# How can we convert 2D pixel art into 3D voxels?

Tihana Krajtmajer
tkrajtmajer@tudelft.nl

Responsible Professor: Petr Kellnhofer | Supervisor: Mathijs Molenaar

TU Delft

## Introduction

Reconstructing 3D models from 2D projections is a well-explored topic in Computer Graphics. There are several standard methods, based on different image cues [1].

This project focuses on implementing **three** of them:

- **silhouette intersection**
- **spatial carving**
- **gradient-based depth estimation**

Additionally, there is no simple tool for artists to convert their existing 2D artwork into 3D voxel models. A simple **user editor** is also implemented.

## Research Question

The main research question is: "*How can multiple orthographic 2D pixel images from different viewpoints be used to construct a 3D voxel representation using various fill methods*".

**Subquestions:**

- how accurate is the model reconstruction?
- how well do the algorithms generalize unseen views?
- how does adding user input improve shape quality?

## Methodology

### Silhouette intersection

- pixels are projected from the 2D image along the corresponding depth axis
- if the projections overlap over a threshold value, the corresponding point in the 3D voxel grid is filled in

### Spatial carving

- the grid is completely filled in at the start, pixels are projected from each image
- if the overlap

### Gradient-based Depth Estimation

- gradients are calculated for each image, filtered to remove edges and concavities are detected
- the depth is estimated based on pixel intensities of the concave regions

### Hybrid Approach

- silhouette intersection was combined with depth estimation

## Limitations

- limited number of models used for results
- accuracy is high, but can be improved

## Results

| | Average Accuracy |
|---|---|
| Silhouette Intersect | 0.714 |
| Spatial Carving | 0.740 |
| Depth Estimation | 0.608 |
| Hybrid Method | 0.750 |

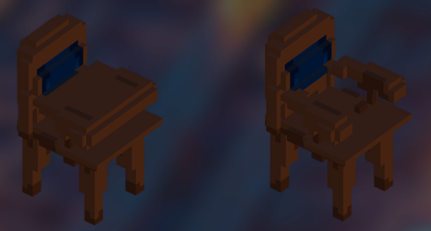

Fig 1: results of using user annotations


Fig 2: results of running the four algorithms

## Conclusion

- the hybrid method performs the best in terms of accuracy
- silhouette intersection and spatial carving give the best results when limited views are available
- user annotations remove excess geometry

- future work should focus on extensions to perspective projections, finding different color merging strategies, adding custom viewpoints

## References

[1] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 1, 2006, pp. 519–528. DOI: 10.1109/CVPR.2006.19.