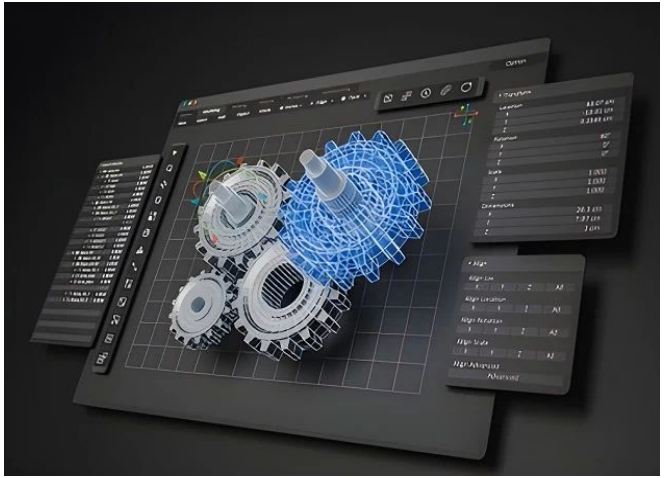
Image-Guided 3D Reconstruction and Natural Language Editing

Group members: Kanishk Sivanandam, Chetan Maviti, Aryan Gadre
Computer Systems Lab: Dr. Yilmaz and Dr. Gabor
Date: 5/21/2025
Periods: 1, 2, 5

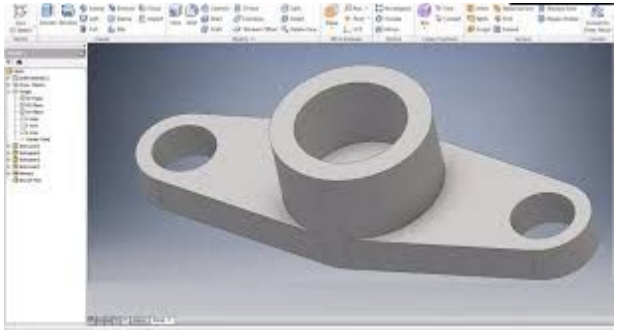
Problem

- Current 3D modeling methods require expertise in specialized software like CAD tools
  - Inaccessible for beginners and time consuming
- Replicating 3D models of real world objects is also a complex task
- No simple methods to modify existing 3D models to better fit a user's needs
- Need for an efficient system that allows 3D model creation without complex software or time.



Background

- 3D modeling softwares such as Blender or Autodesk can create models from scratch, or by meshing multiple pictures
- In the picture, a blender mesh using pictures of a castle can be seen
    - Creating these meshes consumes a lot of time to add appropriate detail
    - Software to create these models consume a lot of storage
- 3D models created by hand are going to be more accurate to the user's wishes, but software is expensive and has a steep learning curve.



Other Solutions

- 3D Modeling Software
  - Blender, Fusion 360, Autodesk
- LiDAR Scanning
  - Uses lasers to measure the distance to a target, generating a point cloud
  - Pros: Very high resolution 3D models, high speed and long range data
  - Cons: High cost and unportable, requiring complex equipment
- Photogrammetry/AI Software
  - Uses multiple photographs to calculate 3D positions to generate a model
  - Pros: Low cost technology, portable, capable of detailed generations
  - Cons: Good lighting conditions necessary
  - Existing AI softwares like 3DFY.AI and NVIDIA Omniverse

Why is Ours Better?

- Simple and Accessible
  - Does not utilize any special hardware or overly complex softwares.
  - Even users with minimal technical experience can benefit; basic commands drive the software.
- Highly flexible
  - Numerous use cases
    - Urban Planning
    - Furniture Customization
    - Custom Case Creation
- Editing capabilities

Novelty

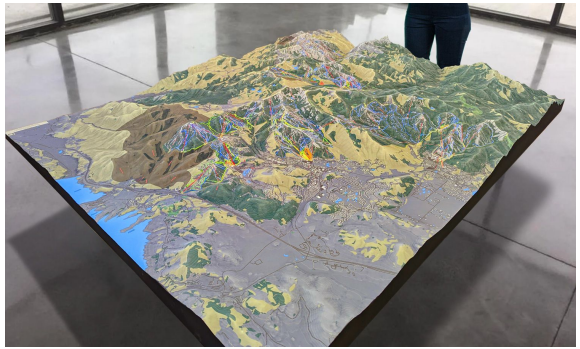
- In our research, we could not find a complete workflow for creating 3D models given a real-world object and be able to edit it given a user's specifications
  - The closest that we have come is NVIDIA's Omniverse, which can create a 3D environment using its library of 3D models
  - However we have not found any software that can create and edit 3D models in one workflow
- Editing capabilities, i.e. extend by 5cm, rotate, scale up/down, etc.:
  - We used scripts to turn instructions from a variety of inputs into an edited and more concrete model
  - There are multiple softwares available that can use AI to create models from scratch
    - However we did not find software that can edit 3D models accurately

Impact

- **Accessibility:** Makes 3D modeling accessible to a broader audience, available to non experts

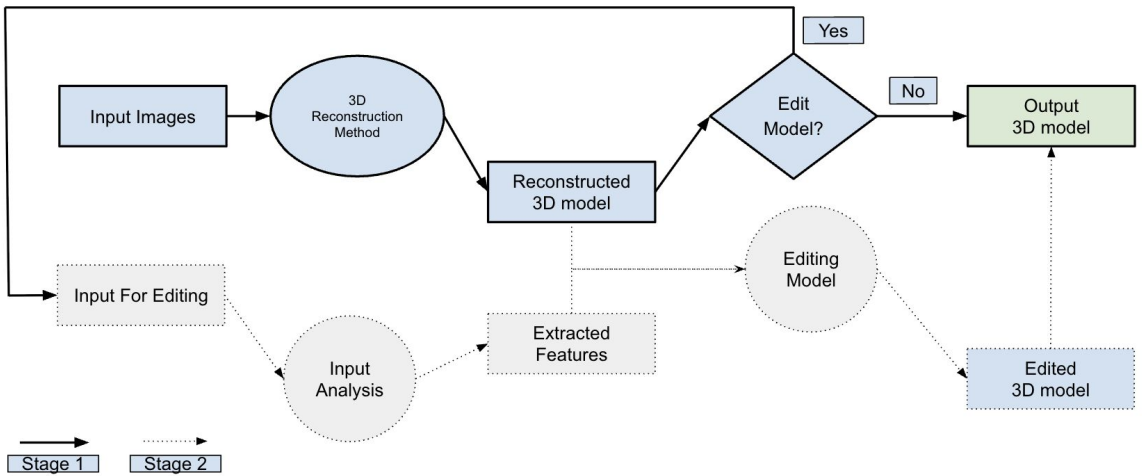
- **Efficiency:** Allows for fast model creation and editing, saving time and resources

- **Allows for creative use cases:**
  - Topographic map modeling
  - Case creation for objects
  - Interior Design



Methods: General Systems Architecture

- The input images are used to create a 3D reconstruction of the initial object
- The input for editing goes through unique input analysis for each use case to extract requested edits
- The requested features are then applied through an editing model for each case and return an edited 3D model



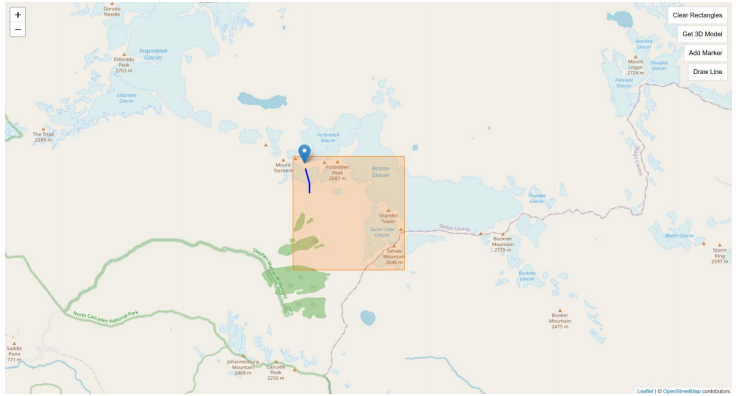
Project Focus Use Cases

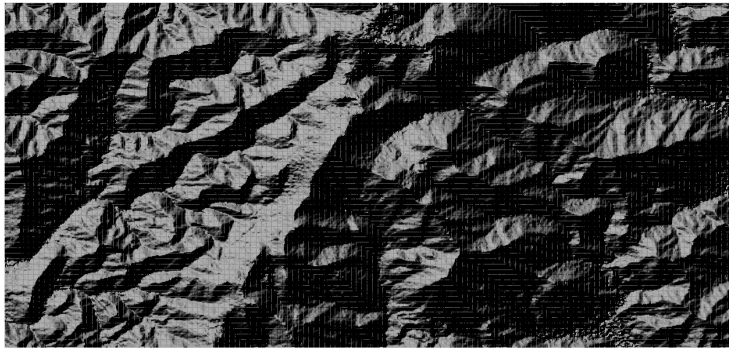
Aryan → Topographic Map Modeling

- Ability to select areas on a map and generate a 3D model of the topographic area
- Topographic maps include representations of elevation, shape, and features of the land
- Edit feature adds annotations to maps
- Useful for engineering and urban design

Methods - 3D Relief Map

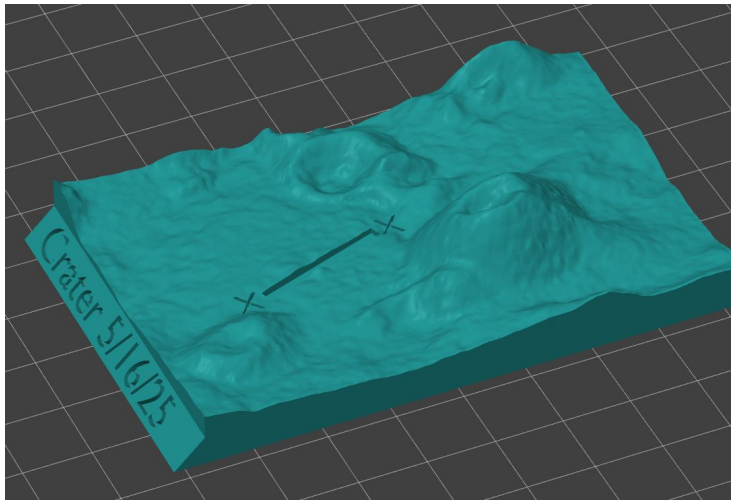
- Area Selection
  - User selects area of interest on web ui as well as name for the map
  - User selects markers and trails to annotate the map
  - The coordinates (latitude, longitude) of the selected area are captured and stored
- DEM Creation
  - Area selected by user is passed on to OpenTopo api to retrieve geographic data in order to create heightmap
  - The returned DEM data can be processed to create a 2D array or matrix that represents elevation of the area



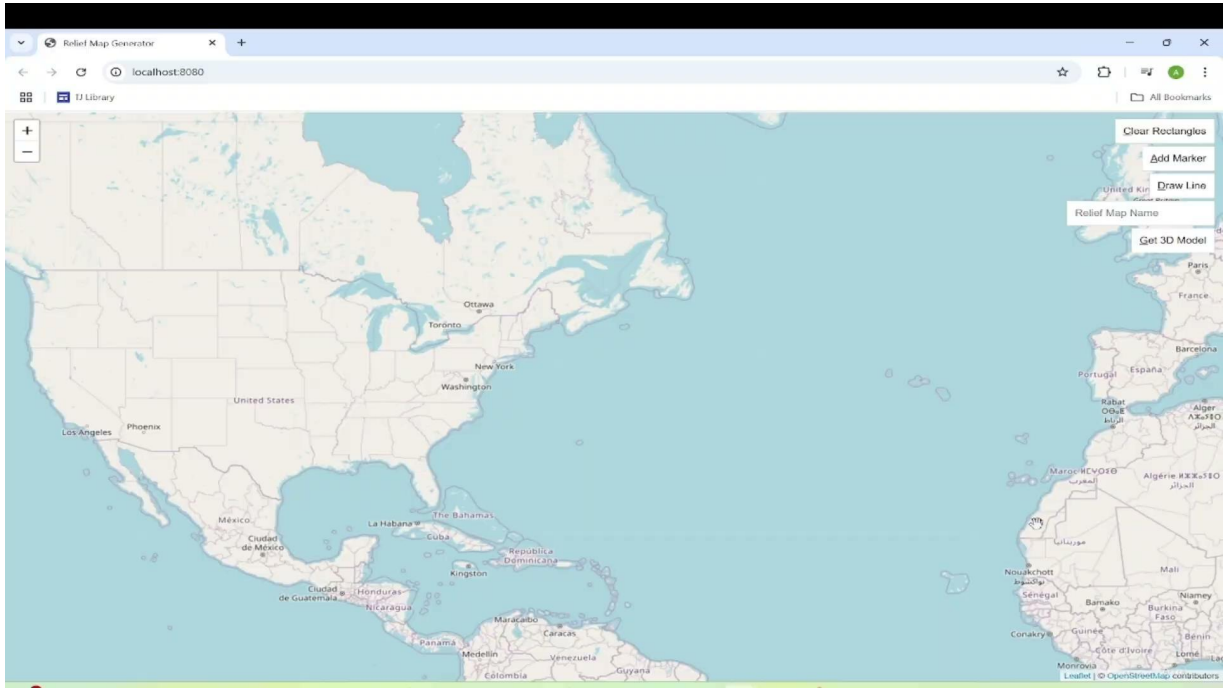


Methods - 3D Relief Map

- Convert Heightmap Data into 3D Model
  - Heightmap is an image where pixel intensity represents elevation
  - Read the map into Python using skimage
- Create Side Walls
  - Use the stl library to generate 3D terrain.
  - Iterate over the map to generate vertices and height values
- Create Contours on the Model
  - Use skimage to extract contour lines from the map
  - Convert the 2D contour points to 3D by assigning Z-values from the heightmap
- Add Markers and Trails
  - Define markers or trails based on specific locations
  - Create 3D representations of these markers and add them to the model
- Adding Name Plate
  - Skimage is used to perform boolean operations to add triangular prism and engrave name & date



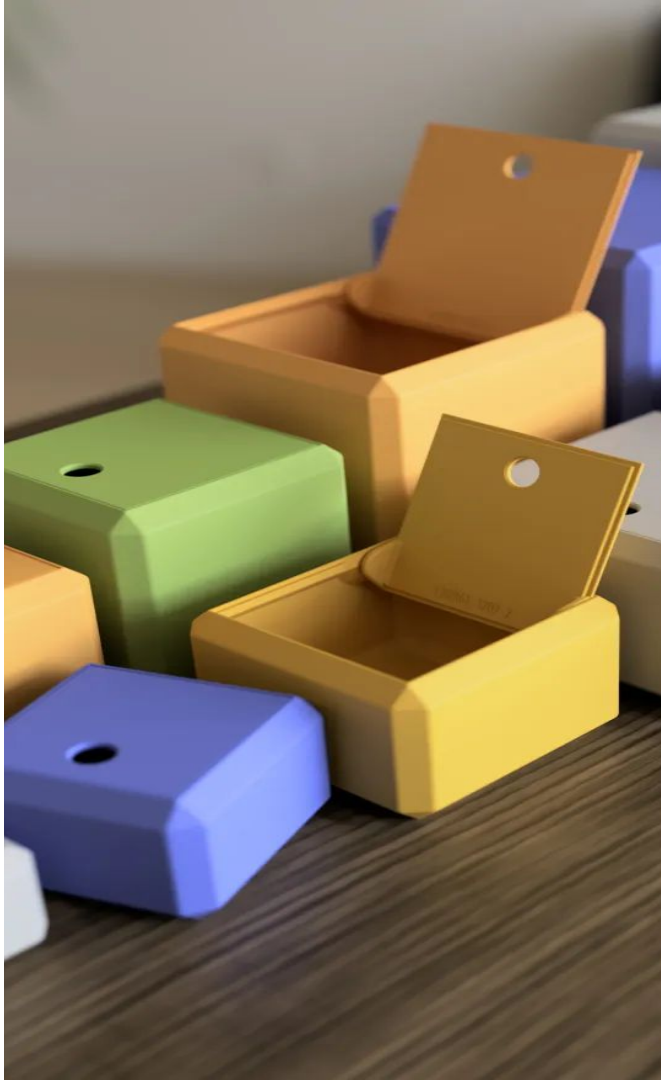
Topographic Map Modeling UI Demo



Project Focus Use Cases

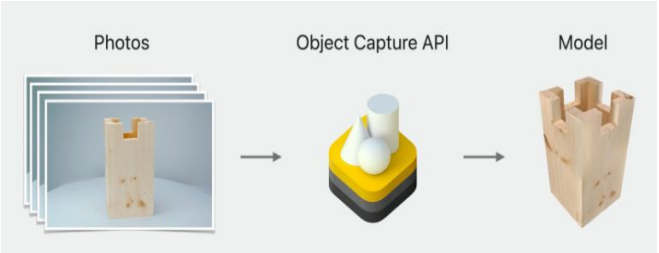
Chetan → Object Case/Package Creation

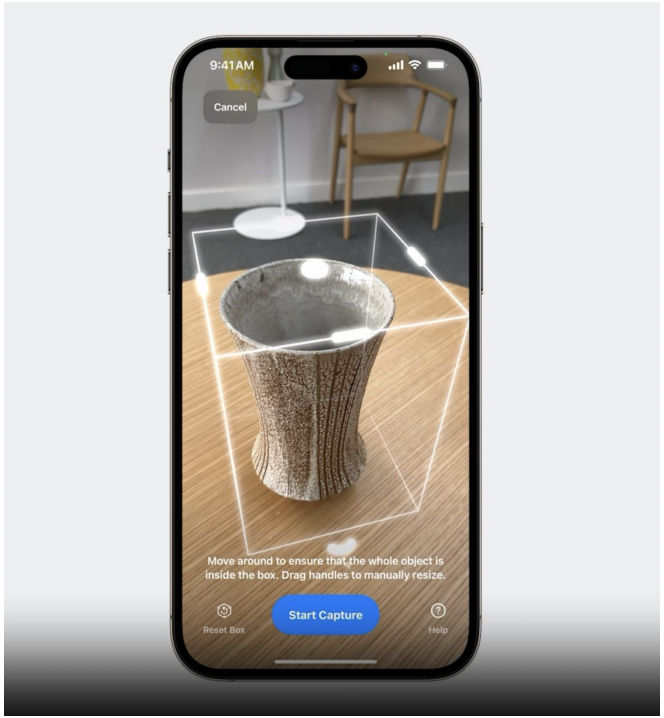
- Creating cases/containers for any object given just its images
- Cases completely covers the object and allow for safe storing and transportation
- Useful in the shipping and sales industry



Methods - Case Creation

1. Image Processing
    - Images of the object (from all angles, 360°) are passed into the model
2. Model Reconstruction
    - Utilize Apple's Object Capture API to create a 3D 1:1 model based on input images
3. Convert .usdz file type to .stl file type
4. Case Creation
    - Uses Blender Scripts to create a custom case for object given 3D model and user parameters
5. Output
    - The final 3D model is rendered and transformed into a 3D printable file type





3D Reconstructio

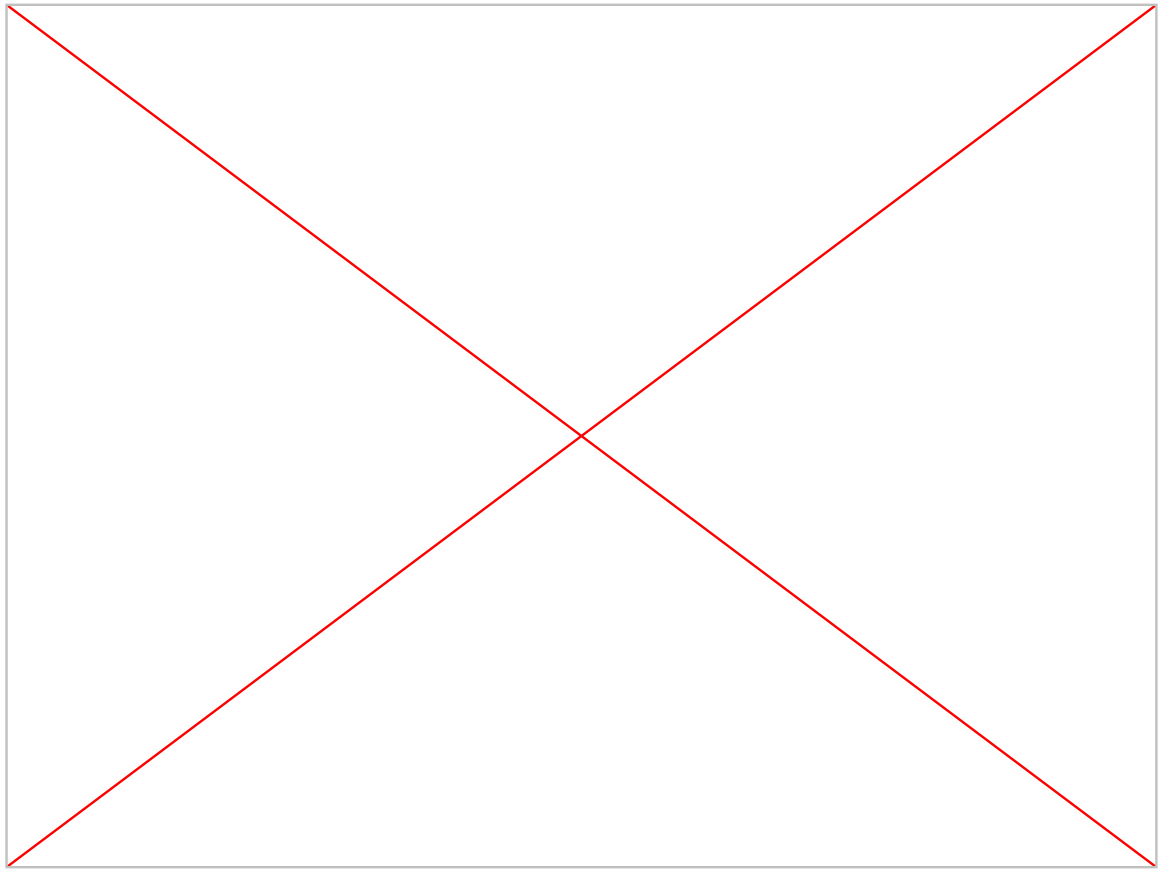
Upload Images for 3D Model

Choose Files No file chosen

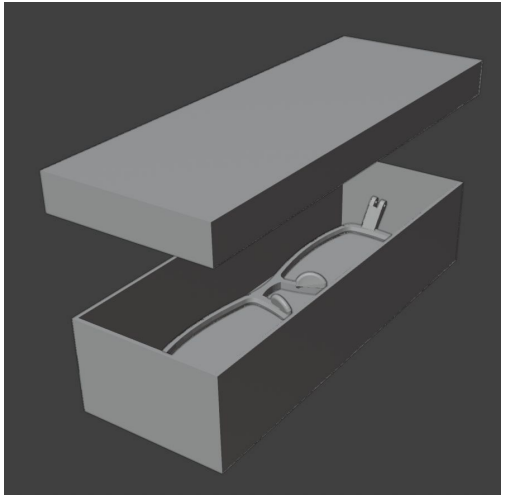
Generate 3D Model

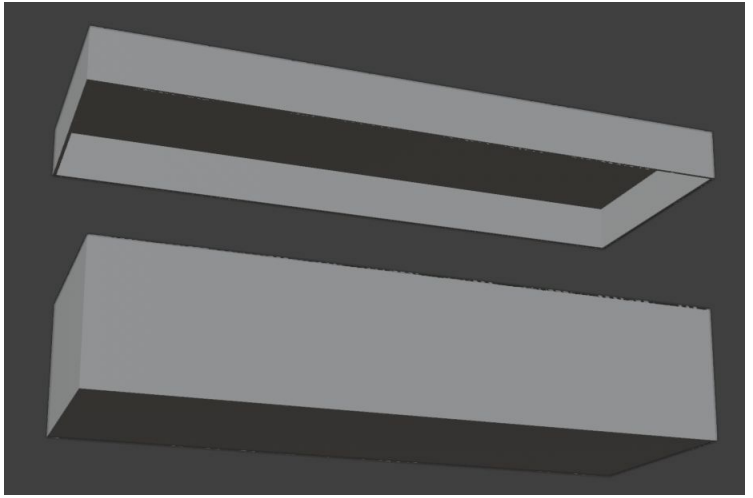
Status: Waiting for upload...

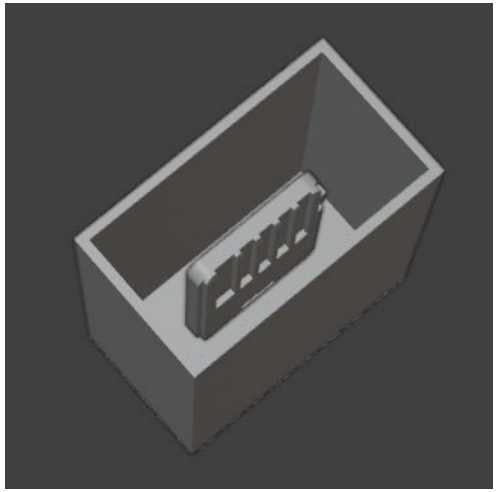
Case Creation

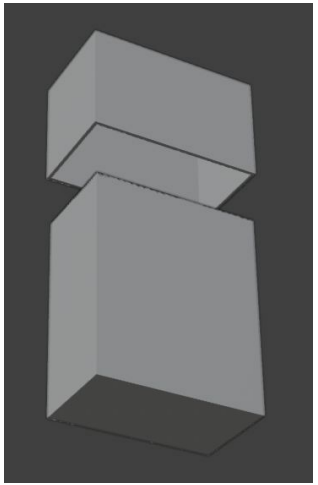


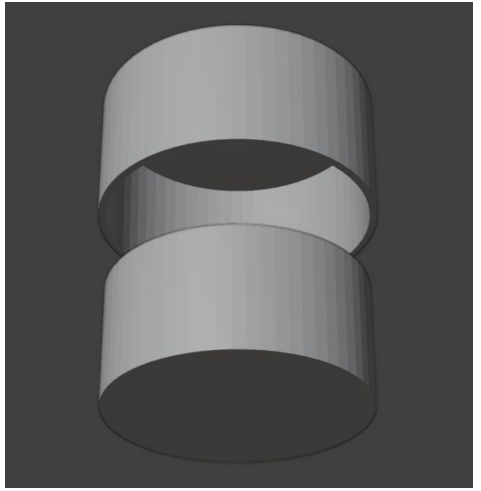
Blender Displays











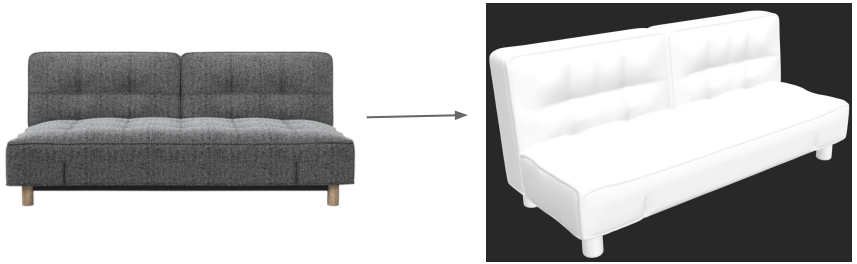
Project Focus Use Cases

Kanishk → Interior Design Modeling

- Ability to represent individual furniture items or interior spaces as 3D models
- Will allow for accurate recreations of room layouts and integrations of furniture models for easy visualization
- Useful for architects and interior designers

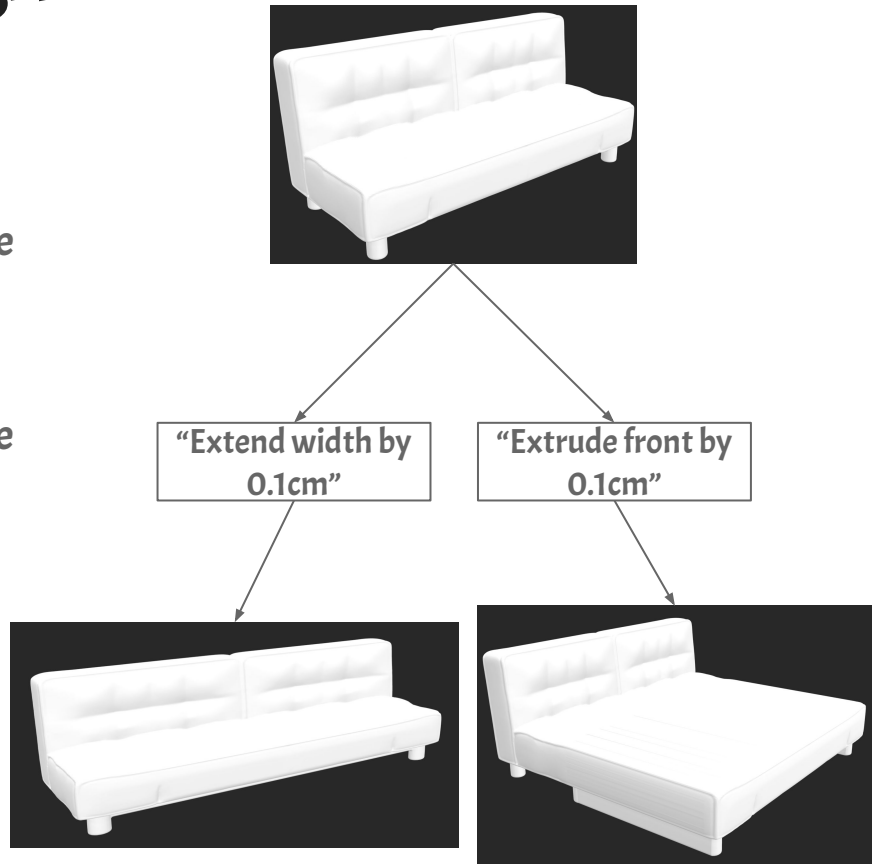
Methods - Interior Design

- 1. Image Processing
  - Images/Drawings of the object are inputted on the website and passed into the pipeline
- 2. Model Reconstruction
  - Background of image is removed
  - Utilize Hunyuan3D 2.0 API to create a 3D model based on input images
    - Hunyuan3D-DiT-v2-0 model handles shape (mesh) creation
    - Depth-aware encoder extracts input image's features like shape and contours
    - Diffusion-based 3D generation model (Denoising Diffusion Transformer)
- 3. Post Processing
  - FaceReducer – Reduces the number of faces in a mesh
  - FloaterRemover – Removes small, disconnected mesh components



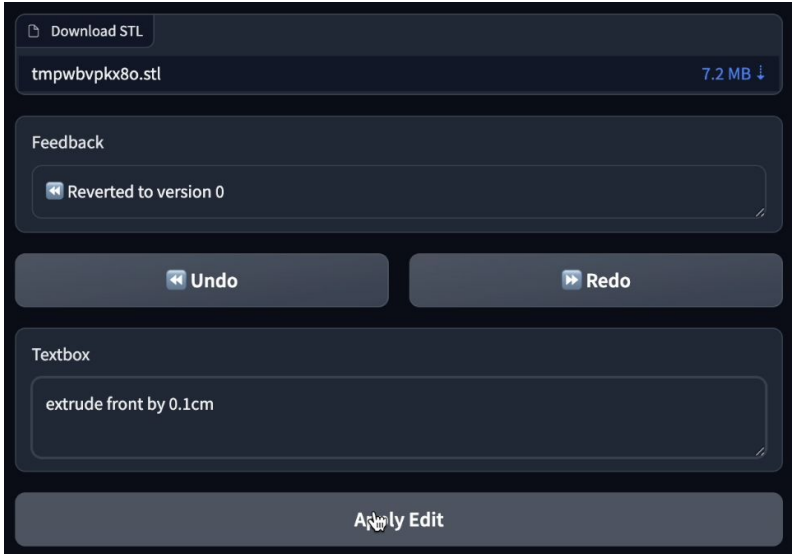
Methods - Interior Design

- 4. Editing
  - User text input is parsed and analyzed for keywords like “scale” and “extrude” using regex
  - 4 major functions: Scale, Rotate, Translate, Extrude
    - **A) Scale**
      - 3 different modes: Uniform, Absolute, Axis
      - Uniform scales all axes by a scale factor
      - Absolute rescales a dimension to a target size
      - Axis scales one axis (i.e X for width)
    - **B) Extrude**
      - Identifies target face by comparing vertex positions with bounding box
      - Find which vertices are closest to the boundary on the target side
      - Move target points outwards by inputted amount

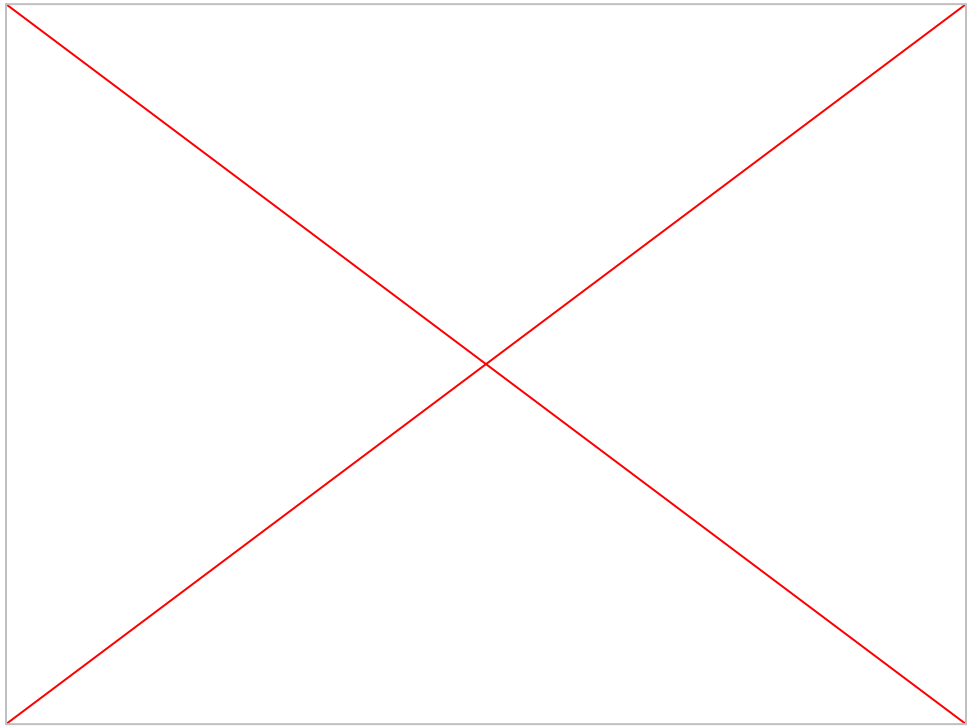


Methods - Interior Design

- 5. Version Control
  - Each edited/generated version of the model is saved in a history list to be indexed between at will
  - Allows for stepping between versions or undoing and redoing changes
- 6. Output
  - The final 3D model is rendered and transformed into a 3D printable file type (.stl)



Interior Design Demo



3D Printing Time-lapse

Limitations

- Image/Complexity Dependence
  - Perform optimally only with clear, well-lit images of relatively simple objects (simple geometric shapes). Objects with interlocked parts will not properly be modeled.
- Computationally Expensive
  - Rendering and editing of 3D models may be especially resource intensive
- (Text Command Limitations)
  - Explicit and clear user input necessary

Results

Heightmap to Terrain Models

- Converted heightmap data into 3D terrain models, effectively representing complex topographies like mountains, valleys, and contours
- Allows editing of 3D relief maps with trails and markers making them more useful in navigating the areas that they represent

Custom Case Creation Using Photogrammetry

- Photogrammetry techniques were used to generate 3D models of objects, where multiple images of objects (e.g., electronics, tools, accessories) were processed to create precise 3D model for packaging or storage design

Image to 3D Model Generation with Natural Language Editing:

- Users can generate models of objects given images or drawings, then modify the 3D models with text commands (e.g., resizing)

Conclusion and Future Work

We were able to create precise models of and make detailed edits to real-world objects in an efficient manner.

Next Steps:

1. Improve model accuracy, especially for fine details.
2. Expand the range of text commands to allow more complex edits.
3. Develop real-time interactive editing capabilities.
4. Explore more diverse use cases, such as home design and game development.
5. Look into automating slicing for STL files

References

Eskandar, S. (2023, April 26). *Exploring feature extraction techniques for Natural Language Processing*. Medium. https://medium.com/@eskandar.sahel/exploring-feature-extraction-techniques-for-natural-language-processing-46052ee6514

Inc., A. (n.d.). *Object capture - augmented reality*. Apple Developer. https://developer.apple.com/augmented-reality/object-capture/

Nvidia Omniverse. NVIDIA. (n.d.). https://www.nvidia.com/en-us/omniverse/

Products using USD. Products Using USD - Universal Scene Description 24.08 documentation. (n.d.). https://openusd.org/release/usd_products.html

Technology. 3DFY.ai. (2023, March 9). https://3dfy.ai/technology

Text-to-3d using 2D diffusion. DreamFusion. (n.d.). https://dreamfusion3d.github.io/

Vision Aerial. (2023, March 9). *Lidar vs. photogrammetry: Which is best for your project?* https://visionaerial.com/lidar-vs-photogrammetry-which-is-best-for-your-project/

Thanks!
Questions?