Supplemental Material: IBR of Cars using Semantic Labels and Approximate Reflection Flow

SIMON RODRIGUEZ, Université Côte d'Azur and Inria
SIDDHANT PRAKASH, Université Côte d'Azur and Inria
PETER HEDMAN, University College London
GEORGE DRETTAKIS, Université Côte d'Azur and Inria

ACM Reference Format:
Simon Rodriguez, Siddhant Prakash, Peter Hedman, and George Drettakis. 2020. Supplemental Material: IBR of Cars using Semantic Labels and Approximate Reflection Flow. *Proc. ACM Comput. Graph. Interact. Tech.* 3, 1 (May 2020), 3 pages. https://doi.org/10.1145/3384535

1 INTRODUCTION

We present additional details on the semantic segmentation training and inference, along with an illustration of the ellipsoid approximation effect when generating reflections. We also present some more detail on limitations of our approach.

2 ISOLATING CAR OBJECTS USING SEMANTIC LABELS

We start with 2D label maps for each input image, obtained with the DeepLab-v2 (Resnet-101) architecture [Chen et al. 2017]. We train it on a subset of the ADE20K dataset [Zhou et al. 2017, 2019], only selecting labels that correspond to object categories that both exhibit the regions we want to detect and are present in cityscapes. We select both object-level and part-level labels among the available ADE20K labels: car, car wheel, car window. We merge all other car related labels into a single 'car' label. All other labels are considered as background. We filter images of this dataset to only keep examples containing instances of those labels. We obtain a training set of 4000 images and a validation set of 500 images. We train our network for 300K epochs [1].

3 ELLIPSOID FITTING AND REFLECTION FLOW COMPUTATION

Effect of radii variations. Each window is approximated by an ellipsoid with longitudinal and vertical radii. Due to shape and physical constraints, the range of admissible curvatures for car windows is quite limited. We use the same range of radii for all windows ($[1m, 40m]$) and sweep it using quadratic steps to sample small values more densely, as small changes to the radii only create noticeably different reflection motion if the radii are small. This is illustrated in Fig. 1, using both axes. A side window is almost planar along an horizontal line, but quite curved along a vertical line. The radius *around the longitudinal axis* is thus small, and even small variations can lead to shifts in the reflections. On the other hand, the radius *around the vertical axis* is quite large, and even large

[1]The following hyperparameters were used for training: batch size=4, learning rate 2.5e^{-4}, momentum=0.9, learning rate decay=0.9 and weight decay=0.0005.

Authors' addresses: Simon Rodriguez, Université Côte d'Azur and Inria, simon.rodriguez@inria.fr; Siddhant Prakash, Université Côte d'Azur and Inria, siddhant.prakash@inria.fr; Peter Hedman, University College London, p.hedman.14@alumni.ucl.ac.uk; George Drettakis, Université Côte d'Azur and Inria, george.drettakis@inria.fr.

variations only cause minor shifts. This motivates our choice of a quadratic sweep schedule, while showing the accuracy required when estimating the radius around the longitudinal axis.

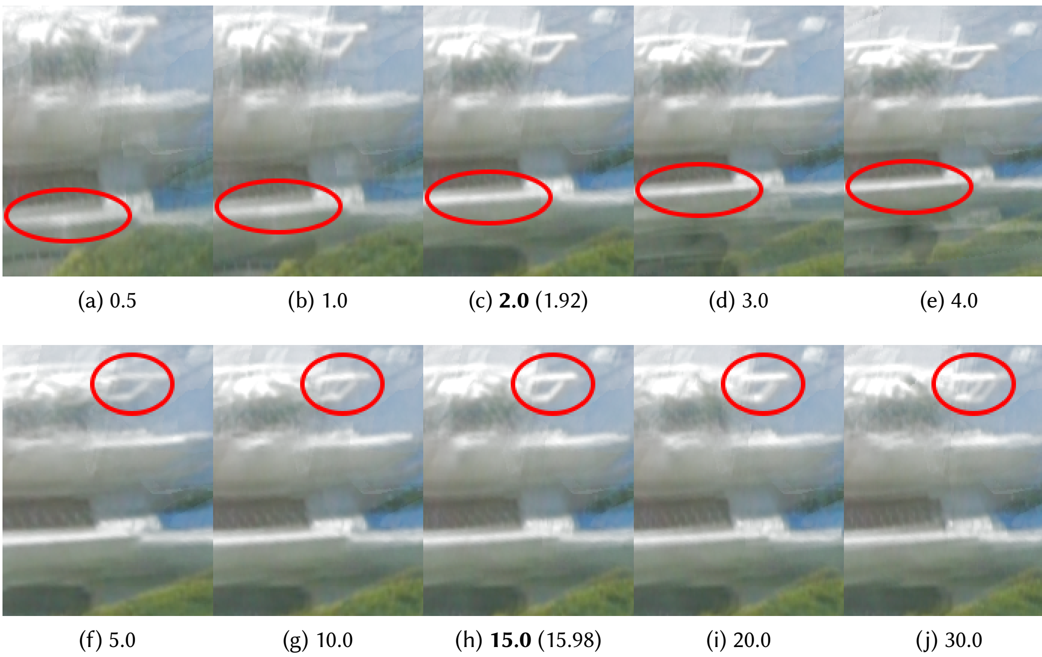

Fig. 1. Evolution of the generated reflection when sweeping the horizontal axis radius from 0.5 to 4.0 (top) and the radius around the vertical axis from 5.0 to 30.0 (bottom). The parameters closest to the radii obtained using our automatic fitting are in bold, the estimated values in parenthesis. Details such as the white horizontal line are sharper when the radii are properly estimated.

Comparison with a planar reflector. Instead of using the ellipsoid approximation, it is possible to solve the reflection flow estimation problem directly for the case of a planar reflector. Each window is is then represented by a plane going through the centroid of the window mesh and oriented by the window mean normal. The intersection with the background sphere is mirrored with respect to the plane before being reprojected in the input view. The planar surface does not manage to capture the specific reflection flow exhibited by slightly curved windows, leading to erroneous alignment between warped specular layers (see Fig. 2).

Fig. 2. A comparison between a planar (left) and ellipsoid (right) representation for each window when computing the reflection flow. The planar simplification leads to strong alignment artifacts and duplications.

4 LIMITATIONS

Our method only produces *plausible* renderings of reflections and transmission for car windows, especially for the car interior. The inaccuracy of the car interior geometry is a limiting factor for the quality of rendering we can achieve. This is shown in Fig. 3(top row) and video.

In the specific configuration where there is a strong discontinuity in a highly transmissive area (typically dark car interior over a bright background with some reflections over the dark areas) our reflective stitching method is not very successful, resulting in rendering artifacts such as ghosting of the interior or duplicated car frames (Fig. 3 (first row second column and video). This is due to the ambiguity of the min composite in this configuration (Fig. 3 and video).

Fig. 3. Top left: novel view rendering with our method. Top right: crop of the view; rendering artifacts are visible. Lower left: closest input view; there is a strong discontinuity on a bright background with some reflections present. Lower right: reflection layer min-composite in this case.

REFERENCES

Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40, 4 (2017), 834–848.

Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. 2017. Scene parsing through ADE20K dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 633–641.

Bolei Zhou, Hang Zhao, Xavier Puig, Tete Xiao, Sanja Fidler, Adela Barriuso, and Antonio Torralba. 2019. Semantic understanding of scenes through the ADE20K dataset. *International Journal of Computer Vision* 127, 3 (2019), 302–321.