KI DATA TOOLING
The Data Kit for Automotive AI

3D Object Augmentation with Photo-realistic Rendering

Paul Robert Herzog, Bosch
Haebom Lee, Bosch

3D Object Augmentation - Overview

Augmenting virtual objects in real images also known as „Mixed Reality“ is a way to leverage „free“ video footage as background material in the process of synthetic data generation. Opposed to fully synthetic content only task relevant objects need to be generated reducing rendering time and scene complexity. However, virtual objects must seamlessly integrate in the real background requiring to estimate the real scene parameters. Hence, our augmentation pipeline requires interfacing with Computer Vision algorithms (Fig. 1).

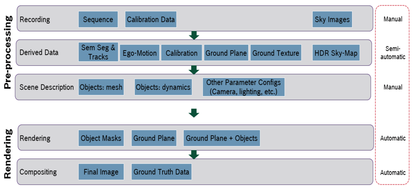

Figure 1: Overview - Augmentation Pipeline

Given the necessary scene and camera settings, we can render virtual 3D objects in front of a “virtual” proxy scene:

Figure 2: Augmented & real vehicles in one of the cameras of the Aschaffenburg sequence. Question: which are the real ones?

Interfacing Computer Vision

We need to estimate a minimal set of scene parameters:

- Camera intrinsic & extrinsic calibration,
- Ground plane,
- Environment lighting (Skymap)

Figure 3 shows some intermediate results from calibration and segmentation that are used to build a proxy background scene that is used to capture shadows and allows virtual object insertion.

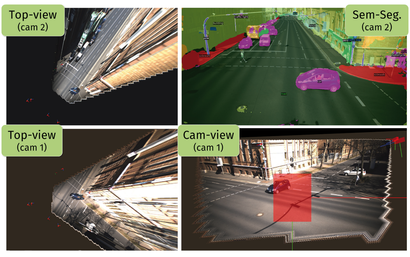

Figure 3: Selected camera views projected onto the ground plane (left), semantic segmentation and virtual camera view of proxy scene (right)

Estimating Lighting

Lighting is assumed to be outdoor lighting, and thus limited to a HDR Skymap. Lighting plays an important role, in particular for contact shadows as shown in Figure 4.

Figure 4: Random lighting vs. „correct“ lighting

We estimate static outdoor lighting using a deep neural network architecture and camera ego-motion [1].

Scene Interactions – Occlusions & Virtual Object Positioning

Where and how to place virtual objects? Semantic segmentation and a physics engine allow automatic instancing of rigid objects (e.g., here blender with bullet physics engine).

Semantic segmentation and the ground plane also allows us estimating image masks (billboards) for occluding the virtual objects.

Future Work

- Generating a dataset of augmented images for testing object detection
- Pedestrian augmentation

References:

[1] Haebom Lee et al., Spatio-Temporal Outdoor Lighting Aggregation on Image Sequences using Transformer Networks; International Journal of Computer Vision, 2022

BOSCH

For more information contact:
paulrobert.herzog@de.bosch.com
haebom.lee@de.bosch.com

KI Data Tooling is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

KI FAMILIE

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag