PAPR in Motion: Seamless Point-level 3D Scene Interpolation

Shichong Peng, Yanshu Zhang, Ke Li
APEX Lab, School of Computing Science, Simon Fraser University

CVPR SEATTLE, WA JUNE 17-21, 2024

Abstract

PAPR in Motion is the first method to address the novel problem of point-level 3D scene interpolation. Building upon the recent Proximity Attention Point Rendering (PAPR) [1] technique, it generates seamless interpolations between both the scene geometry and appearance.

Contributions:

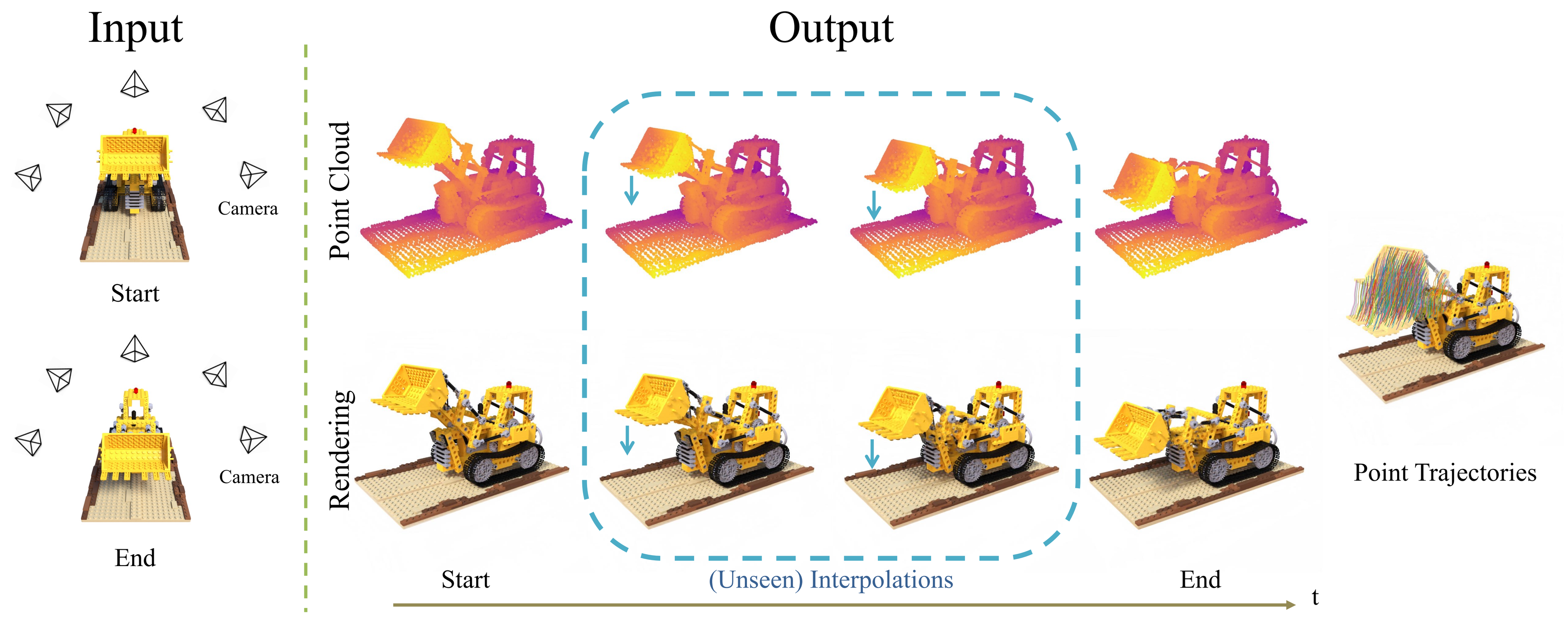
- We introduce the new task of point-level 3D scene interpolation between two distinct states.
- We propose two simple but effective regularization techniques to maintain the temporal consistency of the geometric structures.
- We significantly outperforms the latest dynamic scene point-based method, Dynamic Gaussian [2], on this task.

Problem Setting

Point-level 3D Scene Interpolation

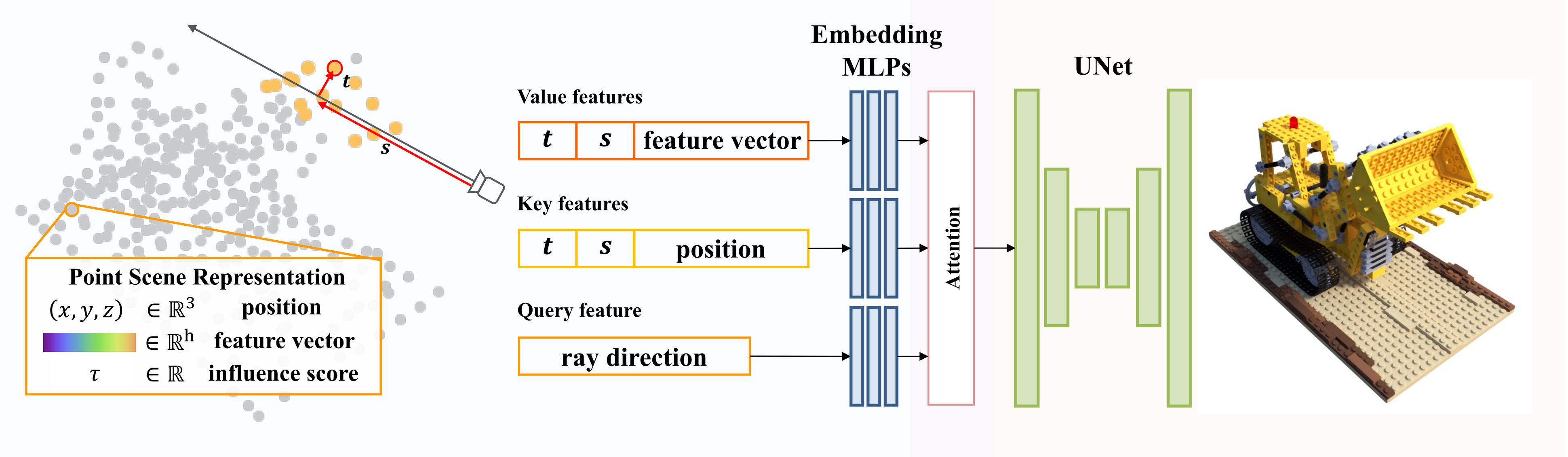
Given observations of a scene at *two* distinct states from multiple views, the goal is to synthesize a smooth point-level interpolation between these states.

- **No intermediate supervision**: only the start and end states are observed, different from typical dynamic scene reconstruction with dense per-frame observations.
- **Point-level interpolation**: as opposed to part-level interpolation, each point in the scene can move freely. i.e., there could be non-rigid transformation.
- **Challenging motion types**: can contain large displacement with multiple moving parts.



Background

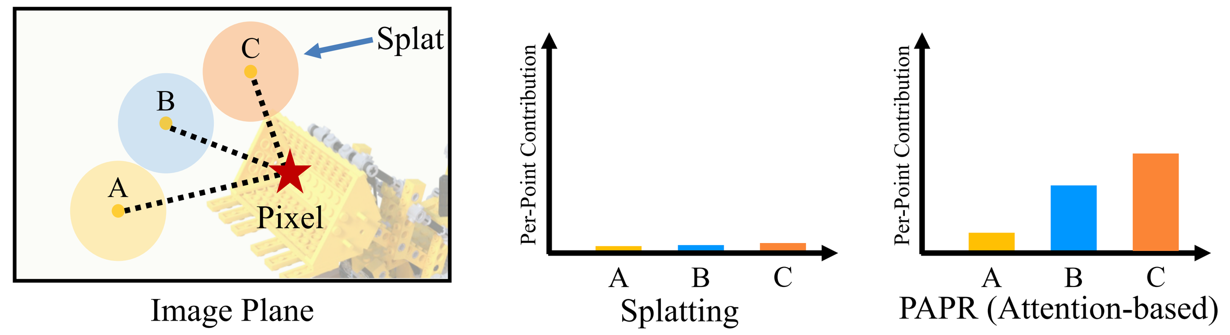
Proximity Attention Point Rendering (PAPR) is a recent method for joint novel view synthesis and 3D reconstruction. It simultaneously learns from scratch an accurate point cloud representation of the scene surface, and an attention-based neural network that renders the point cloud from novel views.



Limitations of Gaussian Splatting

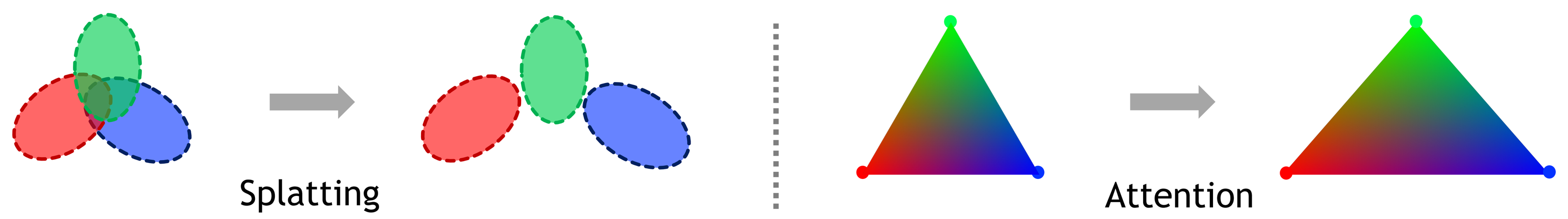
Vanishing Gradient

Consider a pixel with high loss that's not covered by any splat. Splatting-based methods would fail to move splats to cover it because of vanishing gradients. PAPR avoids this issue by using an attention mechanism where the total attention weights always sum up to one, making it more suitable for learning large scene changes.



Rendering Breaks Under Non-rigid Transformation

Gaussian splats may no longer be Gaussian after non-rigid transformations and would introduce gaps in rendering. In contrast, attention-based methods like PAPR render by interpolating nearby points and avoid gaps naturally.



Compared to Gaussian Splatting [3], PAPR can better preserve rendering quality after non-rigid transformation, allowing for rendering non-rigid scene changes during the interpolation process.

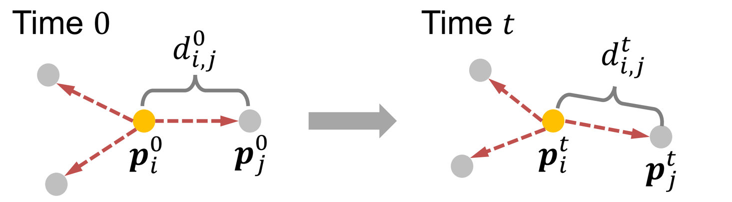
Method

Step 1: Train a PAPR model on the start state from scratch, which serves as a scene template. Each point in this template carries local scene appearance information.

Step 2: Fix all parameters in the scene template except for the point positions, and finetune the template using the multi-view RGB images at the end state. The intermediate point positions in the optimization process serve as the interpolated geometry.

To avoid implausible deformation and ensure motion coherence, we introduce two regularization techniques:

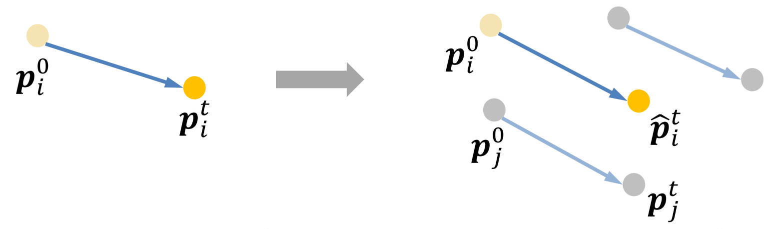
- **Local distance preservation loss ($\mathcal{L}_{rigid}$)**: encourage local rigidity.



$$\mathcal{L}_{rigid} = \frac{1}{kN} \sum_i \sum_{j \in NN_k(i)} |d^0_{i,j} - d^t_{i,j}|$$

total number of points (N); k-nearest neighbours ($NN_k(i)$)

- **Local displacement averaging step (LDAS)**: encourage local uniform movement.

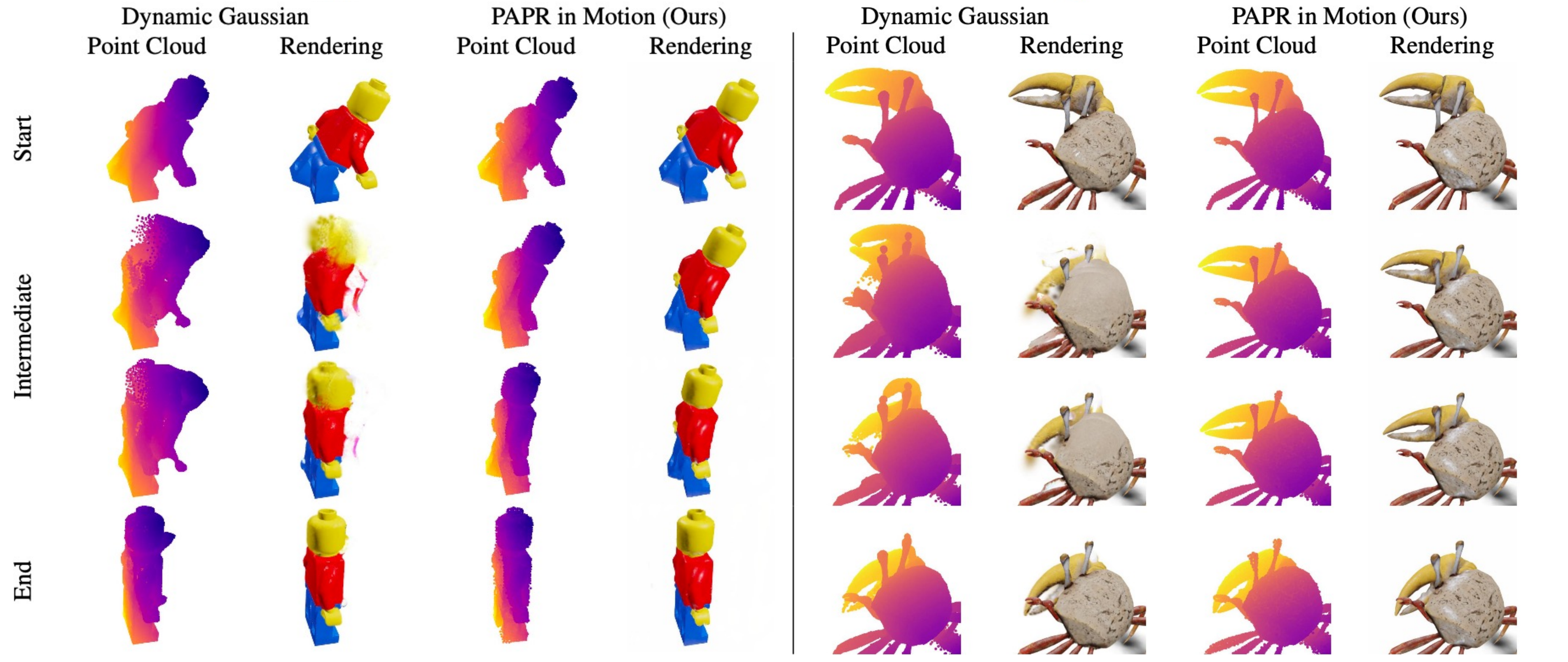


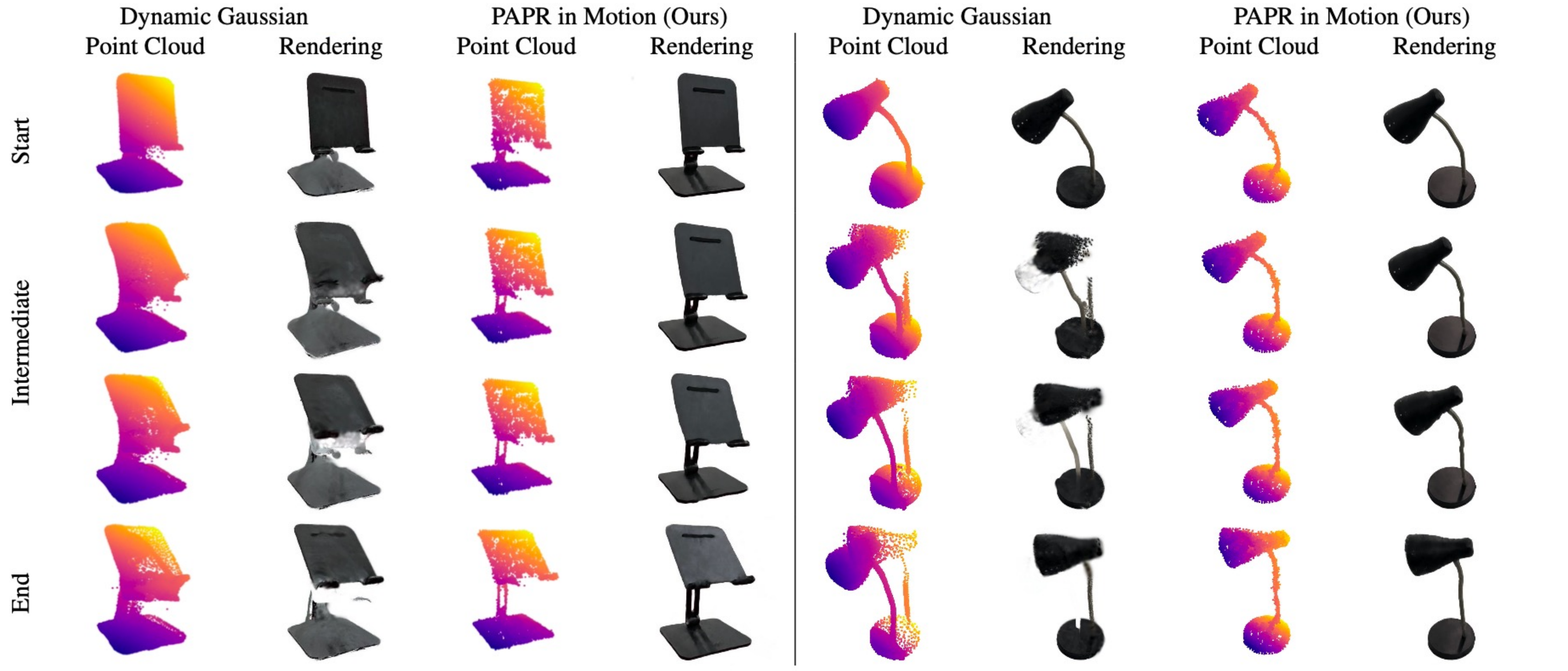
$$\hat{p}^t_i = p^0_i + \frac{1}{k} \sum_{j \in NN_k(i)} (p^t_j - p^0_j)$$

Step 3: Once the geometry converges, fix the point positions, and finetune the point feature vectors, embedding MLPs, and the attention model to capture scene appearance differences (e.g., change in shadows or specular reflections).

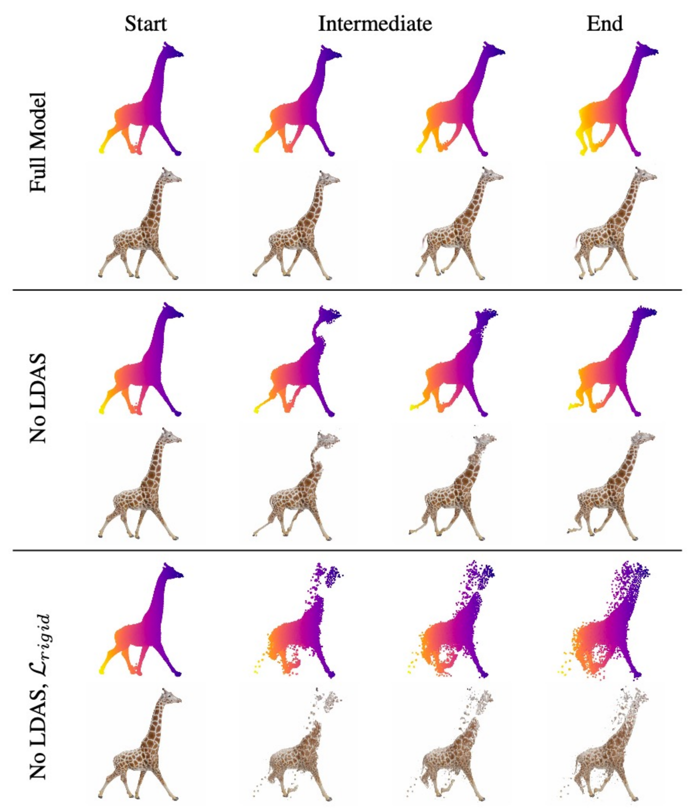
To render an intermediate time step, we interpolate the feature vectors and model weights.

Results

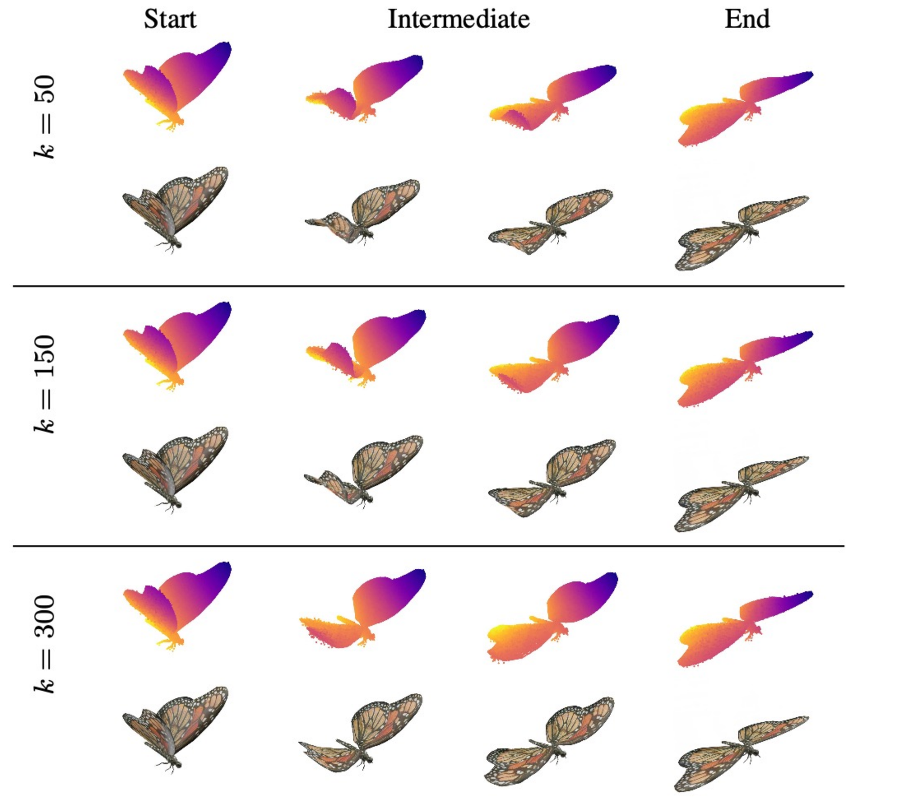




Ablation Study



Sensitivity Analysis



References

[1] Zhang, Y., Peng, S., Moazeni, A., & Li, K. (2023). PAPR: Proximity Attention Point Rendering. *NeurIPS*.
[2] Luiten, J., Kopanas, G., Leibe, B., & Ramanan, D. (2023). Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis. *3DV*.
[3] Kerbl, B., Kopanas, G., Leimkuehler, T., & Drettakis, G. (2023). 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics (TOG)*.