

Temporal Hierarchical Gaussian Mixture Models for Real-Time Point Cloud Streaming

Roland Fischer, Tobias Gels, Rene Weller, Gabriel Zachmann

University of Bremen, Germany

✉ r.fischer@uni-bremen.de
zach@cs.uni-bremen.de

Problem

Robotics, autonomous driving, and telepresence applications rely on point clouds, i.e., for SLAM and scene/avatar reconstruction [4].

Efficient processing and accurate representation are crucial, especially for real-time streaming and rendering.

However, noisy sensor data, huge data loads, and inhomogeneous densities pose a challenge.

Related Work

Voxelization and occupancy-based methods suffer from discretization artifacts and high memory consumption.

Generative probabilistic models such as Gaussian Mixture Models (GMMs) allow for compact representations and high reconstruction fidelity and have been used for registration, compression, and incremental mapping [2].

Hierarchical forms and parallelization can reduce the high computational cost and allow for levels of detail (LODs) [1].

Our Approach

Compact point cloud representation and real-time streaming using a hierarchy of GMMs.

Level-by-level construction for dynamic, progressive transmission and rendering of LODs.

Minimize construction cost by exploiting temporal coherency between consecutive inputs.

Highly parallelized, optimized CUDA implement.

References

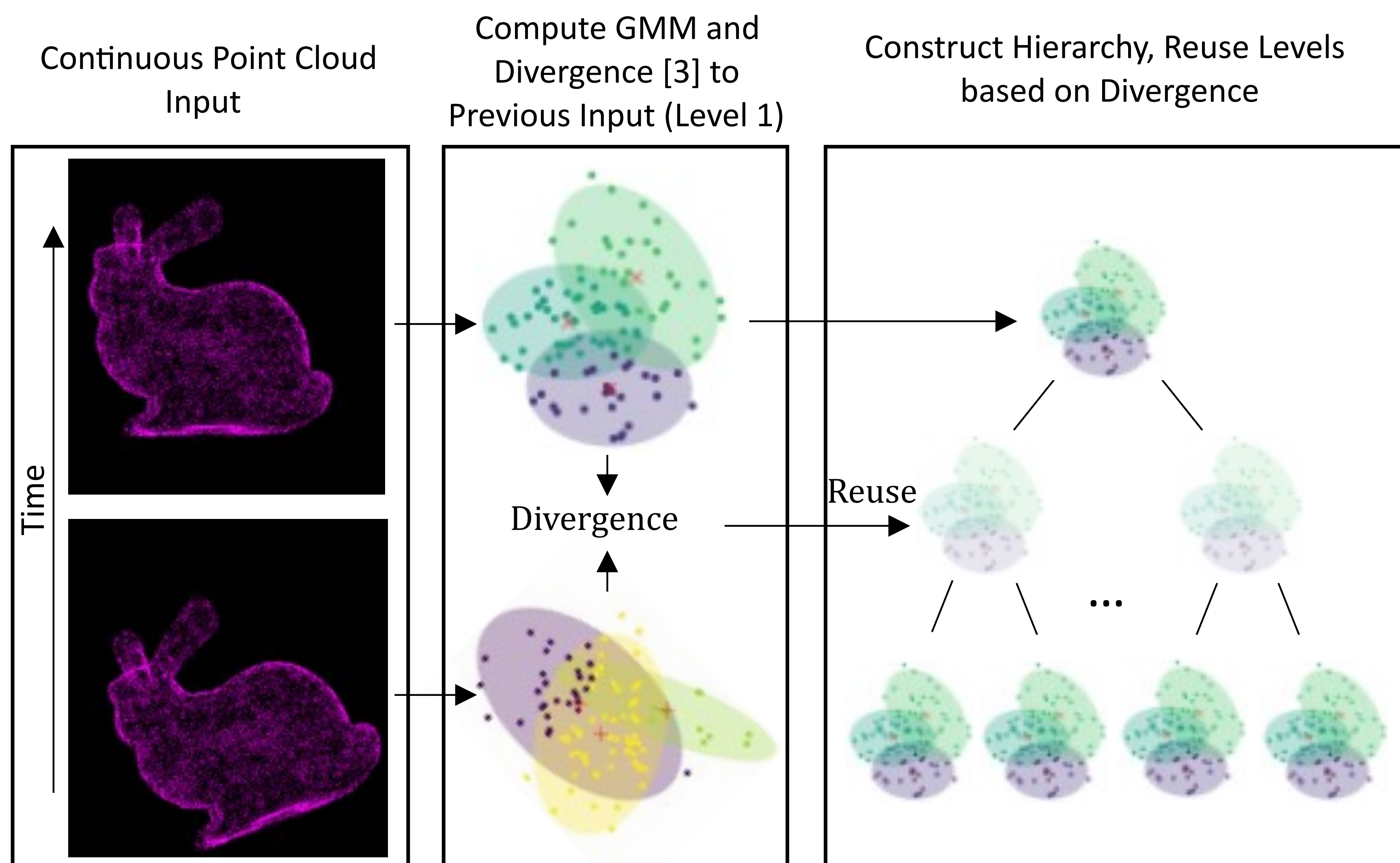
[1] Ben Eckart, Kihwan Kim, Alejandro Troccoli, Alonzo Kelly, and Jan Kautz. 2016. Accelerated Generative Models for 3D Point Cloud Data. In CVPR. 5497–5505.

[2] Kshitij Goel and Wennie Tabib. 2023. Incremental Multimodal Surface Mapping via Self-Organizing Gaussian Mixture Models. IEEE Robotics and Automation Letters (12 2023), 1–8.

[3] John R. Hershey and Peder A. Olsen. 2007. Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models. In ICASSP, Vol. 4. IV-317–IV-320.

[4] Kevin Yu, Gleb Gorbachev, Ulrich Eck, Frieder Pankratz, Nassir Navab, and Daniel Roth. 2021. Avatars for Teleconsultation: Effects of Avatar Embodiment Techniques on User Perception in 3D Asymmetric Telepresence. IEEE Transactions on Visualization and Computer Graphics PP (08 2021), 1–1.

Our Pipeline



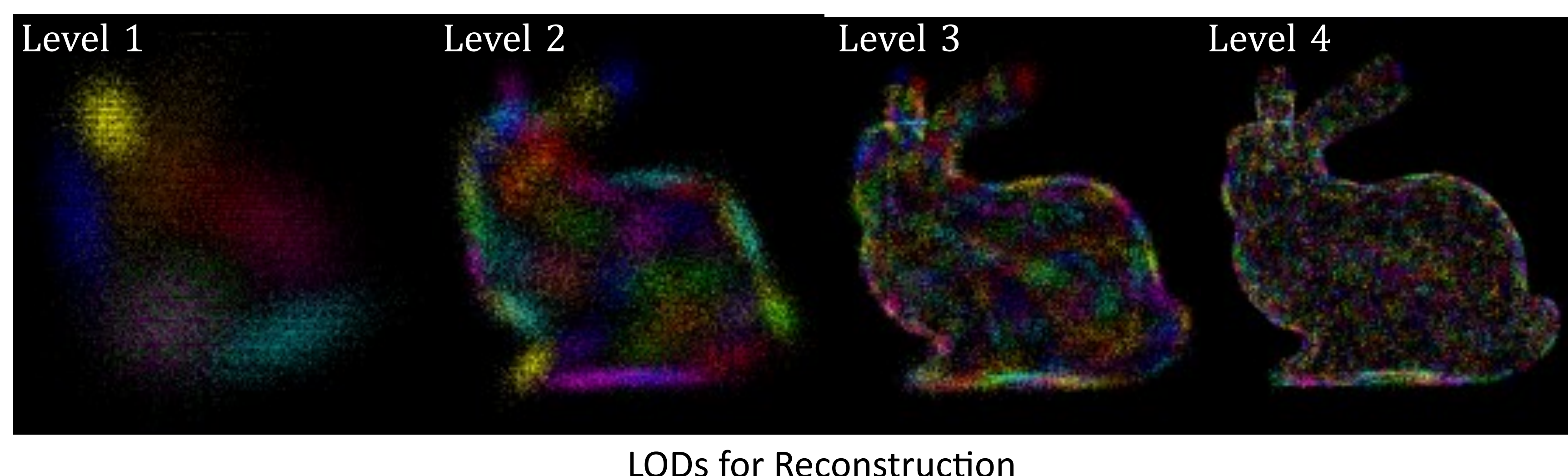
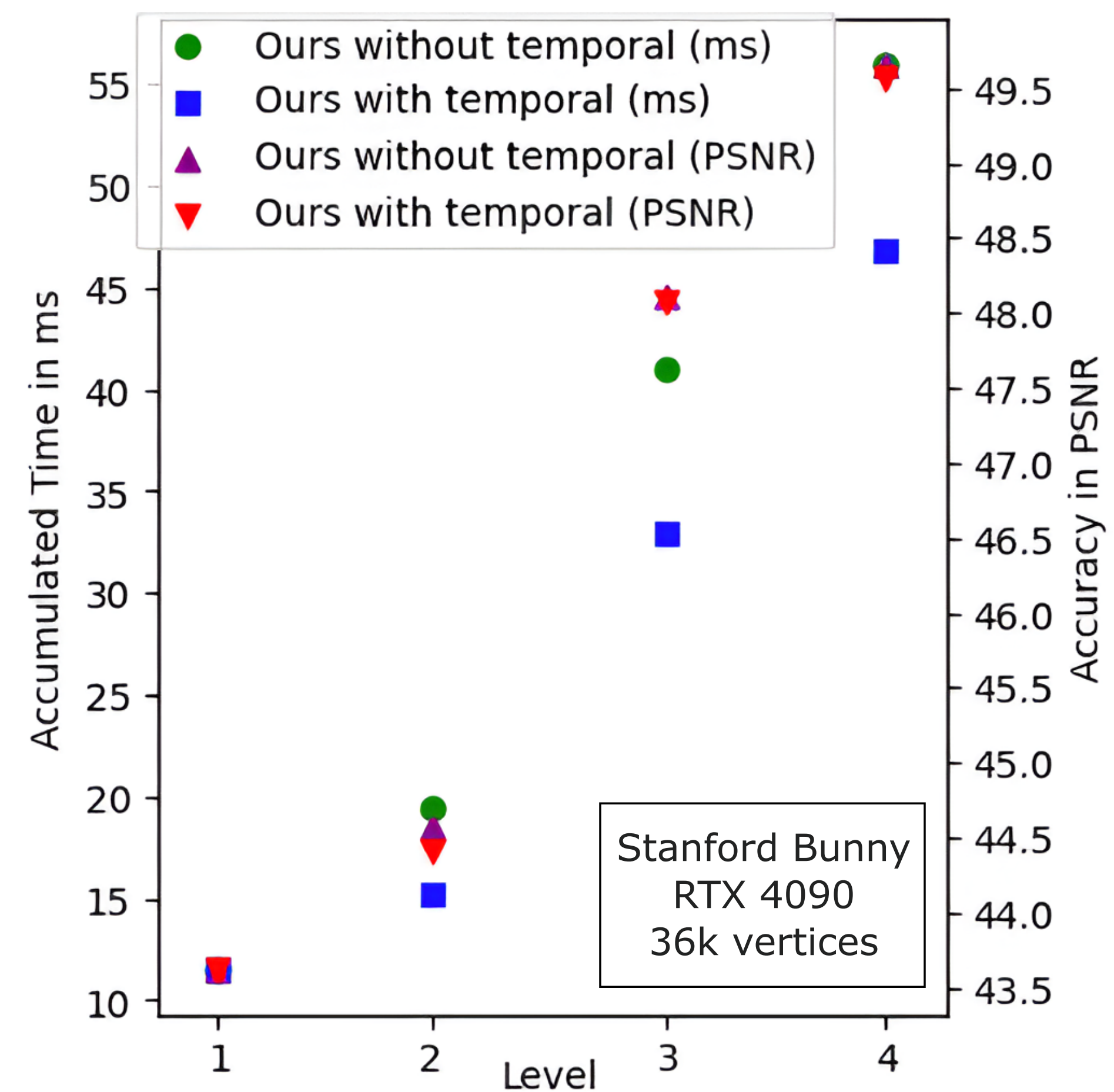
Results

Real-time speeds with high-fidelity reconstructions.

Significantly higher compression factor than previous work with similar accuracy [1]: 4x at level 4, 26x at level 3.

Temporal approach saves 20-36% construction time, without a noticeable loss in accuracy.

Adjustable parameters for balancing speed and accuracy using Kullback-Leibler divergence.



LODs for Reconstruction