

Research on quality evaluation of 3D reconstruction model empowered by AI algorithm

Weihui Zhang^a, Rongyi Wang^b, Zhipeng Ren^a, Cheng Liao^a, Yanqi Huang^a, Shaobin Wang^{*a}

^aSchool of Art and Design, Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China; ^bSchool of Home Furnishing and Art Design, Central South University of Forestry and Technology, Changsha 410004, Hunan, China

ABSTRACT

This paper evaluates four AI-based 3D reconstruction algorithms—TripoSR, Meshy, Instant, and One-2-34—alongside the traditional AR-Code method (which uses scanning for model point cloud reconstruction). By comparing the integrity ratio from point cloud data, depth information from depth maps, and perceived similarity from model dimensions, we identify the AI algorithm that best reconstructs 3D models. This research aligns with the “Opinions on Promoting the Implementation of the National Cultural Digitization Strategy” and provides actionable cases, methods, and strategies. It has practical applications in cultural heritage preservation, product design, and the digital transformation of smart cities, offering resource efficiency and enhanced user experiences in digital environments.

Keywords: Digital model reconstruction under AI algorithm, three-dimensional reconstruction evaluation, digital model evaluation

1. INTRODUCTION

Three-dimensional (3D) reconstruction technology is widely used in computer vision, robot navigation, and digital cultural relics protection¹. The rapid advancement of artificial intelligence (AI) has propelled AI-based 3D reconstruction algorithms to the forefront². Traditional methods, relying on multi-view geometry and structured light³, often struggle with accuracy and efficiency in complex environments and large-scale data.

AI, especially through deep learning and convolutional neural networks (CNNs), offers novel solutions by enhancing accuracy and robustness through vast datasets and feature extraction. AI algorithms can reconstruct high-precision models by interpreting depth information and geometric structures, compensating for partial data loss, and integrating multi-source data, such as RGB images, depth maps, and lidar point clouds.

This study aims to explore and optimize AI applications in 3D reconstruction. We propose an AI-based model quality assessment method, incorporating three indicators: integrity ratio, depth information assessment, and perceived similarity, to comprehensively evaluate and guide the optimization of 3D model reconstruction quality.

2. METHODOLOGY

2.1 Subjects

In studying the optimal path theory for AI-based 3D reconstruction, four algorithms—TripoSR, Meshy, Instant, and One-2-34—were selected based on their popularity. These algorithms were used to model a seat and extract the model's OBJ file. As shown in Figure 1, the seat was chosen for its simple yet distinct structure, clear material contrast, and combination of curved and straight shapes, providing a comprehensive basis for evaluating reconstruction quality. Additionally, the AR Code method, representing traditional 3D reconstruction techniques (using structured light or laser scanning⁴), was employed to scan the seat and extract its OBJ file.

*wangshao1972@qq.com



Figure 1. Reconstruction object.

2.2 Experimental methods

Models were classified into five groups: TripoSR, Meshy, Instant, One-2-34, and AR Code. Point cloud and depth information were extracted, and perceptual similarity was evaluated using computer graphics algorithms. Python was used to extract and visualize point clouds⁵. AI-generated point cloud models were registered with those from AR Code to calculate the integrity ratio (P), a key quality metric indicating the percentage of real object surfaces or point clouds in the reconstruction. A larger P denotes higher quality. The AI reconstruction model's point cloud set is M, and the AR Code model's set is N.

$$P = \frac{M \cap N}{N} (M, N \geq 0) \quad (1)$$

The quality of AI-based algorithms was compared with the AR Code method using depth information from reconstructed models. Python was used to extract and analyze depth maps, assessing integrity, surface smoothness, and noise for a comprehensive evaluation of model accuracy.

For perceptual similarity, the ORB algorithm extracted image feature points, and BFMatcher compared local features. A pre-trained ResNet50 model extracted high-dimensional feature vectors, and cosine similarity measured perceptual similarity, effectively capturing high-level semantic features.





Comparing the four AI algorithms with AR Code identified the optimal AI 3D reconstruction path.






3. RESULTS

3.1 Point cloud evaluation results

From Table 1 and Figures 2 and 3, it can be seen that the four AI 3D reconstruction algorithms based on point cloud evaluation are compared with the AR Code method. The TripoSR algorithm generates the largest number of model points, but the integrity ratio is inferior to the Instant algorithm; the model integrity ratio generated by the Instant algorithm is better than the other three algorithms; the model integrity ratio generated by the One-2-34 algorithm is the lowest, which indicates that there is a large lack of model quality reconstructed by the algorithm among the four algorithms. Instant has the highest integrity ratio.

Table 1. Basic information of subjects.

Algorithm	Point cloud number	Point cloud pictures	Point cloud registration map	Integrity ratio
TripoSR	58096			0.47
Instant	43114			0.49

Algorithm	Point cloud number	Point cloud pictures	Point cloud registration map	Integrity ratio
One-2-34	35876			0.26
Meshy	35552			0.4
AR Code	15758		/	/

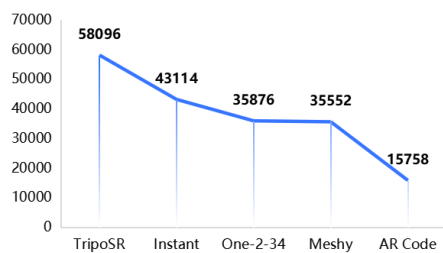


Figure 2. Comparison of model points.

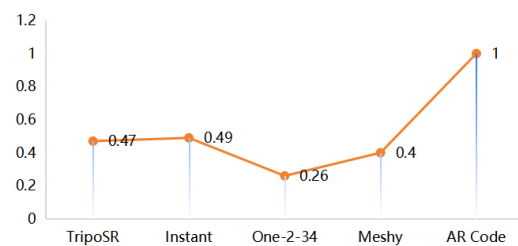


Figure 3. Model integrity ratio comparison.

3.2 Depth information assessment results

Quality assessment using depth maps measures the accuracy and reliability of the 3D model by comparing its depth map with that of the real scene⁶. This evaluation focuses on three dimensions: depth information integrity, surface smoothness, and noise.

Depth information integrity checks if the model captures all essential geometric details without missing areas⁷. Surface smoothness ensures the model is free of unnecessary sharp edges or protrusions, reflecting the real object's smoothness⁸. Noise analysis identifies and evaluates random errors or disturbances from data acquisition, algorithm uncertainties, or environmental factors⁹, determining the model's accuracy and stability and guiding necessary corrections and optimizations.

The depth map of AI reconstruction algorithm based on Python extraction is shown in Figure 4.

According to the extracted depth map, the depth information integrity data and visualization are shown in Table 2 and Figure 5.

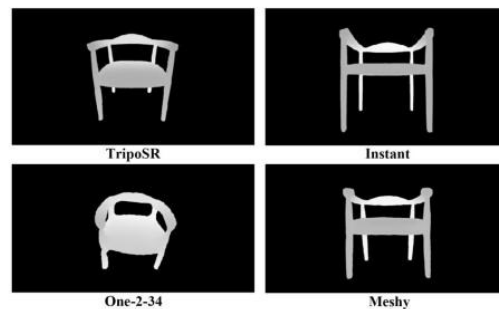


Figure 4. Model depth maps generated by four algorithms.

Table 2. Depth information integrity data.

Algorithm	Mean	Std	Min	25%	50%	75%	Max
TripoSR	-1.288	0.354136	-2.339664	-1.510548	-1.322495	-0.93677	-0.595458
Instant	0.053006	0.499425	-0.739422	-0.467407	0.07406	0.580966	0.8113
One-2-34	-1.48452	0.456114	-2.600965	-1.721241	-1.454214	-1.070674	-0.682965
Meshy	-2.383768	0.423762	-3.53	-2.657742	-2.387465	-1.93283	-1.612

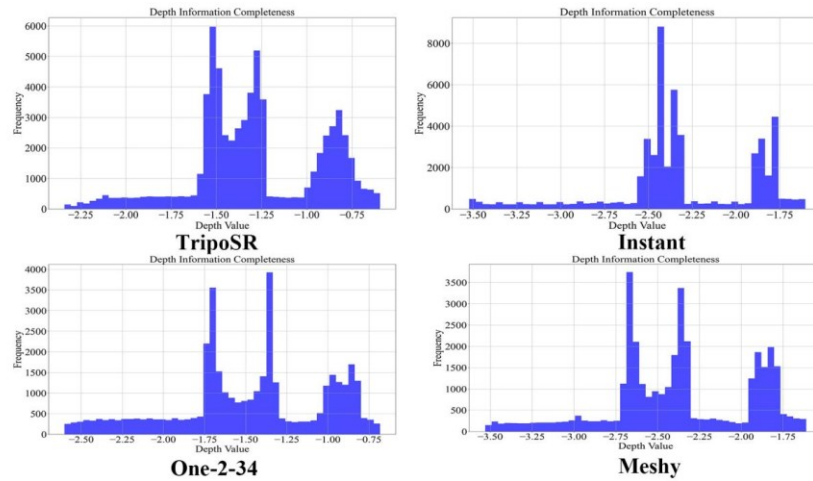


Figure 5. The model depth information integrity map generated by the four algorithms.

From the analysis of depth information integrity in Table 2 and Figure 5, the following conclusions can be drawn:

- (1) The TripoSR algorithm's reconstructed model has an average depth value of -1.288, ranging from -2.339664 to -0.595458, with a standard deviation of 0.354136. This indicates a concentrated distribution and consistent depth information.
- (2) The Instant algorithm's model has an average depth value of 0.053006, ranging from -0.739422 to 0.8113, with a standard deviation of 0.499425. The scattered distribution and large standard deviation suggest significant depth information variation and potential reconstruction bias.
- (3) The One-2-34 algorithm's model has an average depth value of -1.48452, ranging from -2.600965 to -0.682965, with a standard deviation of 0.456114. The scattered depth values indicate lower accuracy in depth information compared to TripoSR.
- (4) The Meshy algorithm's model has an average depth value of -2.383768, ranging from -3.53 to -1.612, with a standard deviation of 0.423762. Although the depth values are concentrated, the overall depth tends to be lower.

In summary, TripoSR performs best in-depth information integrity, exhibiting concentrated depth value distribution and smaller standard deviation, indicating higher accuracy and stability. Meshy, while concentrated, shows lower overall depth values. Instant and One-2-34 have more dispersed depth value distributions, indicating potential depth information deviations.

According to the extracted depth map, the surface smoothness data and visualization of the reconstructed model are shown in Table 3 and Figure 6.

The following conclusions can be drawn from the analysis of the surface smoothness of the model generated by Table 3 and Figure 6:

Table 3. Reconstruction of model surface smoothness data.

Algorithm	Mean	Std	Min	25%	50%	75%	Max
TripoSR	0.000019	0.133434	-0.750202	-0.000451	0.00593	0.042052	0.703654
Instant	-0.000015	0.123193	-0.880765	0	0.00789	0.016561	0.880038
One-2-34	1.426703	9.66497	-2.819716	-2.464393	-3.707469	3.198296	8.883726
Meshy	-0.000017	0.054985	-0.93487	-0.000468	0.000436	0.005531	0.747142

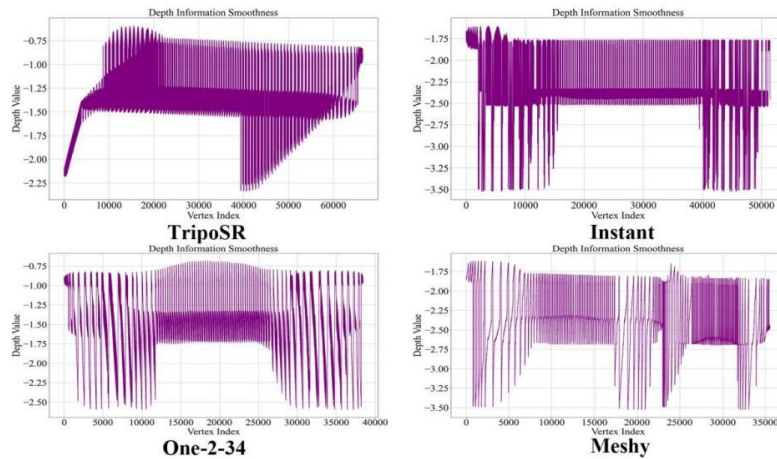


Figure 6. The surface smoothness analysis diagram of the model generated by the four algorithms.

The surface gradient analysis reveals the following:

- (1) TripoSR: Average gradient value of 0.000019, range from -0.750202 to 0.703654, and standard deviation of 0.133434. The surface is smooth with minimal gradient changes.
- (2) Instant: Average gradient value of -0.000015, range from -0.880765 to 0.880038, and standard deviation of 0.123193. The surface is relatively smooth, similar to TripoSR.
- (3) One-2-34: Average gradient value of 1.426703, range from -2.819716 to 8.883726, and standard deviation of 9.66497. The surface is rough with significant gradient changes.
- (4) Meshy: Average gradient value of -0.000017, range from -0.93487 to 0.747142, and standard deviation of 0.054985. The surface is the smoothest with the smallest gradient changes.

In terms of surface smoothness, Meshy performs the best with minimal gradient changes. TripoSR and Instant also exhibit good surface smoothness with small gradient changes. In contrast, One-2-34 has the worst surface smoothness with the largest gradient changes and a rough surface.

The noise data and visualization of the reconstructed model based on the extracted depth map are shown in Table 4 and Figure 7.

Table 4. Reconstruct model noise data.

Algorithm	Mean	Std	Min	25%	50%	75%	Max
TripoSR	-1.288908	0.354136	-2.339664	-1.510548	-1.322495	-0.93677	-0.595458
Instant	-2.319911	0.407868	-3.522756	-2.466499	-2.366907	-1.881537	-1.604812
One-2-34	-1.48452	0.456114	-2.600965	-1.721241	-1.454214	-1.070674	-0.682965
Meshy	-2.383768	0.423762	-3.53	-2.657742	-2.387465	-1.93283	-1.612

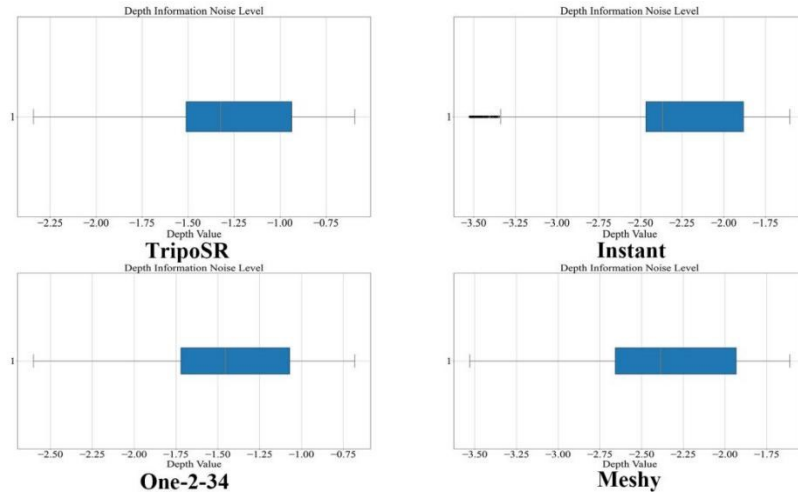


Figure 7. Model noise analysis diagrams generated by four algorithms.

From the noise analysis in Table 4 and Figure 7, the following conclusions can be drawn:

- (1) TripoSR: Average noise value of -1.288908, range from -2.339664 to -0.595458, and standard deviation of 0.354136. The noise distribution is concentrated with a smaller standard deviation.
- (2) Instant: Average noise value of -2.319911, range from -3.522756 to -1.604812, and standard deviation of 0.407868. The noise distribution is more dispersed with a larger standard deviation.
- (3) One-2-34: Average noise value of -1.48452, range from -2.600965 to -0.682965, and standard deviation of 0.456114. The noise distribution is more dispersed with a larger standard deviation.
- (4) Meshy: Average noise value of -2.383768, range from -3.53 to -1.612, and standard deviation of 0.423762. The noise distribution is concentrated with a smaller standard deviation.

In summary, TripoSR performs best in noise analysis, with concentrated distribution and smaller standard deviation, indicating minimal random error. Meshy also has concentrated noise distribution. In contrast, Instant and One-2-34 have more dispersed noise distributions and larger standard deviations, indicating greater random errors.

Table 5 shows the comprehensive comparison of the four AI reconstruction models in depth information integrity, surface smoothness, and noise analysis.

Table 5. The advantages and disadvantages of the depth information of the reconstruction models generated by the four AI algorithms are compared.

Algorithm	Dominance	Inferiority	Overall merit
TripoSR	The depth information is more complete, surface smoothness is better ¹⁰ .	The depth value bias is low.	The overall performance is excellent and the reconstruction quality is high.
Instant	The surface smoothness is good.	The depth information and noise distribution are scattered, and there are large deviations and errors.	The overall performance is general, and the reconstruction quality is medium.
One-2-34	No advantage.	The smoothness is the worst, and the noise distribution is more dispersed.	The overall performance is poor and the reconstruction quality is low.
Meshy	The surface smoothness is the best, and the noise distribution is more concentrated.	The depth value bias is low.	The overall performance is excellent, especially in terms of surface smoothness.

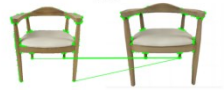


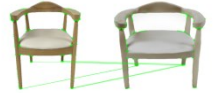
3.3 Perceptual similarity evaluation results

Perceptual similarity measures the visual similarity between two images, focusing on human visual perception rather than traditional pixel-level comparisons like mean square error (MSE) or peak signal-to-noise ratio (PSNR)¹¹. This approach aims to simulate human visual perception for a more accurate assessment of image quality and similarity.

In addition to evaluating point cloud and depth information, perceptual similarity is also assessed. Using deep learning algorithms, the feature points of the 3D reconstruction model and the real object are matched and calculated to determine the similarity between the reconstructed model and the real object.

It can be seen from Table 6 that among the three-dimensional reconstruction models of the four algorithms, the three-dimensional reconstruction model based on the TripoSR algorithm has the highest similarity with the real object, followed by the three-dimensional reconstruction model generated by the Instant algorithm. The model based on One-2-34 algorithm has the lowest similarity, and it can be seen from the picture that the model generated by One-2-34 algorithm has the characteristics of rough surface and incomplete features.

Table 6. Perceptual evaluation results of reconstruction models generated by four AI algorithms.

Algorithm	Feature point matching graph	Perceived similarity
TripoSR		0.975
Instant		0.965
One-2-34		0.628
Meshy		0.882

4. DISCUSSION

In 3D reconstruction, various quality assessment methods have distinct pros and cons. Traditional optimization methods, such as those based on geometric and photometric consistency, are robust and converge well without needing training data. However, they incur high computational costs and require complex strategies to handle noise, outliers, and density changes.

Feature learning-based methods, leveraging deep neural networks and algorithms like RANSAC, excel in accurate matching but demand extensive training data. Their performance drops significantly when faced with unknown scenarios with differing data distributions.

End-to-end learning methods simplify the reconstruction process by transforming registration into a regression problem, merging traditional mathematical theory with deep learning. However, they are sensitive to noise and density differences.

For large-scale 3D reconstruction, combining multi-view images shows promise. Techniques like Structure from Motion (SfM) and Multi-View Stereo (MVS) generate high-precision models for large scenes but need further optimization in computational complexity and data processing efficiency.

This study proposes an optimized AI algorithm to enhance reconstruction accuracy and efficiency while better handling large-scale, complex scene data. This new method is theoretically innovative and practically advantageous, opening new avenues for 3D reconstruction technology development.

5. CONCLUSIONS

A comprehensive evaluation indicates that AI algorithms, through multi-view synthesis, data augmentation, and interpolation techniques, significantly enhance the effectiveness of 3D model reconstruction, resulting in denser and more complete point cloud data. Among these algorithms, TripoSR excels in point cloud integrity, depth information, and perceptual similarity, making it the optimal choice for 3D reconstruction. Meshy performs well in terms of depth information and ranks second in point cloud integrity. Instant demonstrates the best performance in point cloud integrity but falls short in other aspects. The One-2-34 algorithm performs the worst across all dimensions.

Despite the overall superior performance of AI-based 3D reconstruction algorithms, there is still room for improvement in edge processing, such as edge smoothness. Additionally, 3D models reconstructed using AI algorithms need improvements in the number and quality of faces to enhance the model's subdivision level.

Future research could focus on exploring the optimal integration path between AI algorithms and traditional 3D reconstruction scanning methods. Combining the advantages of AI algorithms with the stability and precision of traditional methods may further enhance the overall effectiveness of 3D reconstruction. Moreover, AI algorithms hold great potential in the field of real-time 3D reconstruction. Real-time processing and generation of high-quality 3D models could have profound impacts on fields such as virtual reality, augmented reality, and real-time monitoring. Investigating how to optimize AI algorithms to meet the demands of real-time processing and how to maintain high-precision reconstruction in dynamic environments will be crucial areas for future research.

ACKNOWLEDGEMENTS

2022 Guangdong Philosophy and Social Sciences Lingnan Culture Project 'Lingnan Cantonese Business Building Cultural Heritage Research' (GD22LN06) Phased Research Results.

REFERENCES

- [1] Zhou, L., Wu, G., Zuo, Y., Chen, X. and Hu, H., "A comprehensive review of vision-based 3D reconstruction methods," *Sensors* 24(7), 8-12 (2024).
- [2] Luo, H., Zhang, J., Liu, X., Zhang, L. and Liu, J., "Large-scale 3D reconstruction from multi-view imagery: A comprehensive review," *Remote Sensing* 16(5), 2-5 (2024).
- [3] Zheng, T. X., Huang, S., Li, Y. F., et al., "Key techniques for vision-based 3D reconstruction: a review," *Acta Automatica Sinica* 46(4), 631-652 (2020).
- [4] Rocchini, C., Cignoni, P., Montani, C., et al., "A low cost 3D scanner based on structured light," *Computer Graphics Forum* 20(3), 299-308 (2001).
- [5] Huang, X., et al., "A comprehensive survey on point cloud registration," *arXiv arXiv:2103.02690*, (2021).
- [6] Lorensen, W. E., "Marching cubes: A high resolution 3D surface construction algorithm," *Computer Graphics* 21(1), 7-12 (1987).
- [7] Luo, W., Lu, Z. and Liao, Q., "LNMVSNet: A low-noise multi-view stereo depth inference method for 3D reconstruction," *Sensors* 24(8), 5-7 (2024).
- [8] Hong, Z., Yang, Y., Liu, J., Jiang, S., Pan, H., et al., "Enhancing 3D reconstruction model by deep learning and its application in building damage assessment after earthquake," *Applied Sciences* 12(19), 2-4 (2022).
- [9] Zheng, P., Gao, D., Fan, D. P., et al., "Bilateral reference for high-resolution dichotomous image segmentation," *arXiv arXiv:2401.03407*, (2024).
- [10] Tochilkin, D., Pankratz, D., Liu, Z., et al., "TripoSR: Fast 3D object reconstruction from a single image," *arXiv arXiv:2403.02151*, (2024).
- [11] Yang, L., Kang, B., Huang, Z., et al., "Depth anything: Unleashing the power of large-scale unlabeled data," *arXiv arXiv:2401.10891*, (2024).