Supplementary Material for Coordinate Quantized Neural Implicit Representations for Multi-view Reconstruction

Sijia Jiang, Jing Hua, Zhizhong Han
Department of Computer Science, Wayne State University, Detroit, USA
sijiajiang@wayne.edu, jinghua@wayne.edu, h312h@wayne.edu

1. Network and Training

We use the same network architectures and experimental settings as our baselines for fair comparisons. We report our results with baselines by discretizing continuous coordinates into discrete coordinates.

2. Statistical Analysis

We report more comprehensive analysis with statistics in the section. We show the number of unique points that the network has observed and the number of multi-view consistency that can be triggered with overlapped samples on rays from different views.

2.1. UNISURF

We show the statistical analysis with the baseline UNISURF [3] in the each scene under DTU [2]. The comparisons from Fig. 1 to Fig. 15 highlight our merits over continuous coordinates in terms of reducing variations in the input space and triggering more multi-view consistency constraints. Please watch our video to see the rays with unique samples and rays constrained by multi-view consistency in different iterations.

2.2. NeuS

We repeat the statistical analysis with the baseline NeuS [4] in each scene under DTU [2]. The comparisons from Fig. 16 to Fig. 30 also show that we can reduce variations in the input space and trigger more multi-view consistency constraints with NeuS. We report our statistical analysis with NeuS in each scene in DTU in Tab. 3. The numerical comparison indicates the same conclusion we get with UNISURF, which highlights our dicreteized coordinates in reducing variations in the input sample space and triggering more multi-view consistency constraints. Please watch our video to see the rays with unique samples and rays constrained by multi-view consistency in different iterations.

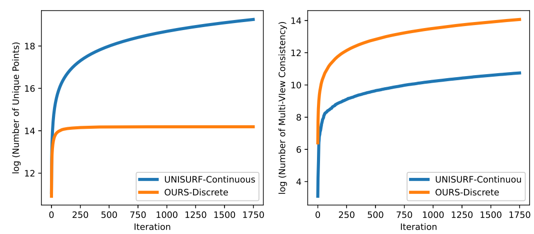

Figure 1. Visual comparisons with UNISURF in scene 24 under DTU.

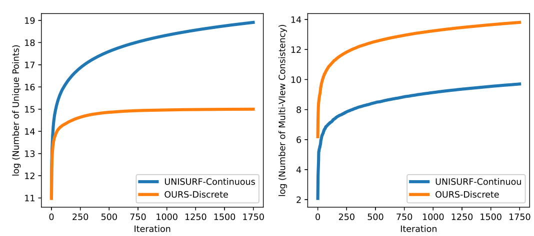

Figure 2. Visual comparisons with UNISURF in scene 37 under DTU.

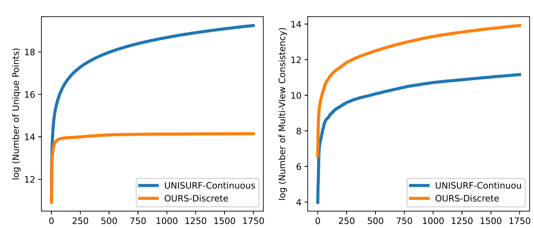

Figure 3. Visual comparisons with UNISURF in scene 40 under DTU.

3. More Optimization Stability Analysis

3.1. SAPE

We show our advaNtages over SAPE [1] based on HF-NeuS [5]. SAPE adds adaptive weights on each element in positional encoding to resist the noises brought by high frequencies. We report visual comparison in Fig. 31 and numerical comparison in Tab. 4. We report results with HF-NeuS without using SAPE adaptive weights, and re-

Scan	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean
NeuS	1.37	1.21	0.73	0.40	1.20	0.70	0.72	**1.01**	**1.16**	0.82	0.66	1.69	0.39	0.49	0.51	0.87
NeuS High Resolution	0.85	0.96	0.84	0.38	**1.0**	**0.6**	**0.58**	1.44	1.23	0.78	0.52	1.18	**0.32**	0.46	0.54	0.78
NeuS Low Resolution	**0.71**	**0.95**	**0.68**	**0.38**	**1.0**	**0.6**	**0.58**	1.40	1.17	**0.79**	**0.52**	**1.07**	**0.32**	**0.43**	**0.45**	**0.74**

Table 1. Neus compares with our method in different resolutions on DTU dataset

Scan	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean
NeuralWarp	**0.49**	0.71	0.38	0.38	0.79	0.81	0.82	1.20	**1.06**	**0.68**	0.66	0.74	0.41	0.63	0.51	0.68
NeuralWarp High Resolution	**0.49**	0.71	**0.37**	0.38	**0.77**	**0.79**	**0.77**	**1.17**	1.18	0.72	**0.62**	0.73	0.37	0.6	**0.49**	0.68
NeuralWarp Low Resolution	0.5	**0.68**	**0.37**	**0.36**	0.81	**0.79**	0.83	1.18	1.09	**0.68**	**0.62**	**0.65**	**0.36**	**0.57**	**0.49**	**0.67**

Table 2. NeuralWarp compares with our method in different resolutions on DTU dataset

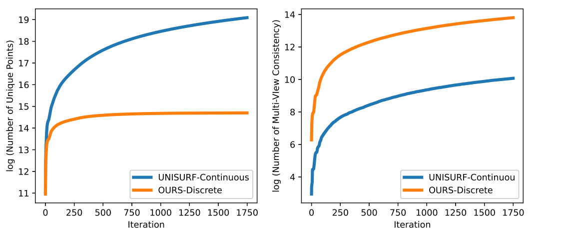

Figure 4. Visual comparisons with UNISURF in scene 55 under DTU.

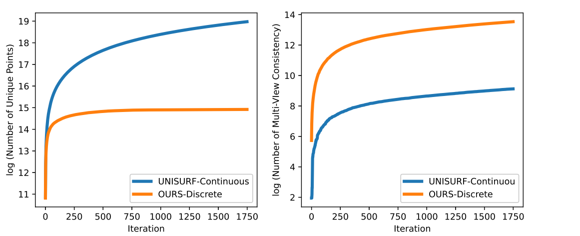

Figure 5. Visual comparisons with UNISURF in scene 63 under DTU.

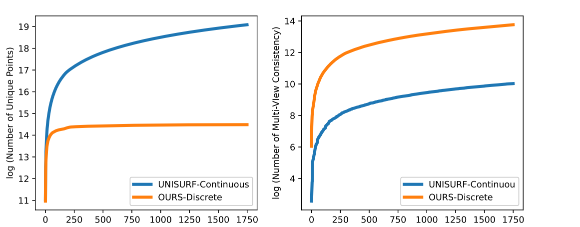

Figure 6. Visual comparisons with UNISURF in scene 65 under DTU.

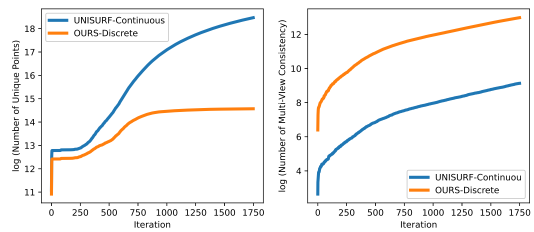

Figure 7. Visual comparisons with UNISURF in scene 69 under DTU.

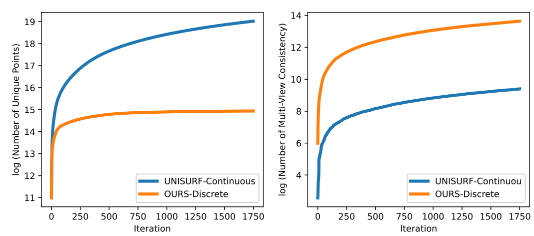

Figure 8. Visual comparisons with UNISURF in scene 83 under DTU.

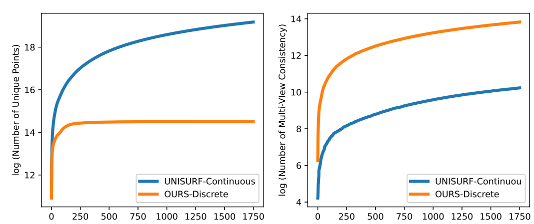

Figure 9. Visual comparisons with UNISURF in scene 97 under DTU.

sults with using our discrete coordinates. We can see that our results are much more accurately with high frequency positional encodings.

3.2. UNISURF

We increase the frequency in positional encodings with UNISURF by using 2 times higher frequency. Fig. 32 shows that our method can stabilize the optimization with high frequency positional encodings, which leads to more accurate reconstruction. We report neumerical comparison in Tab 5.

4. More Results

We show more visual comparisons with the state-of-the-art methods using error maps in our video. Please watch our video for more details.

We also report our results based on NeuS and NeuralWarp with high and low resolution quantized coordinates in Tab. 2 and Tab. 1.

Scan	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean
Unique Ratio	121.8	108.4	105.4	47.6	120.4	116.7	45.4	127.8	118.8	111.3	120.7	110.2	107.2	109.9	119.8	106.1
Consistency Ratio	0.056	0.080	0.047	0.213	0.055	0.075	0.081	0.065	0.238	0.086	0.497	0.262	0.230	0.149	0.035	0.145

Table 3. Statistical analysis for our improvements over baselines NeuS.

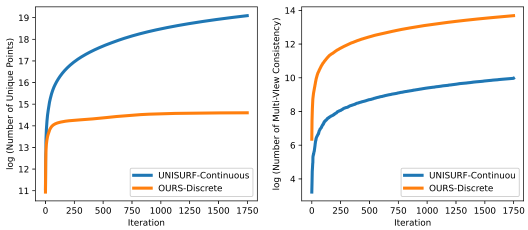

Figure 10. Visual comparisons with UNISURF in scene 105 under DTU.

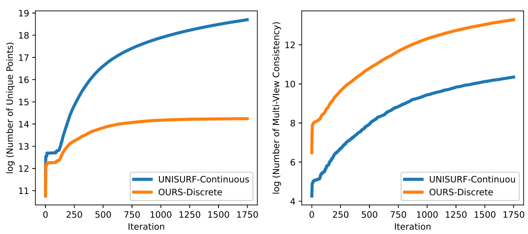

Figure 11. Visual comparisons with UNISURF in scene 106 under DTU.

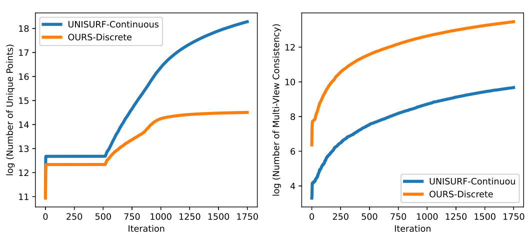

Figure 12. Visual comparisons with UNISURF in scene 110 under DTU.

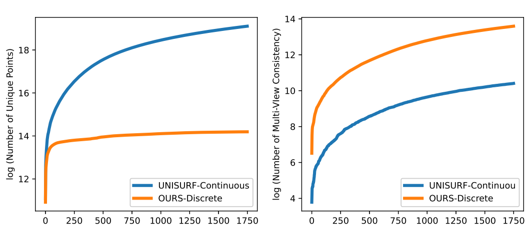

Figure 13. Visual comparisons with UNISURF in scene 114 under DTU.

Scan	37
HF-NeuS	1.32
HF-NeuS w/o Adaptive Weight	1.20
Ours w/o Adaptive Weight	1.08

Table 4. Numerical comparisons with HF-Neus w/o Adaptive Weight in scene 37 under DTU.

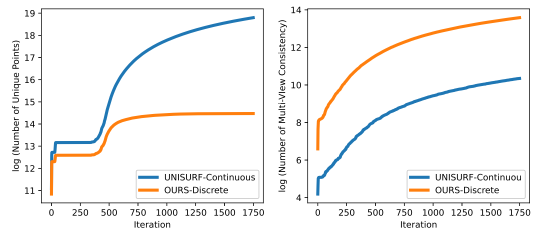

Figure 14. Visual comparisons with UNISURF in scene 118 under DTU.

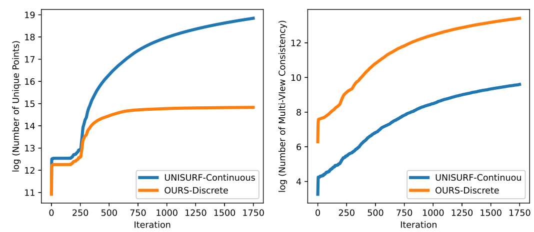

Figure 15. Visual comparisons with UNISURF in scene 122 under DTU.

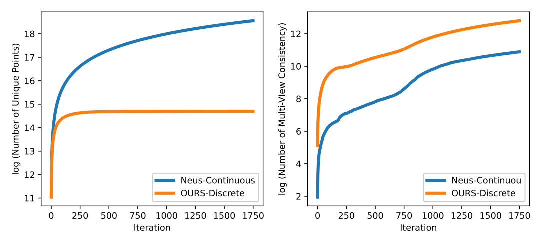

Figure 16. Visual comparisons with NeuS in scene 24 under DTU.

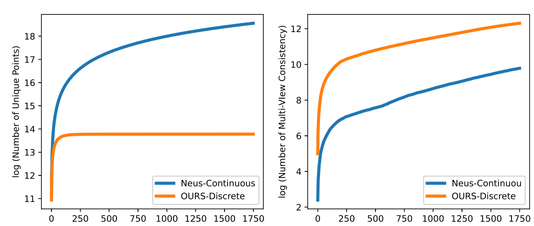

Figure 17. Visual comparisons with NeuS in scene 37 under DTU.

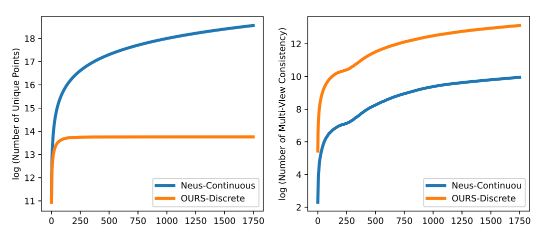

Figure 18. Visual comparisons with NeuS in scene 40 under DTU.

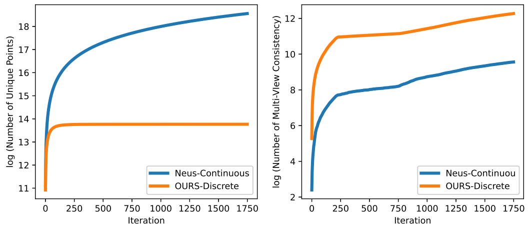

Figure 19. Visual comparisons with NeuS in scene 55 under DTU.

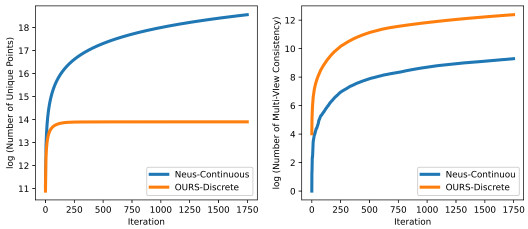

Figure 20. Visual comparisons with NeuS in scene 63 under DTU.

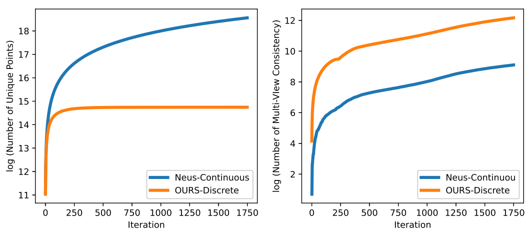

Figure 21. Visual comparisons with NeuS in scene 65 under DTU.

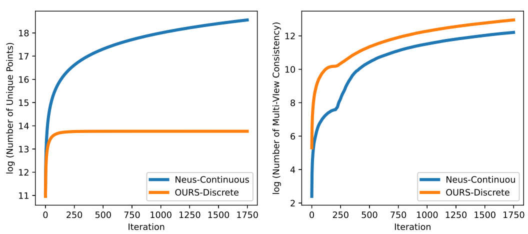

Figure 22. Visual comparisons with NeuS in scene 69 under DTU.

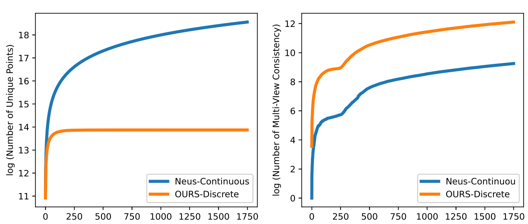

Figure 23. Visual comparisons with NeuS in scene 83 under DTU.

Scan	40	97	110	Mean
UNISURF PE $\times 2$	1.80	1.43	1.55	1.59
Ours PE $\times 2$	**1.07**	**1.26**	**1.24**	**1.19**

Table 5. Numerical comparisons with UNISURF PE $\times 2$ in scene 40 under DTU.

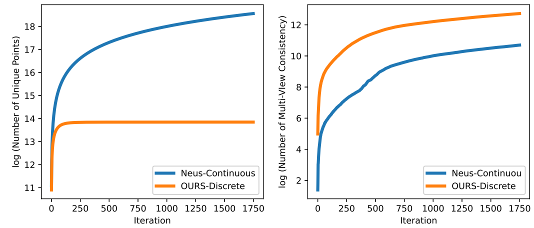

Figure 24. Visual comparisons with NeuS in scene 97 under DTU.

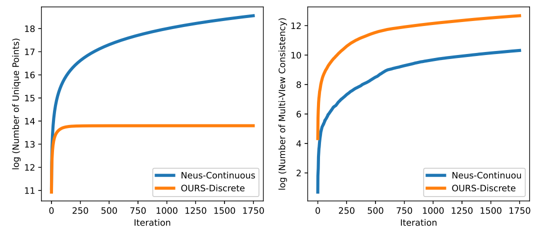

Figure 25. Visual comparisons with NeuS in scene 105 under DTU.

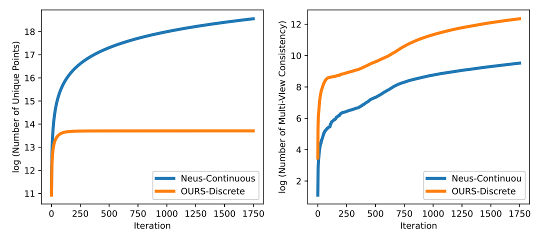

Figure 26. Visual comparisons with NeuS in scene 106 under DTU.

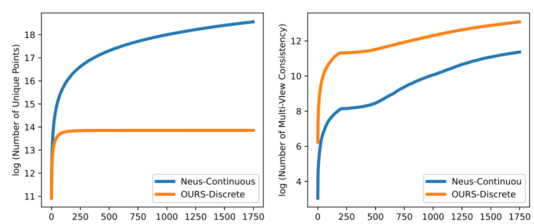

Figure 27. Visual comparisons with NeuS in scene 110 under DTU.

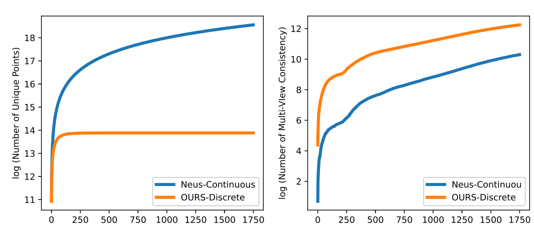

Figure 28. Visual comparisons with NeuS in scene 114 under DTU.

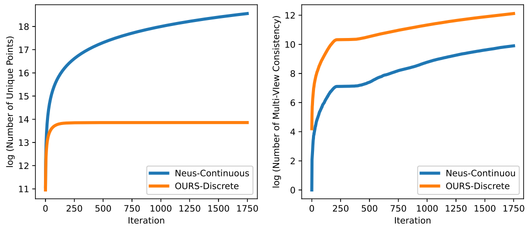

Figure 29. Visual comparisons with NeuS in scene 118 under DTU.

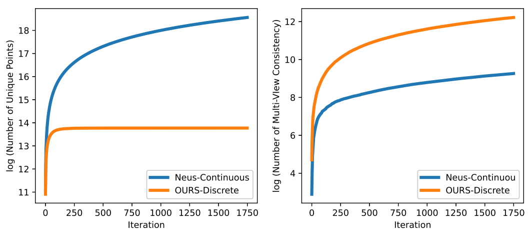

Figure 30. Visual comparisons with NeuS in scene 122 under DTU.

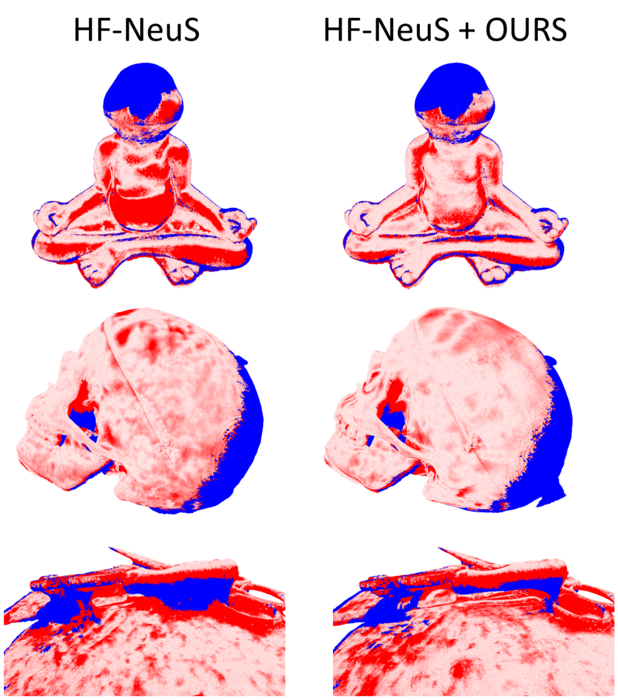

Figure 31. Visual comparisons with HF-Neus in scene 37, 65, 110 under DTU.

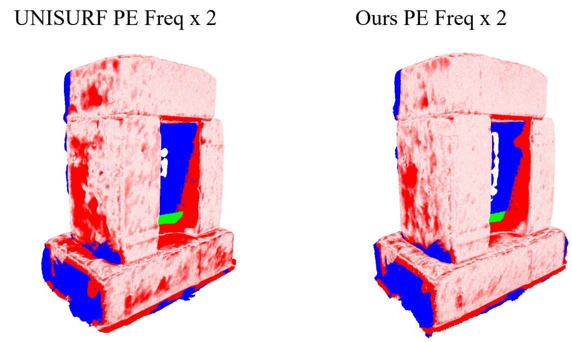

Figure 32. Visual comparisons with UNISURF using 2 times higher frequency in positional encoding in scene 40 under DTU.

References

[1] Amir Hertz, Or Perel, Raja Giryes, Olga Sorkine-Hornung, and Daniel Cohen-Or. Sape: Spatially-adaptive progressive encoding for neural optimization. In *Thirty-Fifth Conference on Neural Information Processing Systems*, 2021. 1

[2] Rasmus Jensen, Anders Dahl, George Vogiatzis, Engil Tola, and Henrik Aanæs. Large scale multi-view stereopsis evaluation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 406–413, 2014. 1

[3] Michael Oechsle, Songyou Peng, and Andreas Geiger. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *International Conference on Computer Vision*, 2021. 1

[4] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. In *Advances in Neural Information Processing Systems*, pages 27171–27183, 2021. 1

[5] Yiqun Wang, Ivan Skorokhodov, and Peter Wonka. HF-NeuS: Improved surface reconstruction using high-frequency details. 2022. 1