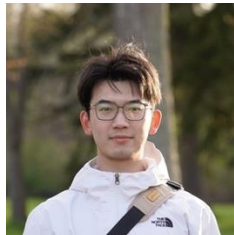


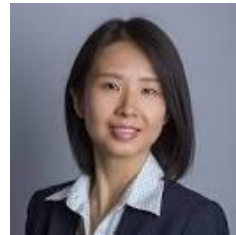
High Fidelity 3D Hand Shape Reconstruction via Scalable Graph Frequency Decomposition



Tianyu Luan,



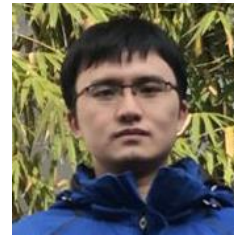
Yuanhao Zhai,



Jingjing Meng,



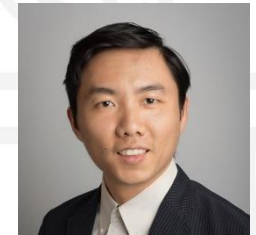
Zhong Li,



Zhang Chen,



Yi Xu,



Junsong Yuan

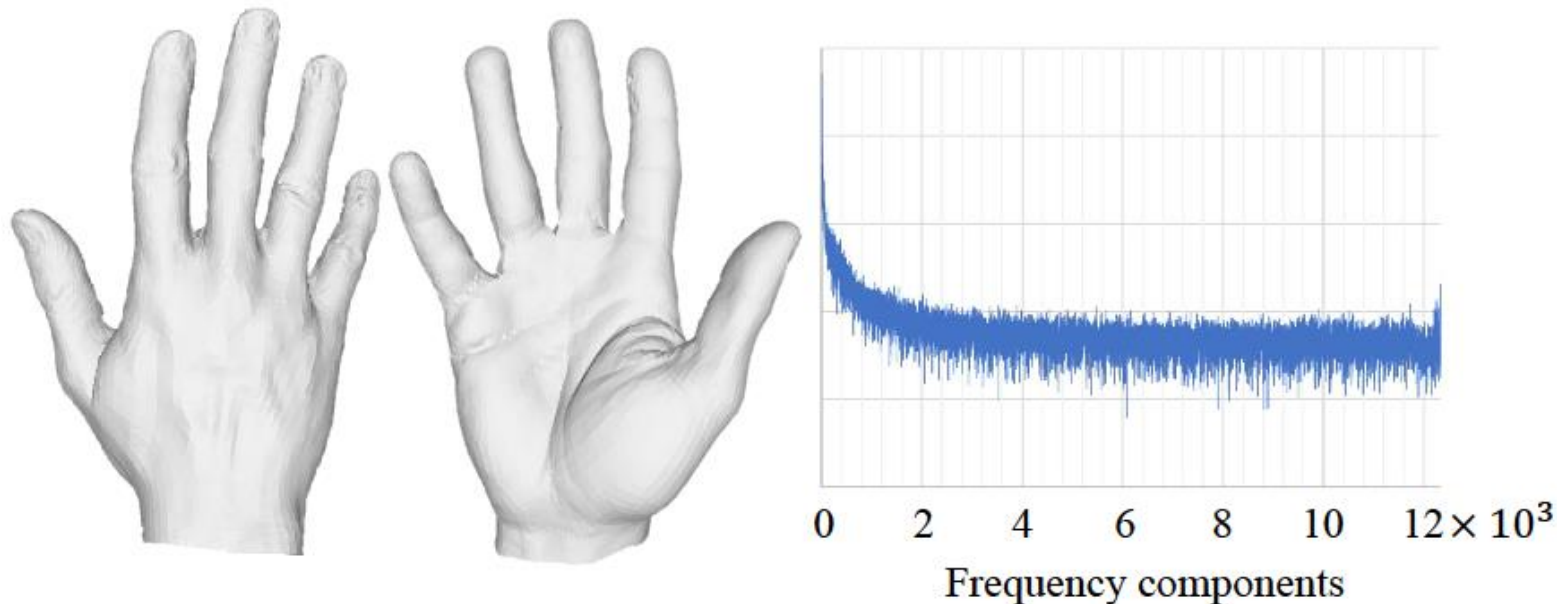


State University of New York at Buffalo

OPPO US Research Center, InnoPeak Technology, Inc.

CONTRIBUTION

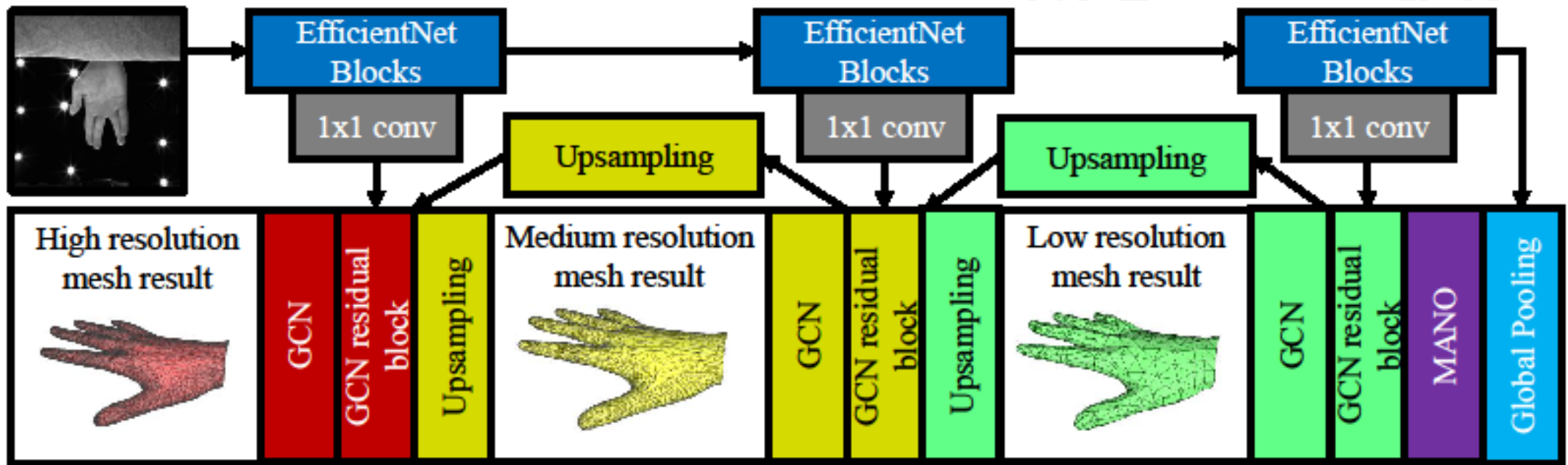
- We design a **high-fidelity 3D hand reconstruction** model based on **frequency decomposition**
- We propose a **frequency split network** architecture to generate high-fidelity hand mesh in a **scalable manner**.
- We propose a **new metric to evaluate 3D mesh details**.



An exemplar hand mesh of sufficient details and its graph frequency decomposition.

PIPELINE

- **Multi-resolution** hand mesh and graph convolution network
- **Multi-scale** coarse-to-fine high-fidelity detail recovery.
- **Multi-stage feature injection.**



FREQUENCY DECOMPOSITION

- Graph Fourier Transform:

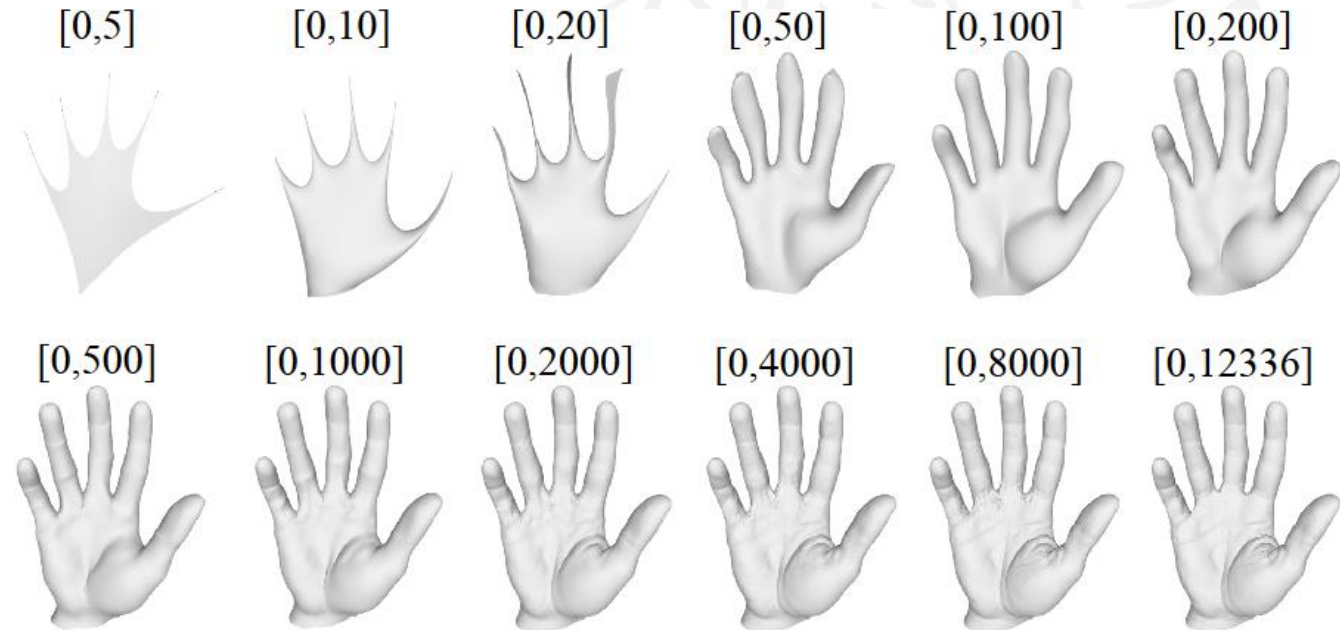
$$\mathbf{L} = \mathbf{U}^\top \mathbf{\Lambda} \mathbf{U}, \quad x = \sum_{i=1}^N \mathbf{U}_i (\mathbf{U}_i^\top x)$$

- Frequency decomposition loss:

$$L_F = \frac{1}{F} \sum_{f=1}^F \log \left(\frac{\|\mathbf{U}_f^\top \hat{V} - \mathbf{U}_f^\top V_{gt}\|^2}{\|\mathbf{U}_f^\top \hat{V}\| \|\mathbf{U}_f^\top V_{gt}\| + \epsilon} + 1 \right)$$

- Mean signal-to-noise ratio:

$$\text{MSNR} = \frac{1}{F} \sum_{f=1}^F \log \left(\frac{\|\mathbf{U}_f^\top \hat{V}\|}{\|\mathbf{U}_f^\top \hat{V} - \mathbf{U}_f^\top V_{gt}\| + \epsilon} \right)$$



Frequency decomposition of 3D hand mesh

EXPERIMENTS

- Performance:

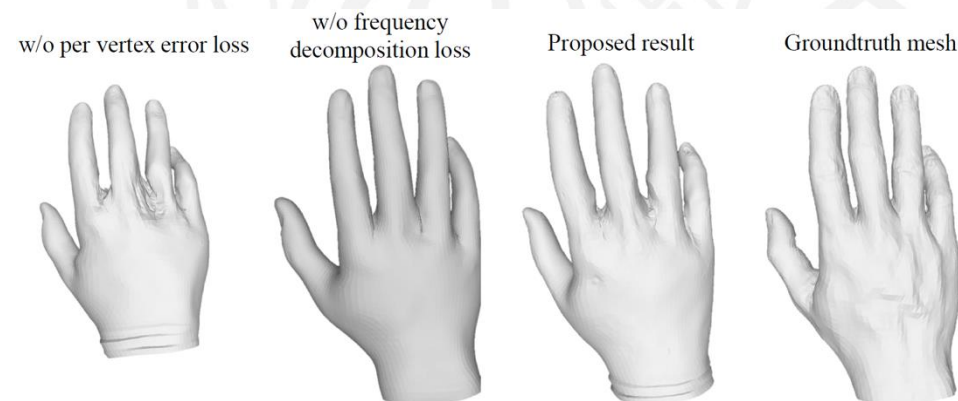
Method	MPJPE/mm ↓	CD/mm ↓	MSNR ↑
MANO	13.41	6.20	-2.64
Ours-level 1	13.25	5.53	-2.70
Ours-level 2	13.25	5.49	-2.62
Ours-level 3	13.25	5.49	-0.68

- Ablation:

Method	MPJPE/mm ↓	CD/mm ↓	MSNR ↑
proposed	13.25	5.49	-0.68
w/o skip connected feature	14.20	5.85	-0.70
w/ average pooling feature	13.95	5.59	-1.10
w/o frequency decomposition loss	14.50	5.86	-1.80
w/o per vertex error loss	14.24	67.8	-0.87

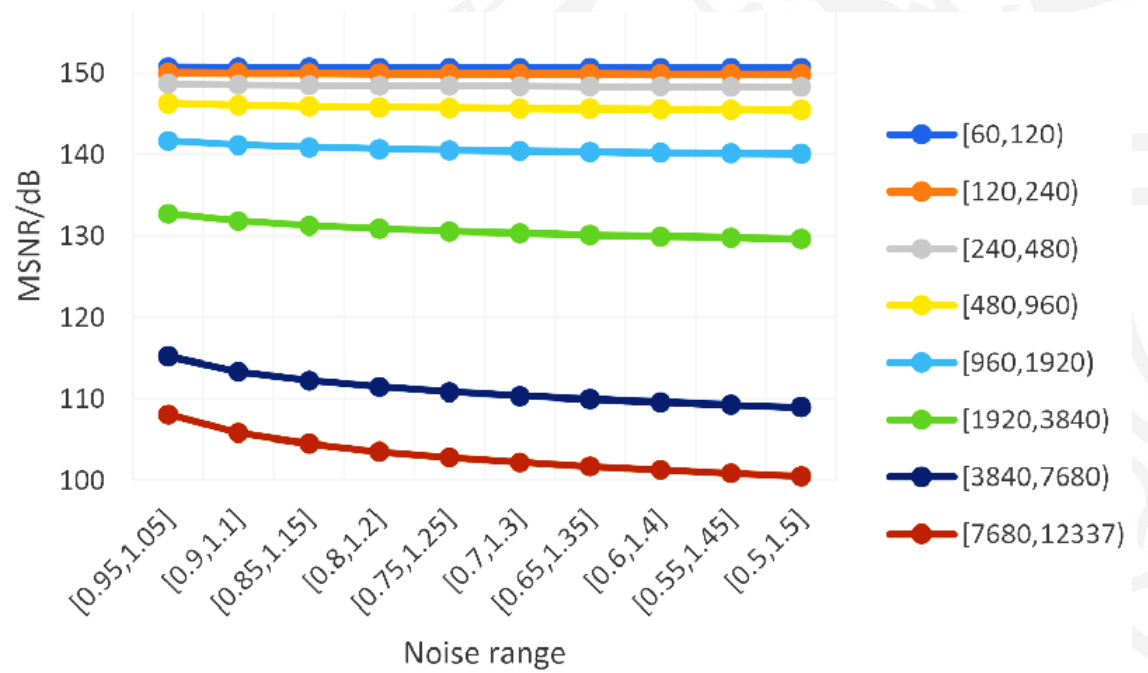
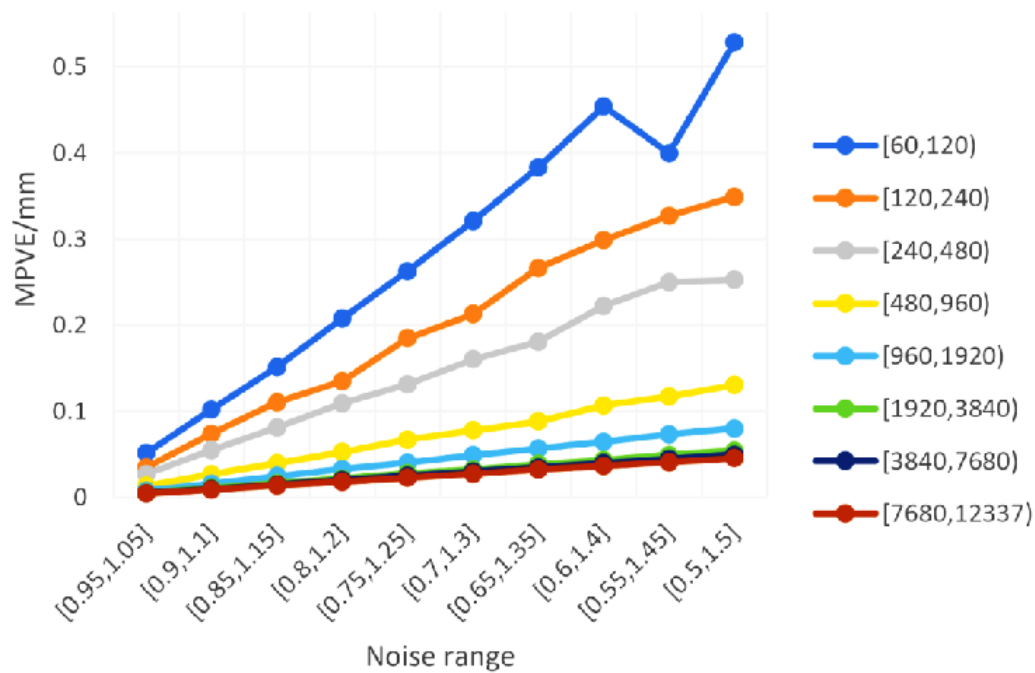
- Scalable efficiency:

Level	#parameter	GFLOPS	#vertices	#faces
baseline	14.5M	1.8	778	1538
1	14.5M	1.9	778	1538
2	14.5M	2.5	3093	6152
3	14.7M	4.8	12337	24608



EXPERIMENTS (cont'd)

- Metric sensitive along the frequency:



VISUALIZATION

