
QuadricFormer: Scene as Superquadrics for 3D Semantic Occupancy Prediction Supplementary Material

Anonymous Author(s)

Affiliation

Address

email

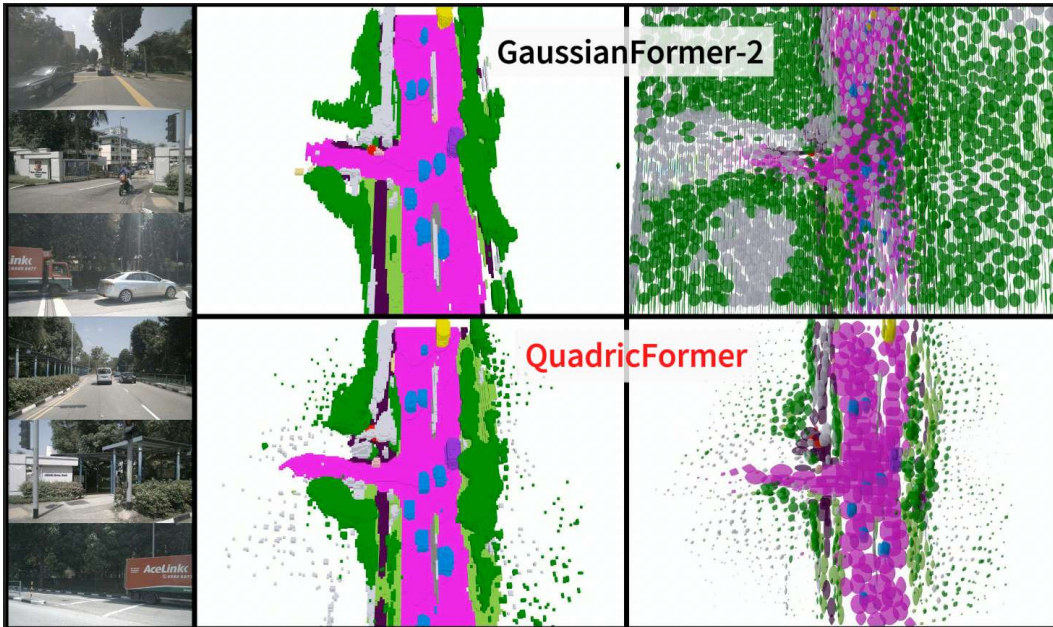


Figure 1: **Visualizations of the proposed QuadricFormer compared to GaussianFormer-2 [2] for 3D semantic occupancy prediction on the nuScenes [1] validation set.** We visualize the six surrounding camera inputs, the corresponding occupancy prediction results, and the primitive representations. The upper row shows the predicted occupancy(left) and the primitive representation(right) by GaussianFormer-2. The lower row shows the prediction results of QuadricFormer.

A Additional Superquadric Details

Superquadrics are a powerful family of parameterized surfaces that can represent various geometric shapes. With just a few parameters, superquadrics can generate shapes ranging from basic ellipsoids, cuboids, and cylinders to more complex shapes with rounded corners, star-like profiles, and smooth transitions between them. This geometric flexibility makes superquadrics ideal for efficiently modeling diverse objects in autonomous driving scenes. The shape of a superquadric is mainly controlled by two groups of parameters. The first group consists of scaling factors (s_x, s_y, s_z), which define the superquadric's dimensions or "radii" along its three principal axes, determining the object's overall size and aspect ratio. The second group includes two key shape parameters (ϵ_1, ϵ_2) that determine the degree of "squareness" or "roundness" of the object. ϵ_1 primarily controls the object's profile in planes containing the z-axis (such as the xz- or yz-plane): smaller values (close to 0.1) create sharper profiles, $\epsilon_1 = 1.0$ produces elliptical outlines, and larger values (up to 2.0) result in flatter contours. Similarly, ϵ_2 controls the shape of the cross-section in the xy-plane. A small ϵ_2 yields a star-shaped

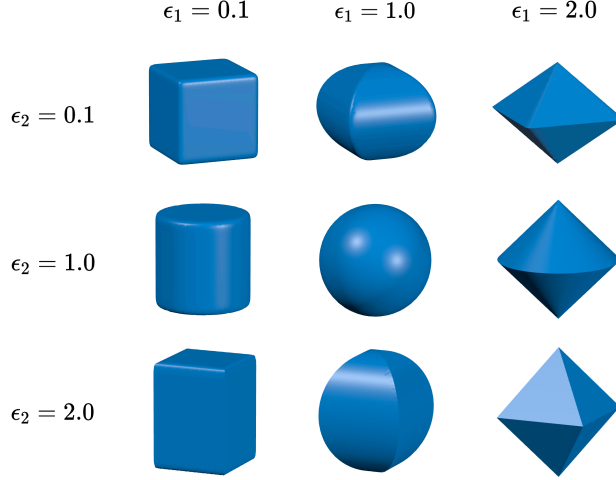


Figure 2: **Superquadrics of different shape parameters.** The figure illustrates how varying ϵ_1 and ϵ_2 produces a wide range of shapes, from star-like and rounded shapes to square-like structures. Such diversity enables superquadrics to flexibly model complex object geometries in 3D scenes.

cross-section, $\epsilon_2=1.0$ gives a circular outline, and large ϵ_2 values produce square-like shapes. As shown in Fig 2, varying ϵ_1 and ϵ_2 of superquadrics results in a wide range of shapes. By combining these scaling and shape parameters, superquadrics can efficiently represent diverse object geometries in autonomous driving scenes. This capability allows them to capture complex structures with significantly fewer primitives than traditional representations (like ellipsoidal Gaussians), highlighting their superior modeling efficiency and expressive power for 3D scene understanding tasks.

B Additional Experiments

We visualize the position distributions of scene primitives using 1600 superquadrics versus 6400 Gaussians in Figure 3. Gaussian-based methods require a dense arrangement of Gaussians throughout the entire 3D space to model the scene, leading to numerous redundant Gaussians and low modeling efficiency. In contrast, our superquadric-based method learns well-structured spatial arrangements, enabling it to effectively model the scene structure with significantly fewer primitives.

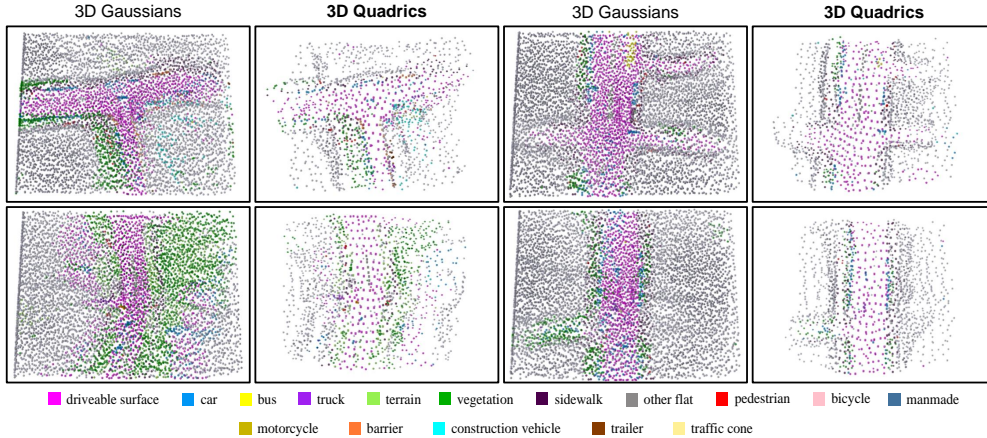


Figure 3: **Visualizations of primitive position distributions learned by different methods.** Our approach produces well-structured spatial arrangements while using significantly fewer primitives.

C Video Demonstration

Figure 1 shows a sampled image from the video demo for 3D semantic occupancy prediction on the nuScenes [1] validation set. Compared to GaussianFormer-2 [2], our QuadricFormer exhibits enhanced modeling capability for complex objects and road surfaces. This demonstrates the effectiveness of our Quadric-based model.

33 **References**

- 34 [1] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush
35 Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for
36 autonomous driving. In *CVPR*, 2020.
- 37 [2] Yuanhui Huang, Amonnut Thammatadatrakoon, Wenzhao Zheng, Yunpeng Zhang, Dalong Du,
38 and Jiwen Lu. Probabilistic gaussian superposition for efficient 3d occupancy prediction. *arXiv*
39 *preprint arXiv:2412.04384*, 2024.