Initial Project Proposal

Year: 2022 Semester: Fall Project Name: Metaporter

Creation Date: March 25, 2022­ Last Modified: 05/09/2022

Team Members:

Member 1: Jehan Shah Email: shah435@purdue.edu

Member 2: Manav Bhasin Email: mbhasin@purdue.edu

Member 3: Kris Kunovski Email: kkunovsk@purdue.edu

Member 4: Sen Wang Email: wang3989@purdue.edu

1.0 Description of Problem:

3D reconstruction of the shape and appearance of objects from images is a problem with applications in many domains such as gaming, archeology, medicine, and more. Many video games exist where players can create their own characters to play with. However, it is often impossible to create a realistic or similar looking character to the player, making the game feel less immersive. We hope to address this problem by providing players with a low-cost tool to create a 3D reconstruction of their faces from a set of 2D images. While we plan to focus on reconstruction of faces, there are many emerging applications of 3D reconstruction such as large-scale reconstructions to create robotic simulations or creating 3D assets for sale as NFTs. By focusing on faces, we hope to learn more about the field, while solving a small sub-problem in the domain.

2.0 Proposed Solution:

Some of the most popular and successful solutions to the 3D reconstruction problem come from a group of methods called multi-view stereo (MVS). This group of algorithms relies on using stereo correspondence as their main cue. The main building blocks of a solution using MVS are as follows [14]:

1. Collect images

2. Compute camera parameters for each image

3. Reconstruct the 3D geometry of the scene from the set of images and corresponding camera parameters

4. Reconstruct the materials of the scene

The quality of the reconstruction produced by this depends on many factors. The image collection process is one of the most important parts of this process and factors such as the quality of the images, the overlap between the images, the background, and the lighting can all make a difference. Reconstruction from images alone can also be challenging if the target is not rich in texture or features. Therefore, we hope to create a more robust reconstruction tool by fusing sensor data from IMUs—to better infer the camera pose, and lidar—to get better depth data. To collect high quality data, we plan to create a circular track. Our target positions themselves in the middle of the ring, and our sensor package moves around the target and collects data from various angles. Then, we will fuse the sensor data, identify features, and use the MVS algorithm described above.

3.0 ECE477 Course Requirements Satisfaction

3.1 Expected Microcontroller Responsibilities

For the proposed project, a microcontroller will be used to interface with various sensor data such as motor, camera, IMU, and LiDAR/ToF in order to conduct 3D image reconstruction. The microcontroller is also responsible for transmitting the data to a different platform for processing, all while accepting basic inputs such as buttons and switches.

3.2 Expected Printed Circuit Responsibilities

The PCB is speculated to house the microcontroller, eeprom, power management, and battery monitoring systems. Depending on the chosen sensor, it will either be integrated into the circuit board or have a dedicated external interface mapped out on the PCB.

4.0 Market Analysis:

According to Mordor Intelligence, the market for 3D reconstruction is booming globally with a 2020 value of USD 816.6 million [1]. The projected value by 2026 is expected to reach USD 1292.22 million; therefore, the market for this technology continues to grow as innovation and demand expands. There is an expansive range of utilization for this type of market including video game development, graphic design, art restoration, and phenotype studies. Medical advances can also be made through CT-like scans using 3D reconstruction to look at bone, tissue, or tumor imagery. In addition, media and consumer products utilizing augmented reality are progressively being used as personal expressions of creativity and personality in modern times. Movies and video games also utilize 3D reconstruction technology to create realistic models for character development.

5.0 Competitive Analysis:

One such competition is a 3D ultrasound imaging patent [2] for reconstructing images of the head or torso. Very similar to our expected design, it focuses on the human body and establishes an accurate representation of the subject. Another patent [3] takes 3D reconstruction a step further by reassembling archaeological artifacts and rendering precise designs of objects. A face reconstruction scanner [9] focuses solely on providing appropriate replications of facial features and movements using sensor fusion and a 360-degree snapshot. Furthermore, another design [4] involves a linear laser and camera to provide accurate readings of any subject at every angle. In addition to laser scanning, there are other compositions of photogrammetry used to lay out these digital representations of different models.

5.1 Preliminary Patent Analysis:

Upon some patent research, there seem to be quite a few patents designed for similar applications. Nonetheless, the scope of these patents is specific and down to the implementations since it all falls into a specific classification G06T7.

Classifications:

G06T (General purpose image processing)

G06T7 (Image analysis)

G06T7/50 (Depth or shape recovery)

G06T7/55 (Depth or shape recovery from multiple images)

5.1.1 Patent #1:

**US Patent Application US201361758699P**

**Patent Title: "Real-time 3d reconstruction using efficient depth sensor"**

**Patent Expiration Date: 01/24/2034**

This patent [5] pertains to a mechanism that utilizes IMU for spatial recognition and a depth sensor to capture the image. Then, it down samples the resolution to reduce the computational complexity, thereby creating a partial mesh with minimal power consumption. This method is analogous to our approach by utilizing the IMU to obtain the spatial relationship, but it has a greater emphasis on reducing power consumption.

5.1.2 Patent #2:

US Patent Application US20170111628A1

**Patent Title: "Method and system for obtaining images for 3d reconstruction"**

**Patent Expiration Date: 09/02/2036**

This patent [2] describes one particular method to obtain a 3D image which is to use at least two light sources which are spatially separated from each other while outputting different brightness levels in a periodic manner. Coupled with at least three cameras which are also positioned spatially, this method produces a high-resolution, high depth-of-field 3D image unmatched by conventional methods. However, it also means that the equipment isn't easily portable and requires vigorous calibration in a lab environment.

5.1.3 Patent #3:

**US Patent Application US8928525B2**

**Patent Title: "Adaptive high speed/high resolution 3D image reconstruction method for any measurement distance"**

**Patent Expiration Date: 04/26/2033**

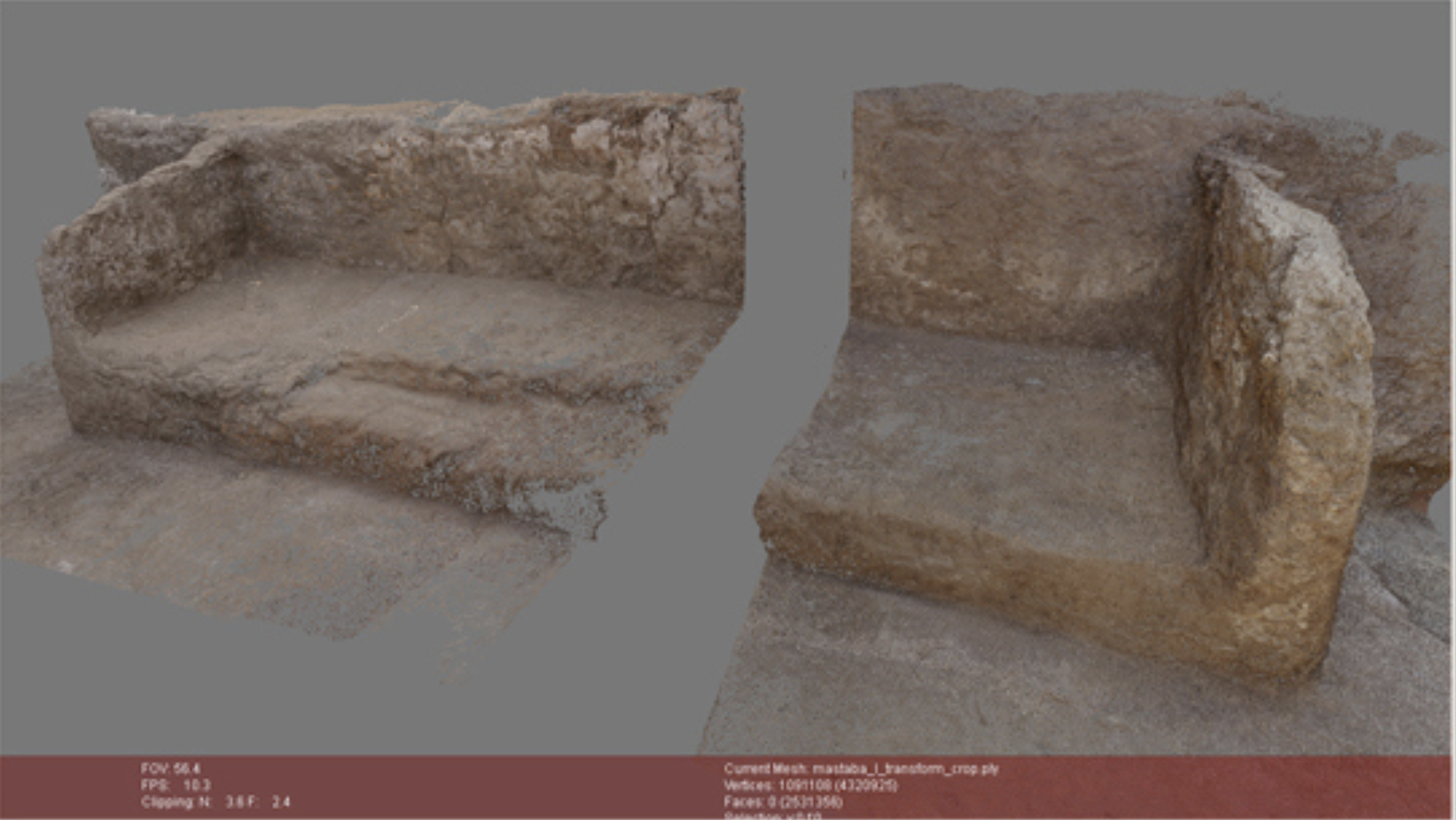
This patent [6] focuses on using predetermined weights and takes multiple images for 3D reconstruction. Then, it will use its in-house algorithms to conduct parallel processing in order to obtain a high-resolution image performed at a high speed. It is highly scalable since it has no limitation on how the physical hardware should behave. Nonetheless, this method requires prior calibration to adjust its algorithm parameters, which makes it unsuitable for different background environments.

5.2 Commercial Product Analysis:

After looking into some commercial products, we have learned that there are many devices in differing industries that can utilize 3D reconstruction techniques. The first two products directly use structure from motion, which is the algorithm we are planning to utilize. The last product uses a different kind of reconstruction from biometric data to provide a 3D rendering of one’s face for a video game.

5.2.1 Commercial Product #1: Agis Metashape

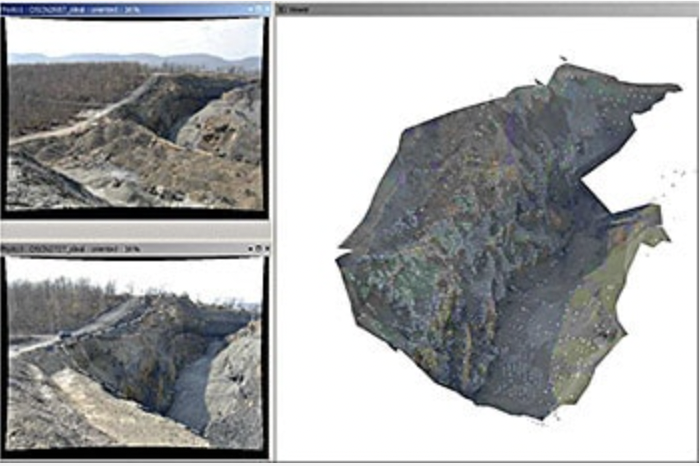
Agis Metashape [7] is a commercial software solution for structure from motion (SFM) primarily used for geological and archaeological reconstruction. Using Agis Metashape you can simply load the pictures as well as any associated GPS (camera pose) positioning, and the software will create a 3D rendering and point cloud. The image below is an example of the rendering capabilities of Agis Metashape. The main difference between the Agis software and our potential project is the absence of sensor fusion. We would like to supplement the images taken with data from sensors to create a more accurate reconstruction.



5.2.2 Commercial Product #2: PhotoModler Scanner

Another commercial product that utilizes structure from motion is called photoModler Scanner [8]. This product is similar to Agis Metashape; however, it has been optimized for UAV mapping. This allows researchers and users to mount cameras to UAVs or drones and then create an accurate 3D rendering of the landscape. As shown in the picture below, the product does a great job of highlighting the depth and intricacies of the terrain. Again, the major difference between this product and our own is the lack of sensor fusion and a target for reconstruction. We want to enhance the reconstruction by supplying data from sensors, and we would like to reconstruct a human body part.

.



5.2.3 Commercial Product #3: NBA 2k face scan app

To give players a much more customizable experience, NBA 2K (basketball video game) developed an app which allows players to scan the front of their face and import it into the game. With this face scan, players can build their own basketball players that look just like them. The NBA 2k face scan app collects unique biometric data which is then linked to a synthetic model of the head [9]. The scan is then overlaid on to the model of the head which allows for a 3D rendering of the user's face.

In our proposal we plan to take a 360-degree scan of the entire human head while the 2k face scan app only captures the front of the head. This would allow for a more robust and accurate portrayal of the user. However, our project would take up much more space compared to the iPhone app that NBA 2K uses.

5.3 Open Source Project Analysis:

5.3.1 Open Source Project #1:

The Multi-View Environment [10] is an implementation of a complete end-to-end pipeline for image-based geometry reconstruction, and it was developed at TU-Darmstadt. This project features SfM, MVS, and Surface Reconstruction. The individual steps of the pipeline are available as command line applications, but most features are also available from a user interface. Additionally, it is widely used in various fields and applications such as geoscience [11].

5.3.2 Open Source Project #2:

Regard3D [11] is an open-source structure from a motion program that supports 3D model recreation from static images of multiple perspectives. It can achieve elevated levels of granularity by going through rigorous phases of image reconstruction. It first detects any unique features from the image, such as edges and corners, and then models them into a mathematical descriptor. This will then be matched with descriptors containing similar features observed in a different angle. Using pairs of descriptors, the algorithm can generate a collection of matches between each image pair. With this information, it will then go through a triangular phase where the algorithm will obtain spatial recognition of the object by calculating the relational difference between each image. Lastly, it uses surface regeneration to render the obtained structure with textures and features. This OSS functions remarkably similar to our project with the difference being that our project is more tailored towards facial features and texture. This requires a higher-level of granularity compared with generic objects, and therefore, an additional layer of depth-map is required to obtain detailed facial features.

5.3.3 Open Source Project #3:

Meshroom [12], is a 3D reconstruction software based on the open-source photogrammetric CV framework called AliceVision [13]. In particular, it features SfM and MVS which are CV pipelines commonly found on similar projects such as project #1. The specific pipelines are as follows: Natural Feature Extraction, Image Matching, Features Matching, Structure from Motion (SfM), Multi-View Stereo (MVS), Depth Maps Estimation, Meshing, Texturing, and Localization. This project has an excellent GUI that is user-friendly, and it also includes a versatile API to interact with various Post-Processing, Retexture, and Retopology software such as Blender, ZBrush, and Maya (or Custom UVs). Overall, Meshroom offers professional experience while keeping the users' technological knowledge requirement at a minimum. Our project has a different focus compared to Meshroom, but some of the pipeline steps are synonymous with ours. We could potentially incorporate some of the pre-existing framework from AliceVision.

6.0 Sources Cited:

[1] “3D Reconstruction Market - Growth, Trends, COVID-19 Impact, and Forecasts (2022 - 2027),” Mordor Intelligence. https://www.mordorintelligence.com/industryreports/3d- reconstruction-market/ (accessed Apr. 9, 2022).

[2] H. Fang and Q. Yin, "Method and system for obtaining images for 3d reconstruction and system for 3d reconstruction," U.S. Patent 20170111628A1, Apr. 20, 2017.

[3] K. Shervais. *Structure from Motion Introductory Guide v3.0. (2015).* Accessed: Apr. 9, 2022. [Online]. Available: https://www.unavco.org/education/resources/modules-and- activities/field-geodesy/module-materials/sfm-intro-guide.pdf

[4] P. Bourke. *Photography based 3D modeling, reconstruction, laser scanning and photogrammetry*. paulbourke.net. <http://paulbourke.net/reconstruction/> (accessed Apr. 9,2022).

[5] ワグナー ダニエル. *Real-time 3d reconstruction using power-efficient depth sensors*. (in Japanese). J.P. Patent 6189483B2, Jun. 8, 2016.

[6] W.J. Byun, Y.H Cho, K.S Kim, M.S Kang and B.S Kim, "Adaptive high speed/high resolution 3D image reconstruction method for any measurement distance," U.S Patent 8928525B2, Jun. 20, 2012.

[7] *Agisoft metashape.* (2022). Agisoft. Accessed: Apr. 9,2022. [Online]. Available: <https://www.agisoft.com/features/professional-edition/>

[8] *PhotoModeler.* (2022). [Online]. Available: <https://www.photomodeler.com/>

[9] F. Robert-Inacio, F. Caudal and C. Rousset, "Biometrics for Human Face Reconstruction in 3D," in *Fortieth Asilomar Conf. on Signals, Systems and Computers*, 2006, pp. 608-612, doi: 10.1109/ACSSC.2006.354820

[10] MVE.(2022). [Online]. Available <https://www.cs.cmu.edu/~reconstruction/MVE.html>

[11] M.J.Westoby, J.Brasington, N.F. Glasser, M.J.Hambrey, J.M. Reynolds, " 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience application." in *Science* *Direct*, Vol 179, pp. 300-314, Aug. 2012. ISSN: 0169-555x. [Online]. Available: <https://doi.org/10.1016/j.geomorph.2012.08.021>.

[12] Regard3D.(2022). [Online]. Available: <https://www.regard3d.org/index.php>

[13] Meshroom.(2022). [Online]. Available: <https://alicevision.org/#meshroom>

[14] Y. Furukawa, C. Hernández. *Multi-View Stereo: A Tutorial (2015).* Accessed: Mar. 25, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/8637125>