S6 MINI PROJECT PRESENTATION 0

2D TO 3D RECONSTRUCTION AND RENDERING

Guided by:
Prof. Dr. Saritha R

CSE S6 B2

TEAM 1	Roll no	KTU ID
Goureesh Chandra	31	TVE22CS069
Ivin Mathew	33	TVE22CS075
Muhammed Farhan	39	TVE22CS094
Rethin Francis	72	LTVE22CS149

Contents

Introduction

- The integration of computer vision and real-time rendering techniques offers a practical solution for creating viewable 3D content from standard photographs.

Motivation

This project addresses the increasing need for efficient 3D content creation in gaming, VR, and digital preservation. It simplifies 3D modeling, making it accessible to non-experts through advanced vision and rendering techniques.

Problem Statement

- Develop an application to accurately transform 2D images into realistic 3D models using advanced computer vision and graphics rendering techniques.

Literature Review

Sl. no.	Name of Publication	Journal / Conference	Author(s)	Date of publication	Remarks
1	Instant Neural Graphics Primitives with a Multiresolution Hash Encoding	ACM Trans. Graph., Vol. 41, No. 4, Article 102	Thomas Müller, Alex Evans, Christoph Schied, Alexander Keller	July 2022	**Description:** This paper introduces a multiresolution hash table for neural graphics primitives, enabling tasks like NeRF and gigapixel image representation with high-quality results in seconds. **Advantage:** Significantly accelerates training times while preserving high reconstruction quality **Disadvantage:** Hash collisions can cause minor artifacts in high-frequency detail reconstructions.
2	Vision Transformers for Dense Prediction	2021 IEEE/CVF International Conference on Computer Vision (ICCV)	René Ranftl, Alexey Bochkovskiy, Vladlen Koltun	10 October 2021	**Description:** Introduces **Dense Prediction Transformers (DPT)**, which utilize vision transformers instead of convolutional networks for dense prediction tasks like monocular depth estimation **Advantages:** Improved Accuracy and Scalability

3	Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer	IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 44, Issue: 3, 01 March 2022)	Rene Ranftl, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun	27 August 2020	**Description:**The paper proposes techniques to improve monocular depth estimation by mixing datasets for better generalization. **Advantage:** Enhances robustness across diverse real-world scenarios. **Disadvantage:**Requires significant computational resources for training.
4	PCRNet: Point Cloud Registration Network using PointNet Encoding	Conference on Computer Vision and Pattern Recognition (CVPR), 2018.	Vinit Sarode, Xueqian Li, Hunter Goforth, Yasuhiro Aoki, Rangaprasad Arun Srivatsan, Simon Lucey, Howie Choset	August 2019	**Description:** PCRNet leverages PointNet for efficient 3D point cloud registration, aligning source and template clouds for tasks like pose estimation. **Advantage:** Fast and robust to noise, suitable for real-time applications. **Disadvantage:** Struggles with unseen shapes or partial point clouds.

5	A Review of Deep Learning Techniques for 3D Reconstruction of 2D Images	12th International Conference on Information & Communication Technology and System (ICTS) 2019	Anny Yuniarti, Nanik Suciati	18 July 2019	**Description:** discussing methods, representations, applications, and research opportunities. **Advantage:** Provides a comprehensive overview **Disadvantage:** Excludes RGB-D-based reconstruction methods.
6	Occupancy Networks: Learning 3D Reconstruction in Function Space	2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowoziny, Andreas Geiger	15 June 2019	**Description:** The paper proposes methods for combining diverse datasets to improve monocular depth estimation models, enabling zero-shot cross-dataset generalization through novel loss functions and multi-dataset training strategies. **Advantages:**Achieves results in monocular depth estimation across diverse datasets.Enables robust generalization to unseen datasets using zero-shot cross-dataset transfer. **Disadvantages:** High computational cost, requiring extensive training resources.Complexity in dataset preparation and alignment due to variations in scale, depth, and quality across sources

7	Pix2Vox: Context-aware 3D Reconstruction from Single and Multi-view Images	International Conference on Computer Vision (ICCV), 2019.	Haozhe Xie, Hongxun Yao, Xiaoshuai Sun, Shangchen Zhou, Shengping Zhang, Wenxiu Sun	2019	**Description**: Pix2Vox is a framework for single-view and multi-view 3D reconstruction, using a context-aware fusion module to combine high-quality parts from multiple 3D volumes and refine them into accurate final outputs. **Advantage**: Outperforms state-of-the-art methods in accuracy and efficiency, with strong generalization to unseen 3D objects. **Disadvantage**: Reconstruction is limited to low resolution and computational complexity increases with higher precision models.
8	An Improved Method for Building A 3D Model from 2D DICOM	2018 International Conference on Advanced Computing and Applications (ACOMP)	Van Sinh NGUYEN, Manh Ha TRAN, Hoang Minh Quang VU	27 November 2018	**Description:** An improved method for faster and accurate 3D reconstruction from 2D DICOM images, aiding medical diagnostics and visualization. **Advantage:**Efficient, accurate, practical for medical use, and cost-effective. **Disadvantage:**Limited comparisons, hardware-dependent, restricted scope, and preprocessing reliance.

9	3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction	European Conference on Computer Vision (ECCV), 2016.	Christopher B Choy, Danfei Xu, JunYoung Gwak,Kevin Chen, Silvio Savarese	2 April 2016	**Description:** The paper introduces a deep learning framework for reconstructing 3D objects from single or multiple 2D images using a recurrent neural network **Advantage:**Efficiently handles single and multi-view reconstruction without needing dense annotations. **Disadvantage:**Faces difficulty with thin structures and objects with high shape variability.
10	WeatherGS: 3D Scene Reconstruction In Adverse Weather Conditions via Gaussian Splatting		Chenghao Qian and Yuhu Guo and Wenjing Li and Gustav Markkula	30 December 2024	**Description**:WeatherGS leverages 3D Gaussian Splatting for reconstructing clear 3D scenes in adverse weather, using a dense-to-sparse strategy to remove weather artifacts like snow, rain, and lens occlusions. **Advantage**: Produces high-quality, artifact-free reconstructions with efficient computation suitable for real-time applications. **Disadvantage**: Relies on pretrained models and may slightly distort fine structural details during preprocessing.

Objective

Accurate Depth Estimation: Leverage MiDaS to generate high-quality depth maps from 2D images.

Efficient 3D Reconstruction: Use Open3D to create refined point clouds and smooth meshes.

Texture and Visual Fidelity: Employ custom shaders to preserve and enhance the image’s original textures.

Software and Hardware Requirements

Programing Language:

- Python

Major Libraries:

- OpenCv:For image processing,depth estimation and features
- Numpy:For numerical computations and handling multi-dimensional arrays
- Open3D:For working data,such as point clouds and meshes

Model Used:

- Midas model is a pytorch-based depth estimation model for converting 2D images int spatial depth maps

Future scope and Applications

Virtual Reality (VR) Content Creation:

Generate realistic 3D assets from photographs for immersive VR experiences.

Satellite Image Rendering:

Create 3D representations of terrains or urban areas from satellite imagery.

Gaming and Entertainment:

Quickly generate 3D assets from reference images for games or animations.

Education and Training:

Produce interactive 3D models for educational tools and training simulations.

Advantages

Cost-Effectiveness: Eliminates the need for expensive 3D scanning hardware.

Time Efficiency: Significantly reduces the time required for 3D asset creation.

Scalability: Can be applied across multiple domains and industries.

High Visual Quality: Ensures realistic textures and details using advanced shaders.

Accessibility: User-friendly interface makes the tool usable by non-experts.

Flexibility: Adaptable to various image types and resolutions.

Real-Time Applications: Optimized for efficient processing suitable for dynamic use cases.

TIMELINE

Week 1
MiDaS Exploration and Initial Setup

Week 2
Open3D Basics and Integration

Week 3
Refining Depth-to-Point Cloud Conversion

Week 4
Research Custom Shading Techniques

Week 5
Integrating Depth and RGB Information

Week 6
Mesh Generation from Point Clouds

Week 7
Custom Shading Implementation

Week 8
Testing and Optimization

Week 9
Refining Reconstruction Quality

Week 10
Adding User Interface and Documentation

Week 11
Final Refinements

Week 12
Testing and Deployment

Conclusion

At this stage of our project, we have successfully completed the exploration of MiDaS for depth estimation and the initial setup, along with gaining foundational knowledge of Open3D and its integration into our workflow. Additionally, we have progressed to researching custom shading techniques, which will play a critical role in enhancing the visual quality of our 3D reconstructions.

References

1. Park, J., Zhou, Q.-Y., & Koltun, V. (2017). Colored point cloud registration revisited. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
2. Song, S., Lichtenberg, S., & Xiao, J. (2015). SUN RGB-D: A RGB-D scene understanding benchmark suite. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
3. Ranftl, R., Lasinger, K., Hafner, D., Schindler, K., & Koltun, V. (2020). Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
4. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*.
5. https://pytorch.org/hub/intelisl_midas_v2/
6. https://www.open3d.org/