECCV)*, 2014. 1
- [2] S. Vicente, J. Carreira, L. Agapito, and J. Batista. Reconstructing pascal voc. *CVPR 2014*, 2014. 1

* Authors contributed equally

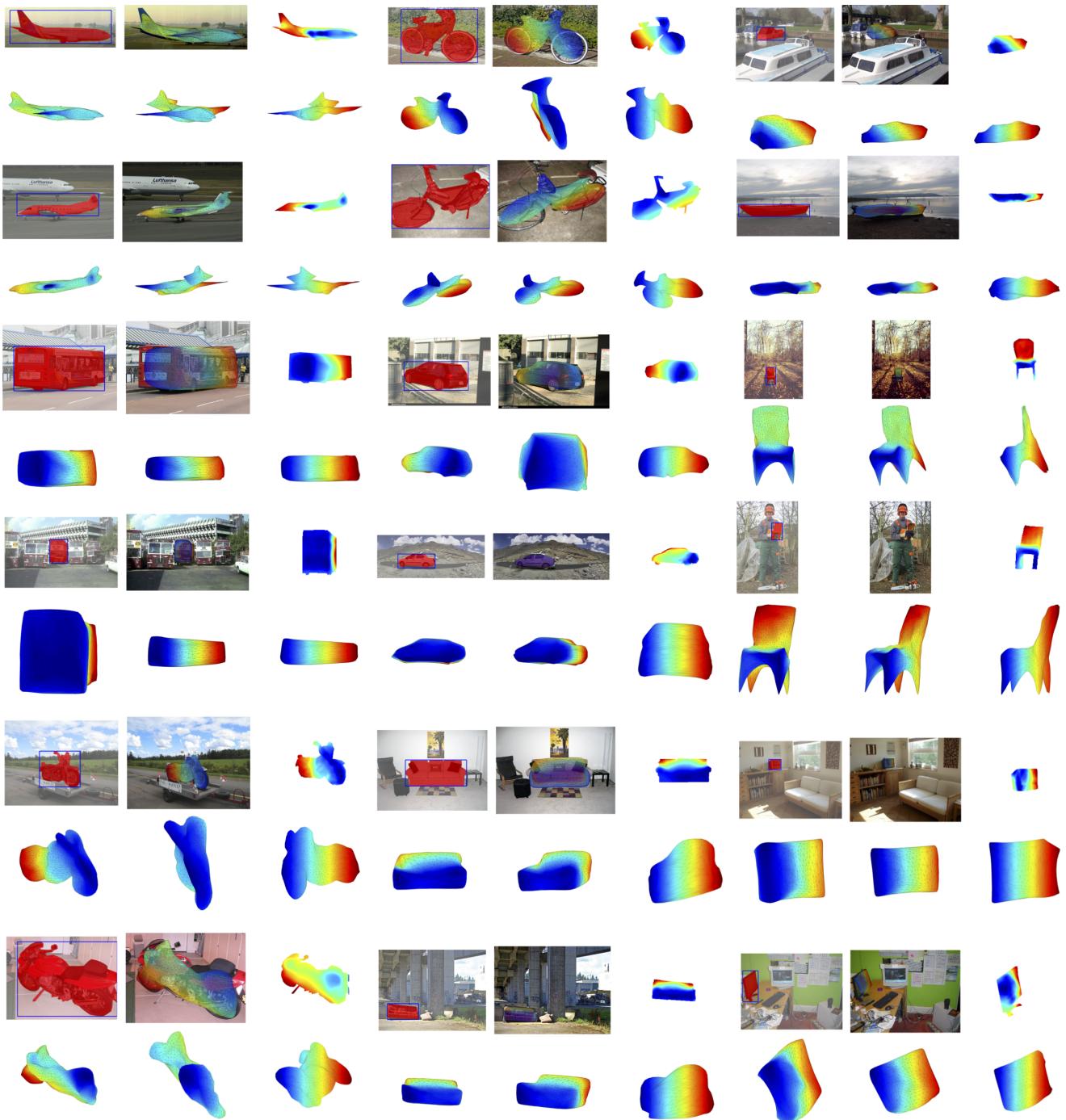


Figure 1: Reconstructions obtained using our algorithm on PASCAL VOC using the fully annotated setting (please see text for details). We show the clean segmentation input, the inferred shape overlaid on the image, a 2.5D depth map (after the bottom-up refinement stage), the mesh in the image viewpoint and two other views. Note that here we use ground truth masks and obtain pose using the ground truth keypoints. Color encodes depth in the image co-ordinate frame (blue is closer). More results can be found at <http://goo.gl/1mALxQ>

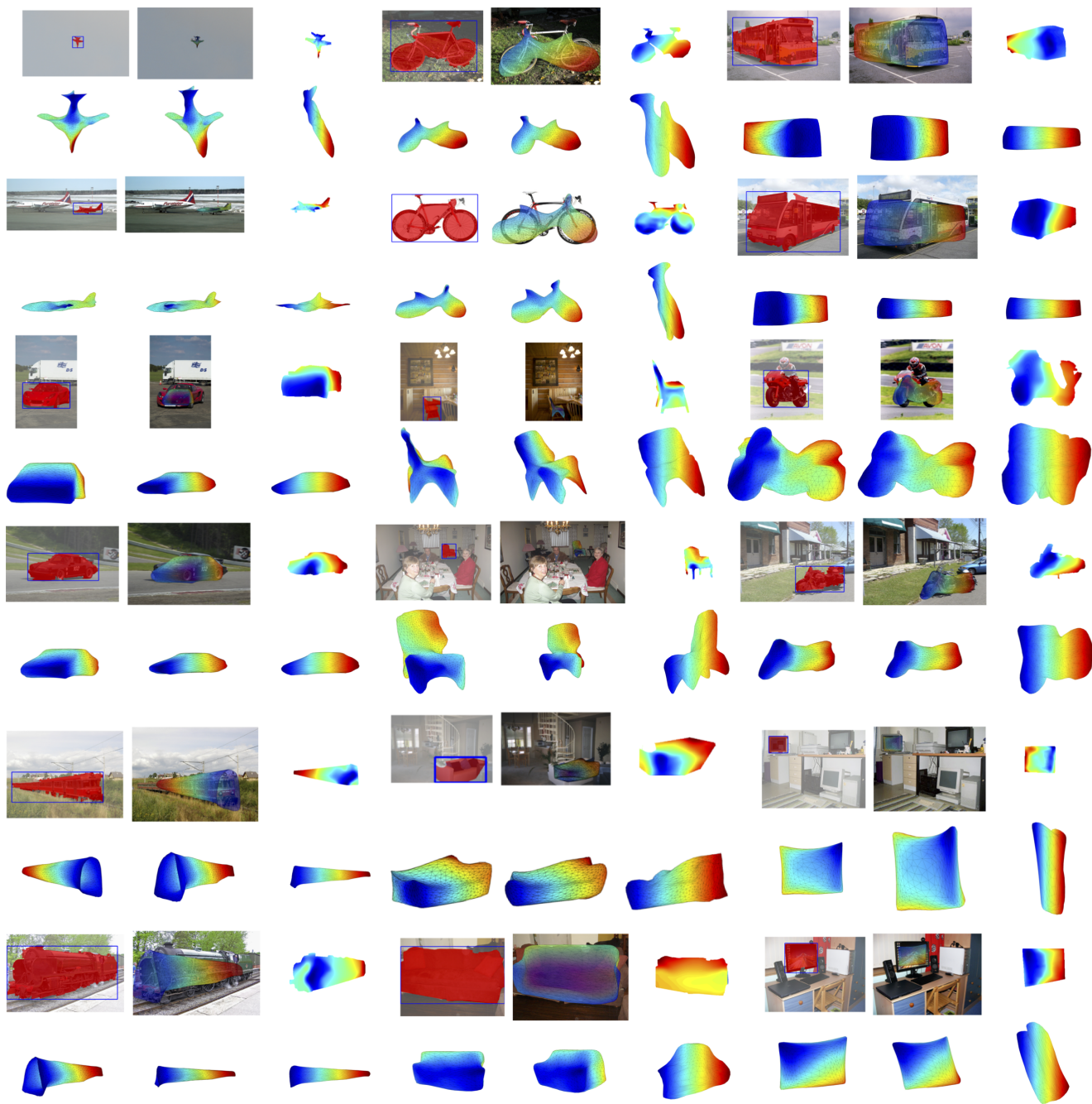


Figure 2: Fully automatic reconstructions on detected instances (0.5 IoU with ground truth) using our models on rigid categories in PASCAL VOC. We show our instance segmentation input, the inferred shape overlaid on the image, a 2.5D depth map (after the bottom-up refinement stage), the mesh in the image viewpoint and two other views. It can be seen that our method produces plausible reconstructions which is a remarkable achievement given just a single image and noisy instance segmentations. Color encodes depth in the image co-ordinate frame (blue is closer). More results can be found at <http://goo.gl/lmALxQ>