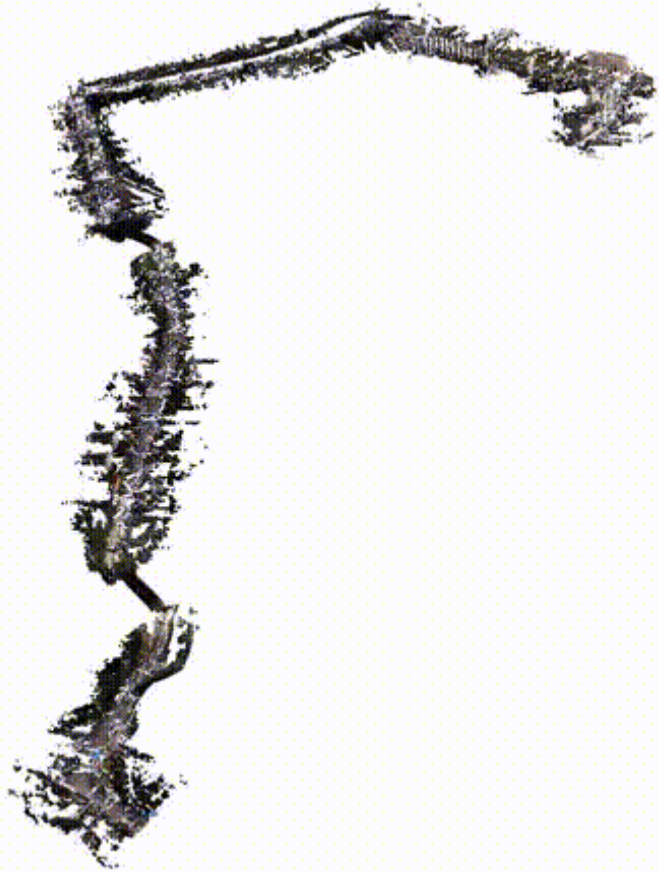
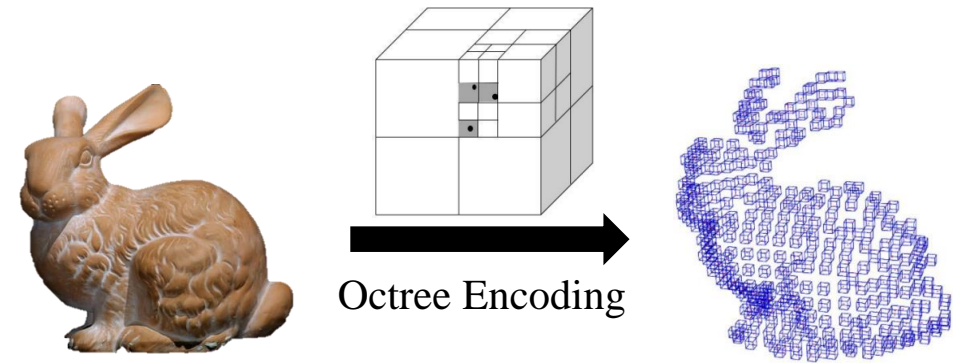




## Background



- Excessively Large Data Volume  
One Frame 2.4GB!
- High Coding Complexity

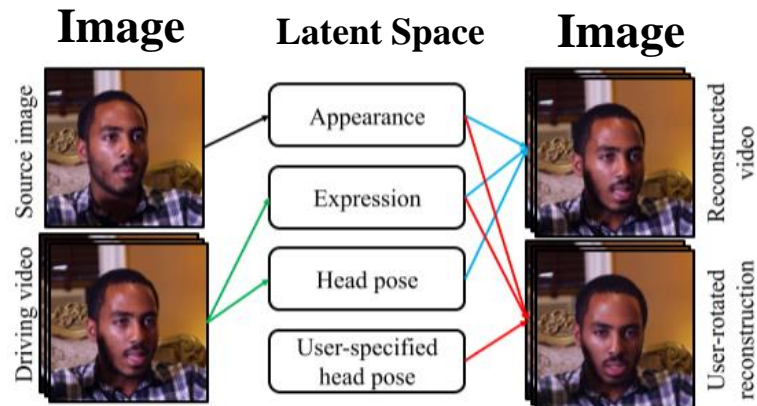




## Background

## Inspired

### One-Shot Free-View Neural Talking-Head Synthesis for Video Conferencing [1]



Can we transfer this method directly to **point cloud** transmission?



## Background

### What is the difference between a 2D image and a 3D point cloud?

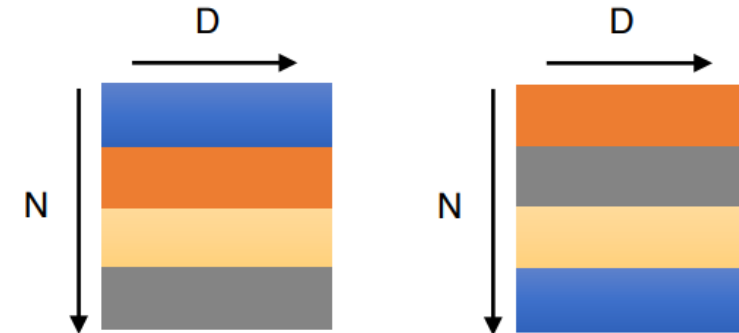
A point cloud is **a set of** data points in space.



- Irregularity [1]

The point cloud data is non-Eulerian structured.

- Permutation invariance

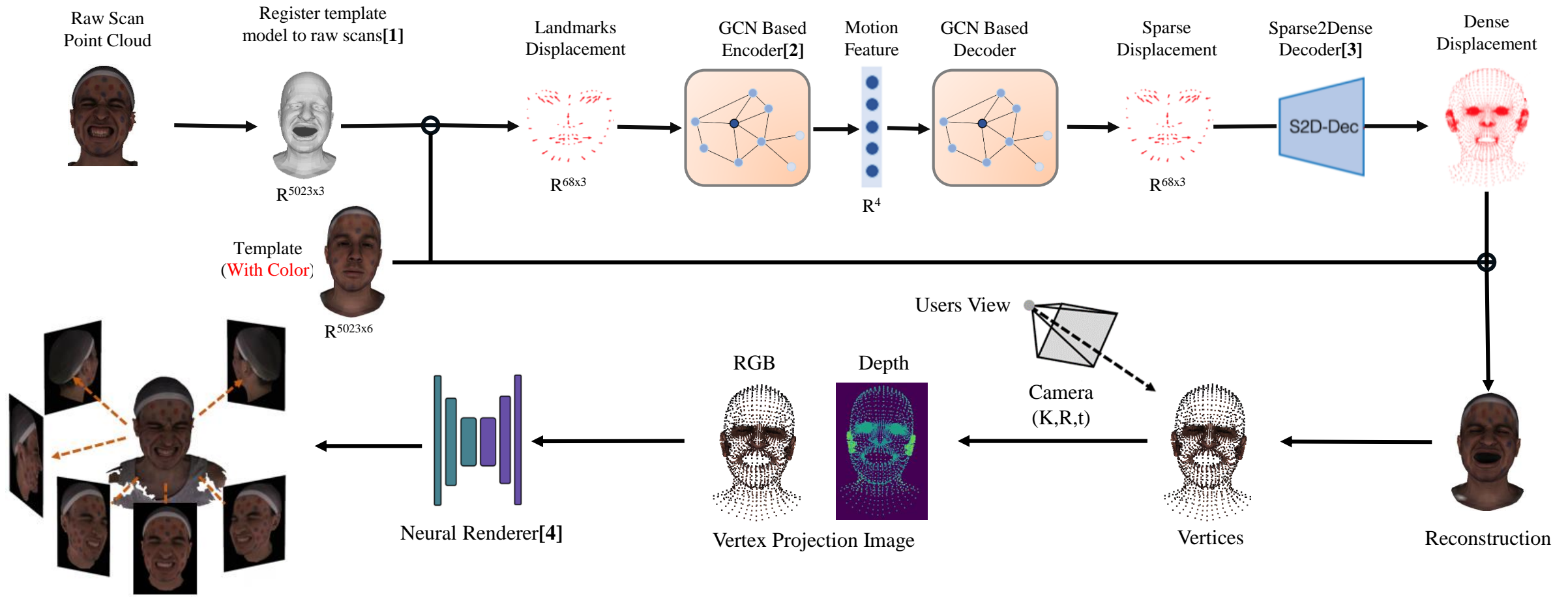


- Transformation invariance

Finding a **latent space** to directly **manipulate point clouds** is not an easy task.



## Method



[1] Li T, Bolkart T, Black M J, et al. Learning a model of facial shape and expression from 4D scans[J]. ACM Trans. Graph., 2017, 36(6): 194:1-194:17.

[2] "Generating 3D faces using convolutional mesh autoencoders." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.

[3] Otterdout N, Ferrari C, Daoudi M, et al. Sparse to Dense Dynamic 3D Facial Expression Generation[C]//Conference on Computer Vision and Pattern Recognition (CVPR). 2022.

[4] Prokudin S, Black M J, Romero J. SmpIpix: Neural avatars from 3d human models[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021: 1810-1819.





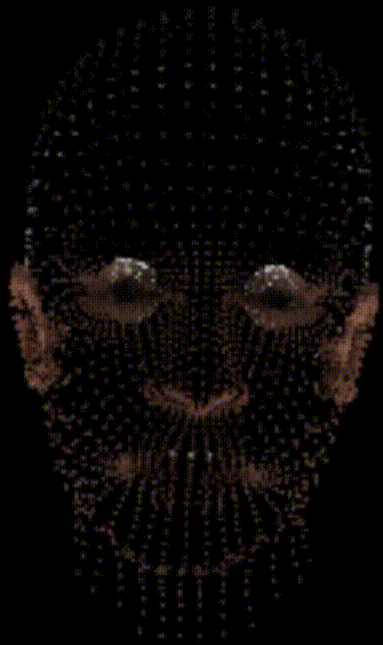
西安电子科技大学  
XIDIAN UNIVERSITY

# Research Experience

## Keypoint-Driven Neural Head Avatars

## Results

---



**Keypoint-Driven**



**Neural Rendering**



**Raw Scan**