Controllable Neural Reconstruction for Autonomous Driving

Máté Tóth, Péter Kovács, Zoltán Bendefy, Zoltán Hortsin, Tamás Matuszka

aiMotive / Budapest, Hungary

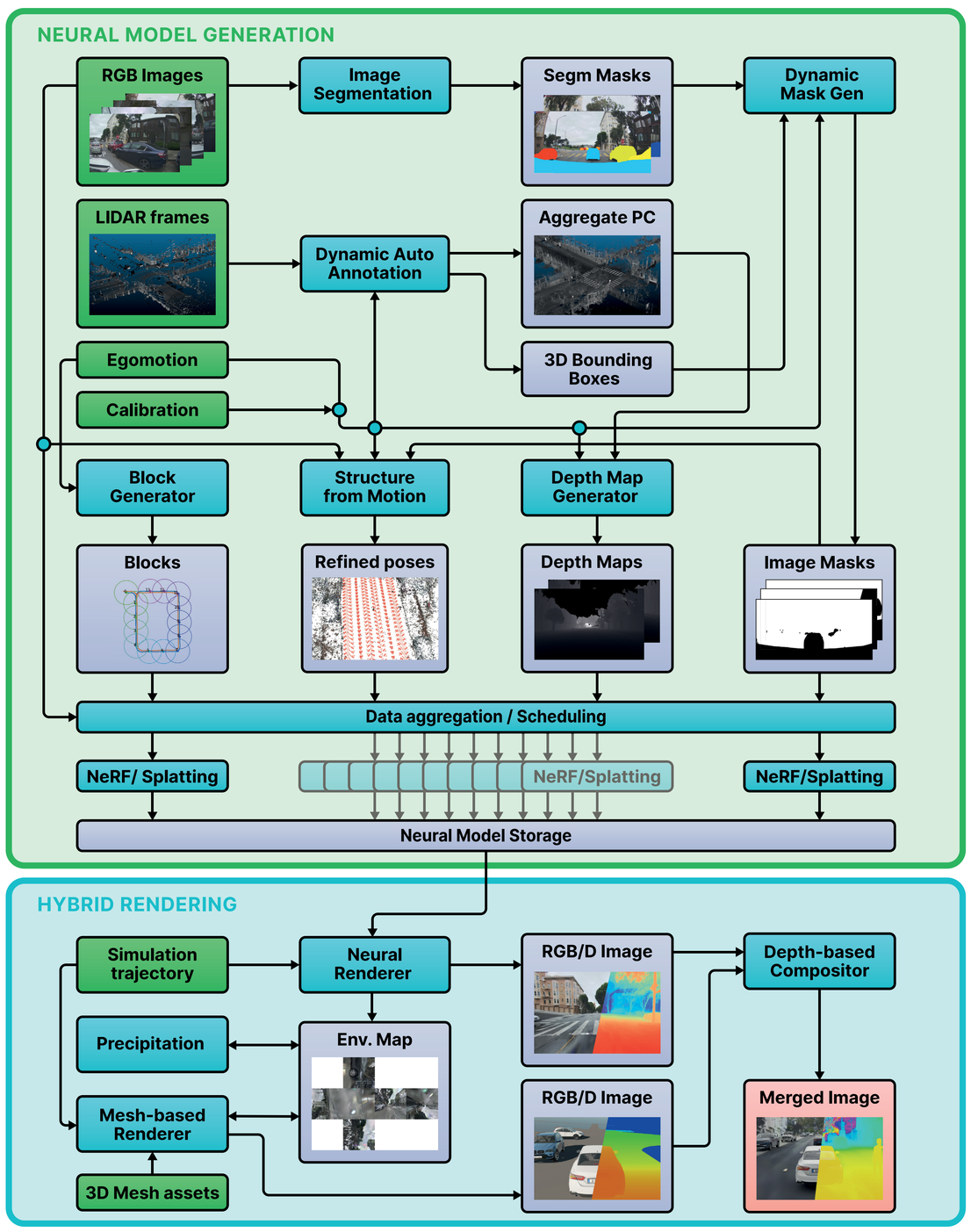
mate.toth@aimotive.com, peter.kovacs2@aimotive.com, zoltan.bendefy@aimotive.com, zoltan.hortsin@aimotive.com, tamas.matuszka@aimotive.com

EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO 2024

OVERVIEW

- We present an **automated pipeline** for training **neural reconstruction** models, utilizing sensor streams captured by a data collection vehicle.
- The reconstructed **scenes can be replayed or manipulated** in a controlled manner, using our in-house simulator.
- The recreated static environment can be **augmented with dynamic agents** with realistic occlusion, lighting and **environmental effects**.
- The demo will provide an interactive experience to the visitors where they can try our hybrid rendering solution in real time.

SYSTEM ARCHITECTURE



APPLICATIONS

- Data augmentation for training
- Open / closed loop evaluation of AD
- Sensor transfer

PREVIOUSLY PRESENTED/PUBLISHED

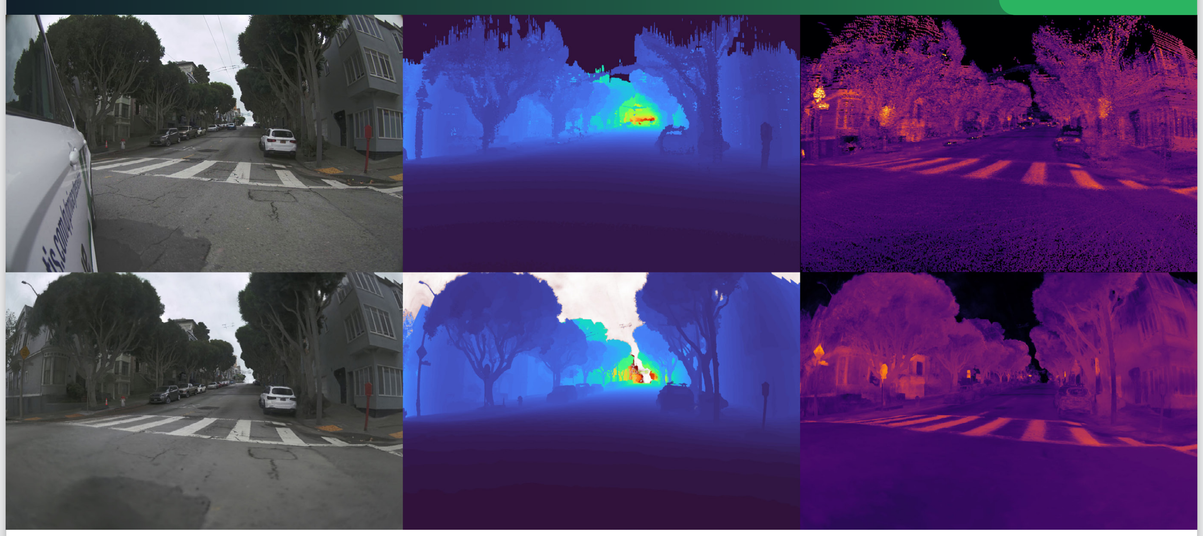
- CVPR 2024 Demo
- SIGGRAPH 2024 Poster

PROJECT PAGE

OUR APPROACH

- As input, calibrated **multi-camera recordings** covering a 360° FoV, aggregated **LiDAR point cloud**, and precise **egomotion** are used.
- **Depth and LiDAR intensity maps** are generated from the point clouds for regularization.
- **Non-stationary objects are masked out**, using 3D bounding boxes.
- Customized **Nerfacto and Splatfacto models** are trained using a block-based approach.
- Mesh-based objects are lit by **image-based lighting**, and added to the background using **depth compositing**.

RESULTS



MULTI-MODALITY RENDERINGS

- Accurate **RGB images, depth and LiDAR intensity maps can be rendered** from arbitrary camera poses as seen above with **GT in the first row**.

CAMERA VIRTUALIZATION

- Arbitrary **virtual camera setups can be simulated**, including different camera alignments and models.
- The figure above shows **simulated** front **fisheye** (left), front **wide angle** (middle) and front **long range** (right) camera renders from a model trained without direct front cameras.

DYNAMIC CONTENT INSERTION

- **Dynamic objects** can be added with realistic **lighting**, and **ambient occlusion**.
- Environmental effects, such as **rain**, **snow** and **fog** can be added for more diverse simulated scenarios.