EMIE-MAP: Large-Scale Road Surface Reconstruction Based on Explicit Mesh and Implicit Encoding (Supplementary Materials)

Wenhua Wu[1], Qi Wang[2], Guangming Wang[3],
Junping Wang[4], Tiankun Zhao[4], Yang Liu[4], Dongchao Gao[4],
Zhe Liu[1]*, and Hesheng Wang[2]*

[1] MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University
[2] Department of Automation, Shanghai Jiao Tong University
[3] University of Cambridge
[4] Hozon New Energy Automobile Co., Ltd
{2142431167wwh, liuzhesjtu, wanghesheng}@sjtu.edu.cn

1 Overview

In this document, we present more details and several extra results. In Sec. 2, we elaborate on the implementation details of our method. In Sec. 3, we present the road surface reconstruction results for additional scenarios.

2 Further Implementation Details

In this section, we provide additional detailed implementation information.

2.1 Hyperparameters

The resolution of the road surface mesh is set to 0.1m, with a mesh range extending 15m on each side of the vehicle trajectory. Each vertex stores coordinates p, semantics sem, and implicit color encoding l_c. The coordinates are in three dimensions (x, y, z), semantics include five categories: lane marking, curb, manhole, road, and background. The color encoding has a dimension of 16. (x, y) is encoded into a 22-dimensional coordinate vector. The elevation residual MLP consists of eight layers with a width of 128, using the ReLU activation function. The color MLPs consist of two layers with a width of 16, with the ReLU activation function in the intermediate layer and the Sigmoid activation function in the output layer.

The parameters to be optimized include the parameters of the elevation residual MLP and color MLPs, as well as the semantic and color encoding. We use the Adam [1] optimizer to optimize these parameters. The learning rate for the

* Corresponding author.

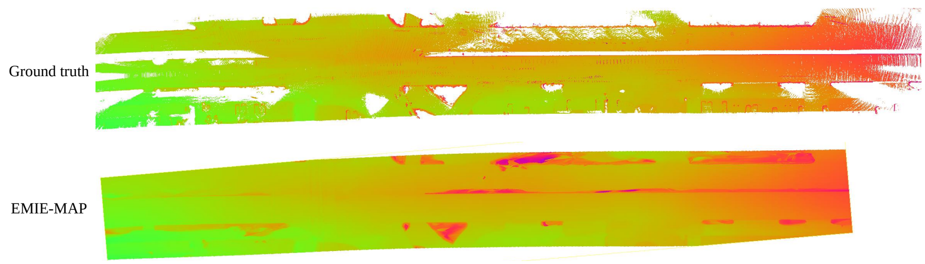

Fig. 1: Elevation Evaluation. The top image shows the ground truth of the road surface obtained by concatenating the LiDAR point clouds, while the bottom image displays the road elevation map reconstructed using EMIE-MAP. For LiDAR points within the reconstruction area, we calculated the average distance from the LiDAR points to the reconstructed road surface as Elev-error.

elevation residual MLP is set to 0.01, while the learning rate for the color MLPs is set to 0.005. The semantic learning rate is set to 0.1, and the learning rate for the color encoding is set to 0.005. The loss weights are set as $\lambda_{rgb} = 1.0$, $\lambda_{sem} = 1.0$, $\lambda_z = 1.0$, and $\lambda_{smooth} = 1.0$. For each scene, a total of 5 epochs are trained with a batch size of 8.

Due to the sufficient coverage of the road surface by the front-facing wide-angle camera and the front-left and front-right cameras, we only utilize images from these three perspectives for the reconstruction. All experiments are conducted on a server equipped with an NVIDIA A100 GPU.

2.2 Evaluation Metrics

For road surface color and semantics, we project the reconstructed road surface onto the perspective of each camera to obtain rendered images. We evaluate the results using the Peak Signal-to-Noise Ratio (PSNR) for color fidelity and mean Intersection over Union (mIoU) for semantic segmentation accuracy. For road surface elevation, we evaluate the performance by calculating the average distance between the Lidar ground points and the reconstructed road surface, defined as Elev-error.

We performed object detection on each frame of the LiDAR point cloud to filter out vehicles and pedestrians on the road surface. Then, we concatenated all the LiDAR point clouds together based on vehicle trajectories to generate a local LiDAR map as the ground truth for the road surface. Fig. 1 shows the ground truth of the road surface in the city street scene and the elevation map reconstructed using EMIE-MAP. For LiDAR points within the reconstruction area, we calculated the average distance from the LiDAR points to the reconstructed road surface as Elev-error.

3 More Scene Road Surface Reconstruction Results

In this section, we provide road surface reconstruction results for more diverse scenarios, including highways, tunnels, curve, early morning and night-time scenes. The dataset is sourced from our data collection vehicles. The public dataset used in the RoMe consists of flat scenes, with consistent color observation from different cameras. However, our surround-view cameras vary in model, resulting in inconsistent colors across different views, significantly increasing the challenge of this task. The results demonstrate the versatility of our algorithm, which achieves good reconstruction performance in various complex environments. Our method is capable of reconstructing RGB maps obtained from different camera models and semantic maps, and it accurately captures road surface elements such as lane lines and arrows. From the elevation maps, it is evident that our road surface reconstruction closely matches the height information from the LiDAR map. Our method not only effectively models changes in slope along the road but also captures height variations in details such as curbs. Compared to LiDAR maps, our reconstruction results provide a dense road surface mesh with color and semantic information, making them more suitable for downstream tasks such as BEV perception and data annotation.

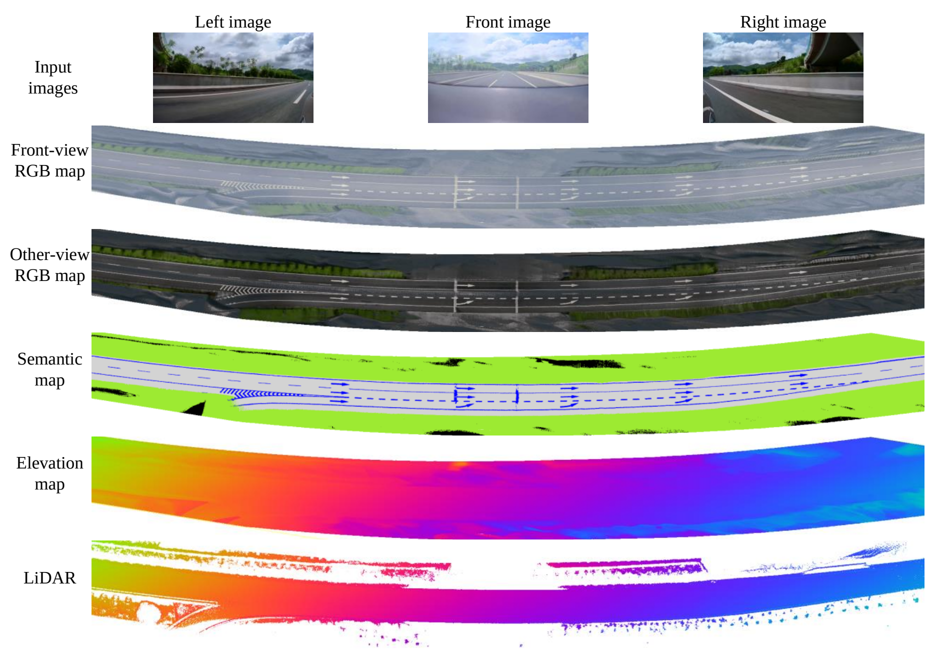

Fig. 2: Road surface reconstruction results of EMIE-MAP in a highway scene.

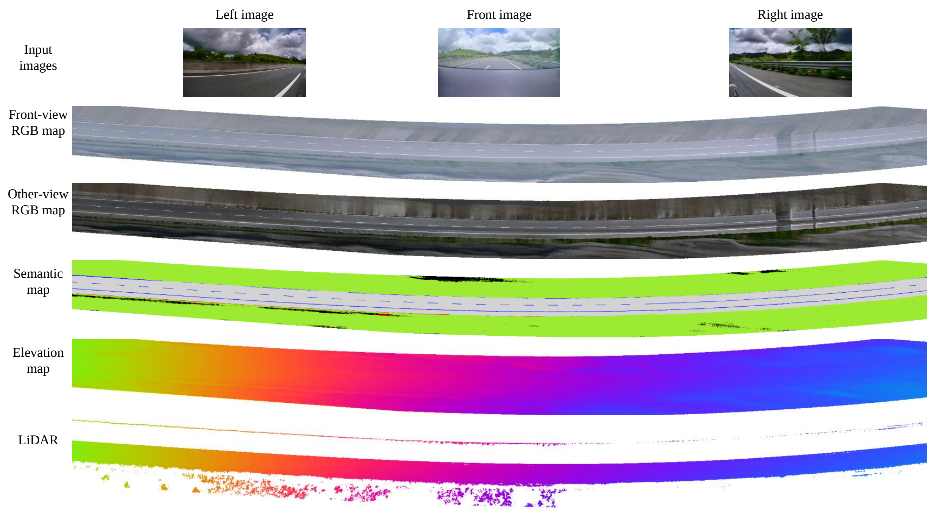

Fig. 3: Road surface reconstruction results of EMIE-MAP in a highway scene.

References

1. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

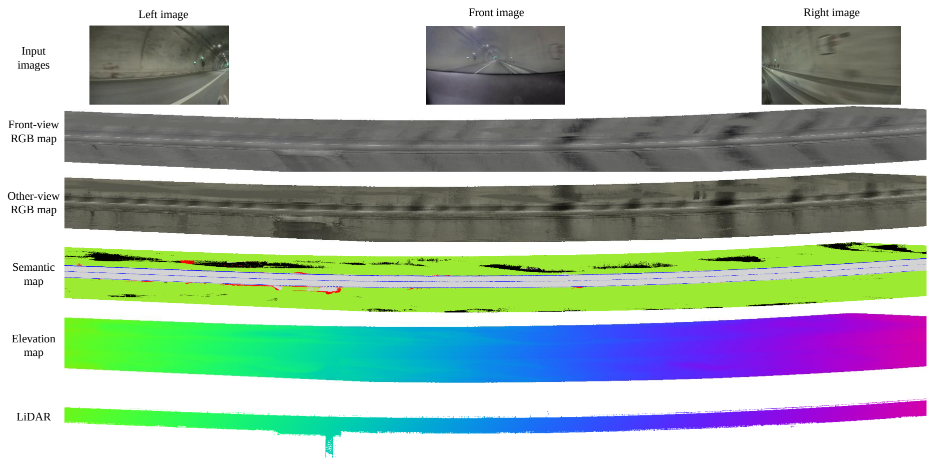

Fig. 4: Road surface reconstruction results of EMIE-MAP in a tunnel scene.

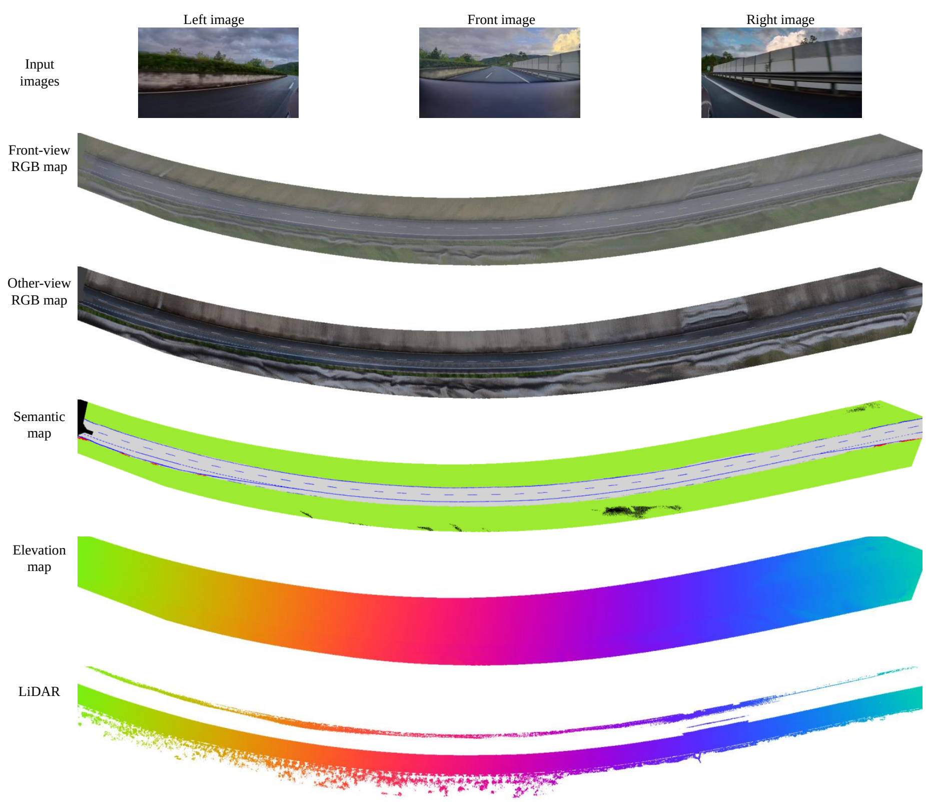

Fig. 5: Road surface reconstruction results of EMIE-MAP in a curve scene.

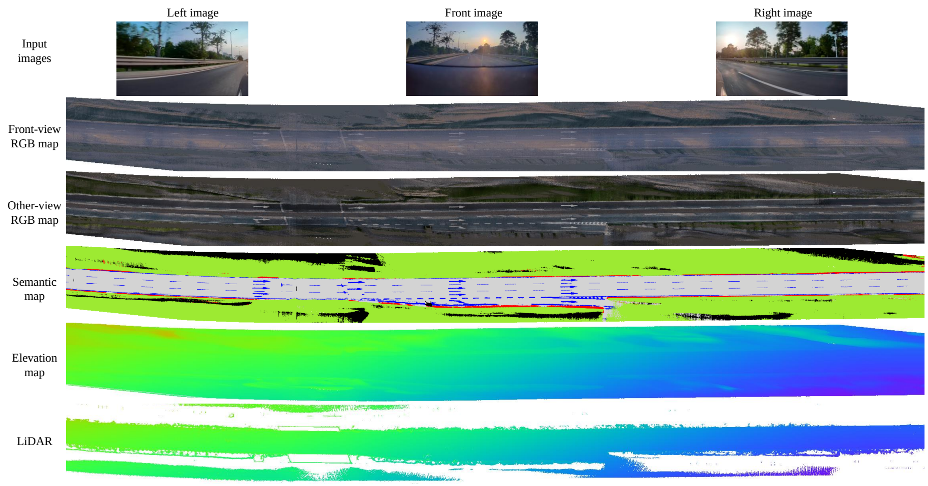

Fig. 6: Road surface reconstruction results of EMIE-MAP in a early morning scene.

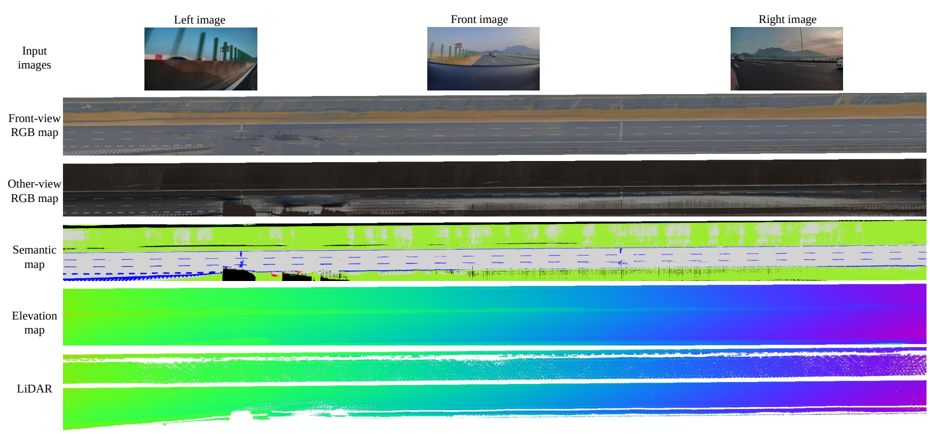

Fig. 7: Road surface reconstruction results of EMIE-MAP in a early morning scene.

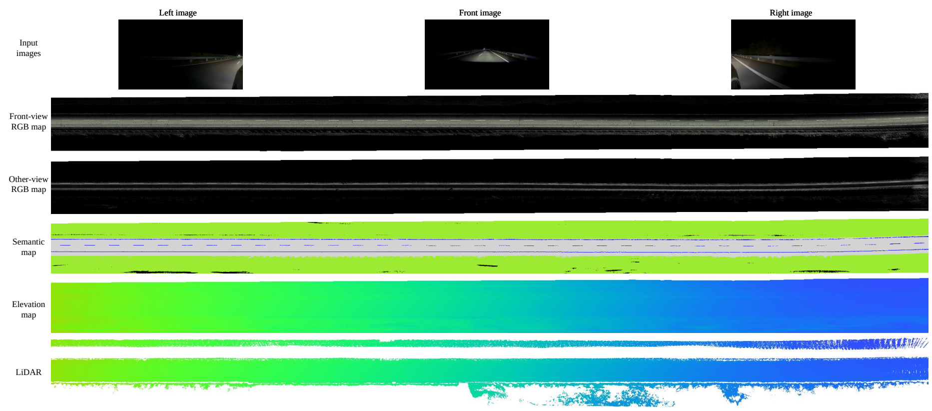

Fig. 8: Road surface reconstruction results of EMIE-MAP in a night scene.

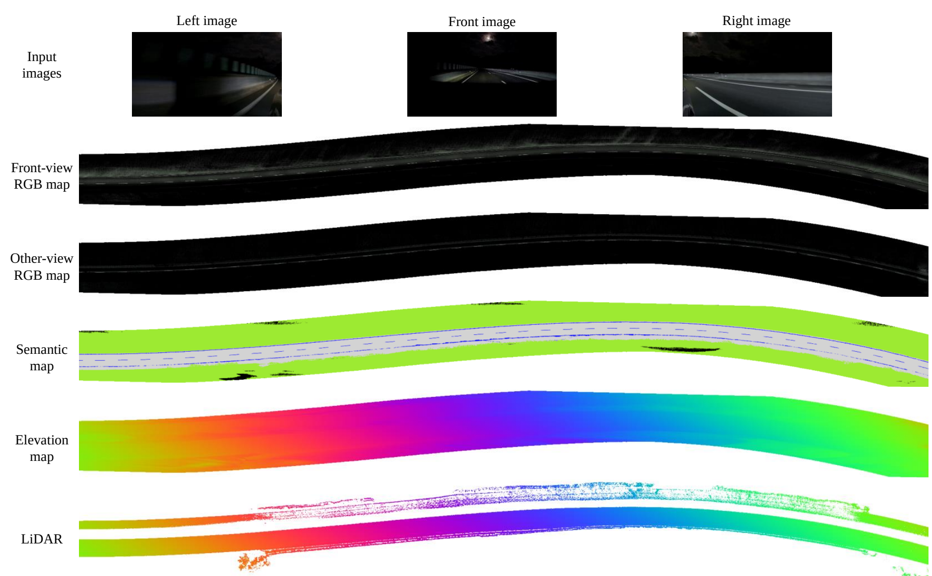

Fig. 9: Road surface reconstruction results of EMIE-MAP in a night scene.