1.12 Plausibility Verification for 3D Object Detectors Using Energy-Based Optimization

Abhishek Vivekanandan | FZI

KI WISSEN
Automotive AI Powered by Knowledge

Motivation

Object detectors lack a built-in safety layer, raising concerns about their reliability, especially given their susceptibility to adversarial attacks that could endanger autonomous vehicles. This issue is exacerbated when uncertainty is not integrated into an end-to-end design. To mitigate this, we propose a parallel module to validate 3D object proposals from the MonoRUn[1] model, using a plausibility framework that incorporates cross-sensor data to minimize false positives. Our verification approach introduces a composite energy function based on four energy functions informed by prior knowledge, enhancing the plausibility of the predictions. Additionally, we've developed a two-step process to refine the optimization of this energy model.

Key Components

- **Chamfer Distance Energy Function**: Measures the compatibility between the point cloud observation and optimized shape vectors. Changing PCA components yields a different shape. Optimize for better shape to match the point clouds.
- **Silhouette/Differential IoU Energy Function:** Energy value is derived from the comparison between two different segmentation masks. MASK-RCNN[2] provides the initial mask, and a secondary mask is created via the projection function with the help of Chamfer Distance Energy Function.
- **Height and Rotation Energy Function:** One simple requirement for a plausible hypothesis is, for the object to be on the ground. Especially for detections of the class car, this is a strong requirement. For objects being close to the ground plane, the energy function will be close to 0, while deviations are punished quadratically. Rotation consistency checks whether the four wheels of the car touch the ground else they emit a higher energy value.

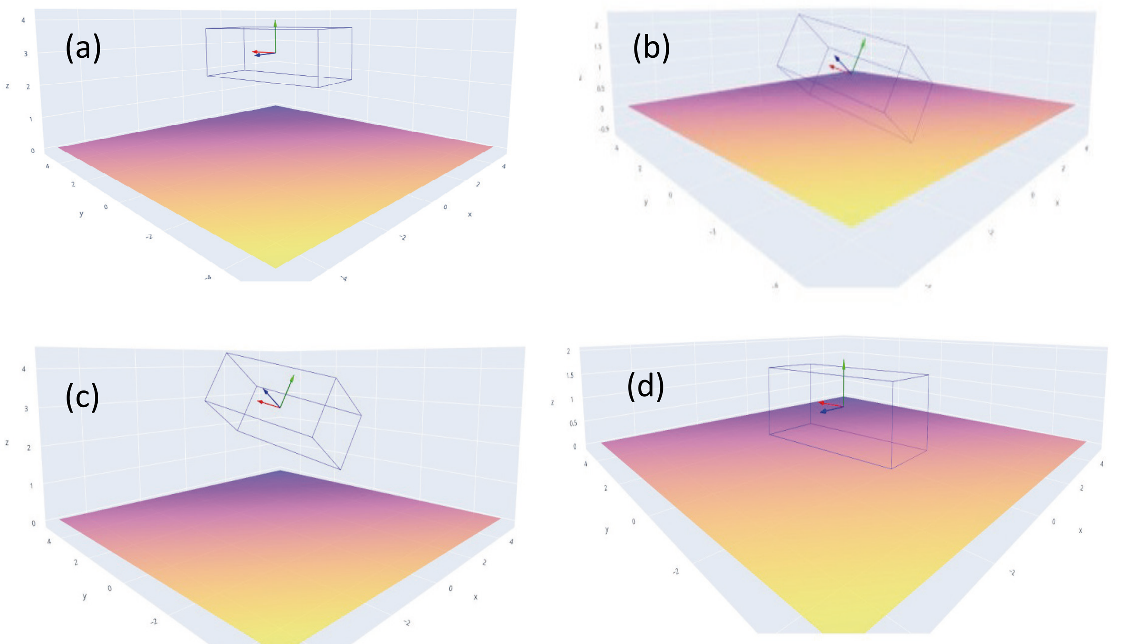

Figure 2: Height-over-Ground and Rotation energy prior-based optimization for a given ground plane and a hypothesis (blue bounding box) with random position and orientation, as shown in (a). The result of optimizing the priors individually and jointly is shown in (b), (c) and (d) (© 1. FZI)

Results

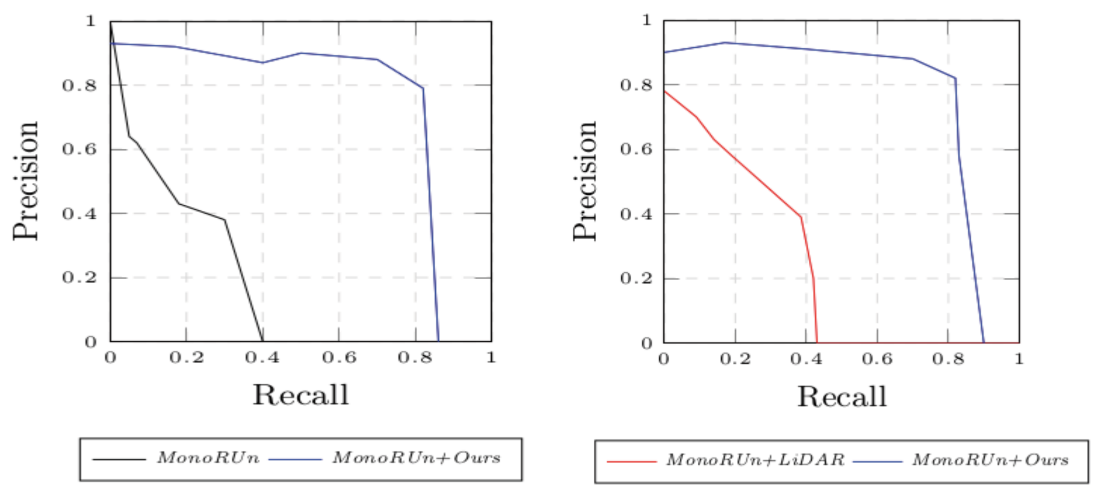

Figure 3: Quantitative evaluation of MonoRUn network and ours on KITTI dataset. The left chart shows the baseline for the network which uses no LiDAR supervision when compared with the right.

Our qualitative evaluation of the energy-based threshold filter, both with and without LiDAR supervision, reveals that the filter's Average Precision markedly outperforms that of the MonoRUn. The precision plots underscore a substantial true positive rate, affirming the effectiveness of our approach in diminishing false positives while concurrently achieving a reduced false negative.

References

[1] Chen, H., Huang, Y., Tian, W., Gao, Z., Xiong, L.: Monorun: Monocular 3d object detection by reconstruction and uncertainty propagation.
[2] He, K., Gkioxari, G., Doll´ar, P., Girshick, R.B.: Mask R-CNN. In: IEEE International Conference on Computer Vision, ICCV 2017

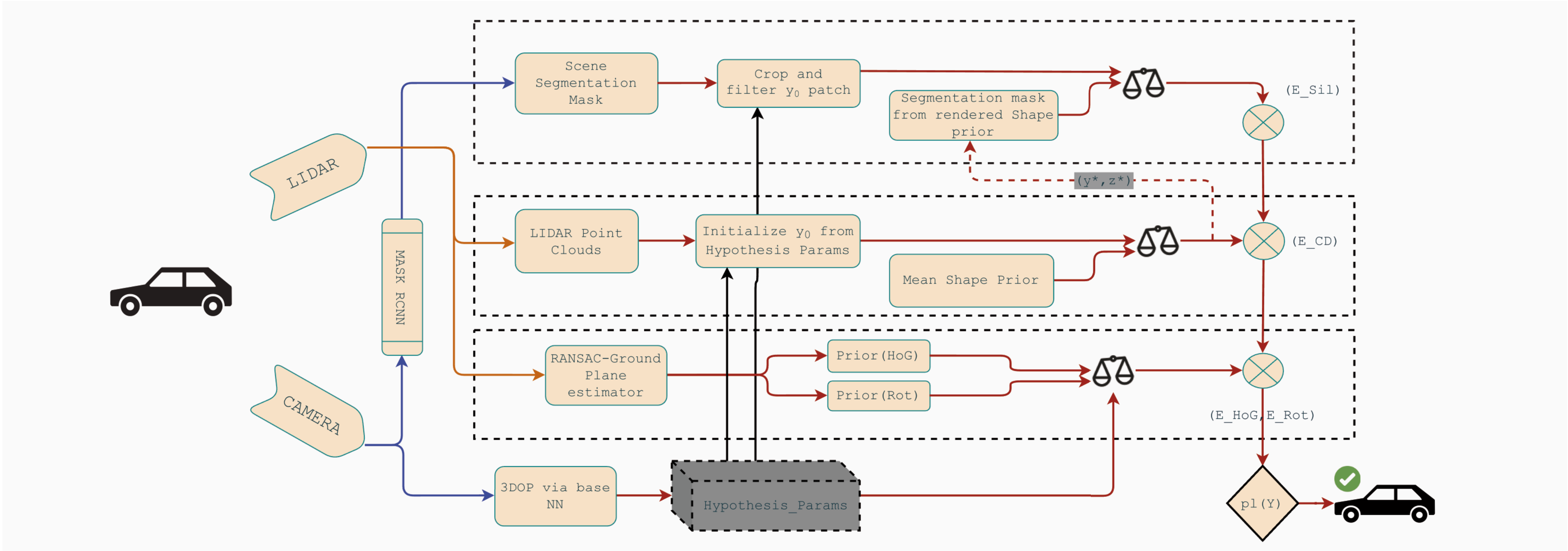

Figure 1: Parallel verification layer to obtain energy values for alternative plausibility measurement score (© 1. FZI)

Detection

Partners

External partners

For more information contact:
vivekanandan@fzi.de
hubschneider@fzi.de

KI Wissen 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.

www.kiwissen.de
@KI_Familie
KI Familie

Funded by the European Union
NextGenerationEU

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