DOI: 10.53469/jpce.2025.07(08).04

From 2D to 3D: A Review of Monocular Depth Reconstruction Algorithms and Their Application in Rendering Pipelines

Shilpi Gupta[1], Wani Ladu Worja[2]

[1]Professor, Department of CSE, Jaya Engineering College

[2]Assistant Professor, Department of CSE, Jaya Engineering College

Abstract: ***Depth reconstruction from 2D images is a critical process for generating realistic 3D visualizations used in applications such as virtual reality (VR), augmented reality (AR), robotics, and autonomous systems. Over the past decade, numerous methods have been developed to estimate depth from 2D data, ranging from traditional geometric approaches to advanced machine learning models. This paper provides a comprehensive survey of these techniques, categorizing them based on methodology, computational efficiency, and application domains. Additionally, the role of depth reconstruction in enhancing 3D rendering quality is discussed. The survey concludes by highlighting current challenges and potential research directions to bridge the gap between academic advancements and real - world implementation.***

Keywords: Depth reconstruction, 2D images, 3D rendering, Virtual reality, Augmented reality, Geometric methods, Machine learning

1. Introduction

The transformation of 2D images into 3D representations is fundamental for numerous applications that demand realistic visualization and spatial understanding. Accurate depth reconstruction is a key component of this transformation, enabling systems to infer the spatial arrangement of objects and scenes. Applications such as immersive VR/AR experiences, autonomous navigation, and 3D modeling rely heavily on robust depth estimation techniques.

This paper surveys the evolution of depth reconstruction methods, from traditional algorithms based on geometric principles to modern approaches leveraging deep learning. By examining their advantages, limitations, and real - world applicability, the study aims to provide a holistic view of the field while emphasizing the importance of depth data in high - quality 3D rendering.

2. Categories of Depth Reconstruction Methods

2.1 Traditional Geometric Methods

- **Stereo Vision**: Uses disparities between two images captured from slightly different viewpoints to estimate depth. Effective in controlled settings but limited by occlusions and texture - less regions.
- **Structure - from - Motion (SfM):** Analyzes motion between frames to estimate depth, suitable for dynamic environments but computationally intensive.
- **Shape - from - Shading**: Infers depth based on light and shadow analysis. While it works well in uniform lighting conditions, it struggles with complex lighting.

2.2 Machine Learning - Based Methods

- **Supervised Deep Learning**: Convolutional Neural Networks (CNNs) and transformer - based models predict depth from single or multiple images. These methods require large labeled datasets but provide high accuracy.
- **Unsupervised Learning**: Eliminates the need for labeled data by leveraging constraints such as photometric consistency. However, they are often less precise compared to supervised models.

2.3 Hybrid Approaches

Combining traditional geometric methods with machine learning has emerged as a promising trend. These approaches use geometric models to provide initial depth estimates, which are then refined by ML models, achieving a balance between accuracy and efficiency.

3. Depth Reconstruction in 3D Rendering Applications

The quality of 3D rendering is significantly influenced by the accuracy of depth data. Key applications include:

3.1 Virtual and Augmented Reality

Accurate depth reconstruction enhances immersion by aligning virtual objects with the real - world environment in AR applications and creating realistic simulations in VR.

3.2 Gaming

Depth data facilitates dynamic interactions between virtual characters and environments, enabling more lifelike experiences.

3.3 Robotics and Automation

Robots rely on depth information for navigation and object manipulation, particularly in dynamic or unstructured environments.

3.4 Digital Content Creation

Depth reconstruction enables the generation of 3D models from 2D sketches or photographs, streamlining workflows in industries such as animation and architecture.

4. Comparative Analysis of Methods

The table below summarizes the key features of various depth reconstruction methods:

Method	Accuracy	Speed	Challenges	Applications
Stereo Vision	High	Moderate	Occlusions, texture - less areas	VR, robotics
Structure - from - Motion (SfM)	Moderate	Low	High computational cost	AR, film production
Shape - from - Shading	Moderate	High	Complex lighting conditions	Gaming, content creation
Supervised ML	High	Moderate	Dataset dependency	All
Unsupervised ML	Moderate	High	Less precision	VR, robotics
Hybrid Methods	Very High	Moderate	Model complexity	Real - time applications

5. Challenges in Depth Reconstruction

5.1 Occlusions and Texture - Less Regions

Most methods struggle to estimate depth in occluded areas or regions lacking texture, leading to incomplete reconstructions.

5.2 Computational Efficiency

Real - time applications demand fast and efficient depth estimation, which remains a challenge for high - accuracy methods.

5.3 Scalability and Generalization

Many models perform well in controlled environments but fail to generalize to diverse and dynamic scenarios.

6. Future Research Directions

- **Multi - Modal Depth Estimation**: Integrating data from LiDAR, infrared sensors, and cameras to enhance depth accuracy.
- **Resource - Efficient Models**: Developing lightweight models for deployment on mobile devices and AR/VR headsets.
- **Dynamic Scene Adaptation**: Improving algorithms to handle rapid scene changes in real time.

7. Conclusion

Depth reconstruction from 2D images is a cornerstone of modern 3D rendering applications. This paper provides a detailed survey of existing methods, categorizing them based on their approaches, strengths, and limitations. Traditional geometric techniques, while foundational, face limitations in complex environments, whereas ML - based methods provide robust solutions at the cost of computational demands. The integration of these methods into hybrid frameworks represents a promising path forward, addressing both accuracy and efficiency. Future advancements in this field will play a pivotal role in revolutionizing industries that rely on realistic 3D visualizations.

References

[1] Scharstein, D., & Szeliski, R. (2002). A Taxonomy and Evaluation of Dense Two - Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision.*

[2] Eigen, D., et al. (2014). Depth Map Prediction from a Single Image Using a Multi - Scale Deep Network. *Advances in Neural Information Processing Systems.*

[3] Godard, C., et al. (2017). Unsupervised Monocular Depth Estimation with Left - Right Consistency. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

[4] Dosovitskiy, A., et al. (2021). An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*