Functional Specification

Year: \_2022\_ Semester: \_\_Fall\_\_\_ Team: \_05\_\_ Project:\_\_\_\_Metaporter\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Creation Date: \_\_\_August 30, 2022\_\_\_ Last Modified: September 03, 2022

**Member 1: Jehan Shah** **Email:** [**shah435@purdue.edu**](mailto:shah435@purdue.edu)

**Member 2: Manav Bhasin**  **Email:** [**mbhasin@purdue.edu**](mailto:mbhasin@purdue.edu)

**Member 3: Kris Kunovski** **Email:** [**kkunovsk@purdue.edu**](mailto:kkunovsk@purdue.edu)

**Member 4: Sen Wang** **Email:** [**wang3989@purdue.edu**](mailto:wang3989@purdue.edu)

Assignment Evaluation:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item** | **Score (0-5)** | **Weight** | **Points** | **Notes** |
| **Assignment-Specific Items** | | | | |
| **Functional Description** |  | x3 |  |  |
| **Theory of Operation** |  | x3 |  |  |
| **Expected Usage Case** |  | x3 |  |  |
| **Design Constraints** |  | x3 |  |  |
| **Writing-Specific Items** | | | | |
| **Spelling and Grammar** |  | x2 |  |  |
| **Formatting and Citations** |  | x1 |  |  |
| **Figures and Graphs** |  | x2 |  |  |
| **Technical Writing Style** |  | x3 |  |  |
| **Total Score** |  | | |  |

5: Excellent 4: Good 3: Acceptable 2: Poor 1: Very Poor 0: Not attempted

General Comments:

*Relevant overall comments about the paper will be included here*

1.0 Functional Description

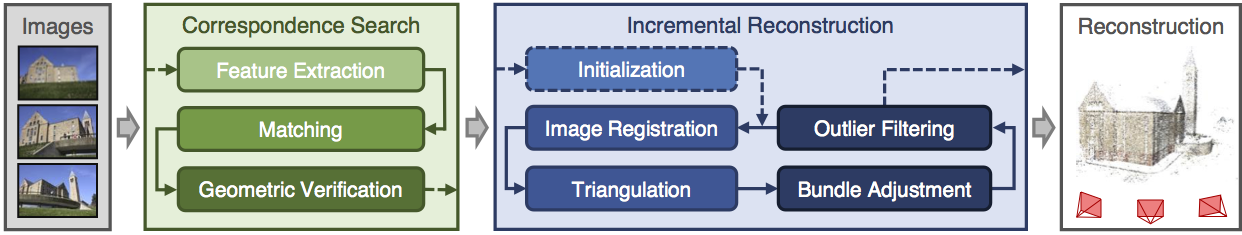
Metaporter is a handheld, plugged-in 3D scanning and reconstruction tool. With Metaporter, users can scan still objects and get a dense 3D mesh for it within minutes. Metaporter creates accurate 3D models by fusing multi-modal sensor data from sensors such as IMU, LiDAR, and Camera and estimating the volumetric radiance-and-density fields using neural networks. However, it has some limitations, and the quality of the reconstruction depends on the quality of data collection including the amount of data acquired. Metaporter may also have a limited ability to reconstruct certain types of surfaces and may not work under certain lighting conditions. Users will use a keypad matrix to select modes and interface with the device. A display on the device will show useful information such as the system state and the data recording time. The reconstruction will be viewed on the host machine.

2.0 Theory of Operation

There are multiple scientific and engineering principles that enable Metaporter to function. To break it down, we will begin by describing the underlying scientific principles behind each of our sensors: LiDAR, camera, and IMU. Then, we will talk about the sensor fusion algorithm used to combine sensor data and accurately estimate the camera pose. Finally, we will touch upon the neural network architecture responsible for performing the volumetric rendering and creating the 3D mesh.

A Light Detection and Ranging sensor (LiDAR) measures the distance from an object by emitting a short laser pulse and recording the amount of time it takes for the pulse to bounce off the object and return to the sensor . By using LiDAR, Metaporter can measure the distance of the target object and estimate its relative position in space.

A camera provides high-resolution RGB data which aids in inferring the color and texture of the target. Using a structure-from-motion pipeline, a monocular camera is used to infer the depth information and estimate the pose of the camera.



An IMU provides positional information in the form of roll, pitch, and yaw (or equivalent using direction cosine matrix, or quaternion). A 6 degrees of freedom (DoF) IMU commonly comprises of an accelerometer and a gyroscope, 9 DoF IMUs also contain magnetometers. IMUs measure the 3D rotation from one coordinate frame to another. The absolute orientation of Metaporter (i.e., body coordinate frame), relative to the standard coordinate frame (North, East, Down) can be estimated using just an accelerometer and a magnetometer. This data is rectified from calibration to remove the interference of hard iron sources—that generate their own magnetic fields—and soft iron sources—ferromagnetic material. The gyroscope measures the angular rate. By multiplying the angular rate with the sample time, we can calculate the change in angle during that time. By integrating (dead reckoning) over time, we can measure the current orientation relative to the starting orientation. Adding a gyroscope helps remove noise from corrupting linear accelerations. However, dead reckoning is prone to drift as it also integrates the noise over time. Therefore, data from all three sensors is fused together for a more accurate estimate using a sensor fusion algorithm such as Extended Kalman Filter (EKF).

Together using data from each of these sensors and a simultaneous localization and mapping algorithm, we estimate the camera extrinsics (camera position in space) and feed them along with the images to the neural network.

A neural emittance radiance field (NeRF) is a dense neural network that uses 2D images and their camera poses to reconstruct a volumetric radiance-and-density field that is visualized using ray marching.

Through this process, Metaporter constructs a dense 3D mesh of the target object.

3.0 Expected Usage Case

Metaporter is a handheld device with which users can scan still objects in order to create dense 3D meshes. Since Metaporter will be tethered to the wall for it to remain powered, it will best be used indoors. To guarantee a better reconstruction it is recommended that Metaporter be used in an environment with consistent lighting. Any age group above 13 should be able to use and operate the handheld device, however downloading and using the software required to view the reconstruction may require some technical proficiency. Users should also make sure that they have enough space around their target, so that they can collect data from various angles and heights.

4.0 Design Constraints

4.1 Computational Constraints

The primary computational functions of Metaporter include:

* Read IMU and LiDAR data
* Package sensor data and transmit it to the Jetson Nano via UART
* Receive and interpret user input from a keypad matrix
* Display system state to an LCD display via SPI
* Perform sensor fusion on multi-modal data
* Reconstruct a volumetric radiance-and-density field

With regards to the computational constraints on the microcontroller, it is mostly I/O bound as the microcontroller is responsible for transferring IMU and LiDAR data to the Jetson Nano via UART. Our microcontroller of choice is the STM32F091RCT6. It has 32KB of SRAM and 256KB of flash ROM. The LiDAR samples at a maximum rate of 1000Hz producing a 16bit distance reading for each sample. We estimate a total data sampling time of 1.5 minutes. This would produce 2000 Bytes/second and a total of 180KB data which exceeds our total memory capacity. However, not all data needs to be kept in memory as this will be offloaded to Jetson via UART. Using a baud high rate of 115200 we would have a data rate of 10472 bytes per second which should allow for sufficiently fast data transfer. The microcontroller also has two DMA channels that can be configured for the USART receiver and transmitter which we may choose to use instead if it proves more efficient. With 2 independent SPI channels (1 for display), and 2 independent I2C channels (for LiDAR and IMU), it has sufficient support for all our peripherals.

To interface with the camera, we require a CSI-2 interface which is available on the Jetson Nano. The Jetson Nano is also GPU enabled for efficient onboard image processing and with a wifi-card, it is able to wirelessly connect to our host machine.

To compute the final reconstruction, we are using a neural-network architecture called NeRF that is extremely computationally demanding, and we will offload it to a host machine wirelessly from Jetson Nano. The implementation that we are using requires a Nvidia GPU where more tensor cores increase performance. For our prototype, we expect to use a Nvidia RTX 2080Ti. It also requires a C++ 14 capable compiler, CUDA v10.2 or higher, and CMake v3.21 or higher.

4.2 Electronics Constraints

The major electrical components of Metaporter are a microcontroller, a camera, a LiDAR, an IMU, a keypad matrix, an LCD display and an Nvidia Jetson Nano. From an overarching perspective, the project consists of two components, a sensor hub and the Jetson Nano.

The sensory hub will be a microcontroller that directly interfaces with the sensors (IMU and LiDAR) via I2C (with pull-up resistors) for ease of implementation, as well as reserving valuable SPI and UART channels for more I/O intensive tasks. The only IMUs that are readily available is 6-axis. It doesn't include the magnetometer that is vital for providing a frame of reference during 3D reconstruction. Therefore, depending on the future part availability, the team may have to decide between forgoing with IMU or obtaining a separate magnetometer which would increase the complexity of the system. On the other hand, the LiDAR sensor will provide precision depth information in high frequencies with the expense of power draw.

In addition to the sensors, the microcontroller will also be responsible for receiving user input from the keypad matrix and displaying the system state to an LCD display via SPI.

Lastly, the microcontroller will also have a UART interface that connects to an Nvidia Jetson Nano. Note that a memory buffer is required for UART communications which is built-in to the microcontroller of our choice (stm32f091). This UART interface will transmit all sensory data collected by the microcontroller.

4.3 Thermal/Power Constraints

Given the power-demanding nature of pre-processing on an Nvidia Jetson Nano, the device will need to be plugged-in. Several power rails will have to be designed on the PCB as each component takes in different operating voltages, ranging from 1.8V,3.3V, and 5V. The project power dissipation is anticipated to be around 15W TDP.

Heat dissipation will also be an important design consideration for our team, as both LiDAR and Jetson Nano are power hungry with Nano requiring active cooling. According to IEC regulation, IEC 62368-1 regulates the touch temperature limits on non-metal surfaces to be 70ᵒC (158 ᵒF) [7]. Nonetheless, our target maximum package temperature is aimed to be 35ᵒC (95 ᵒF), as higher temperatures will be unpleasant for consumers to hold for a long duration of time.

4.4 Mechanical Constraints

Metaporter consists of a handheld package meant to be moved around a target to capture data from various sensors. This means that Metaporter should be light enough so that an average human would be able to move it around freely. All in all, we estimate that metaporter will be no more than two to three pounds. Since metaporter is a handheld device, this means that we also must have packaging around the device to protect users from exposed wiring and electrical components. The package surrounding the electrical components will need to be vented to accommodate any cooling needs. Although we are designing the device to be packaged, we cannot guarantee that it will be weatherproof or environment proof, so it is important to use Metaporter indoors.

4.5 Economic Constraints

While there are no commercial products distinctly identical to Metaporter, open-source projects like Multi-View Environment [1], Regard3D [2], and Meshroom [3] similarly utilize structure from motion and multi-view stereo to render 3D models. Considering this, other competitors will undeniably put economic pressure on Metaporter’s business; therefore, it is important to set a reasonable price for our product. Working with the constraints of our costliest components, namely the LiDAR and Jetson nano, we set our initial product cost to $650 USD. In addition to the Metaporter apparatus, there is another external device that customers will have to purchase to fully utilize our product, and this would be an NVIDIA GPU-enabled system to run the 3D modeling software and UI.

4.6 Other Constraints

In order to accurately recreate a person’s face, the camera should be able to record clear images of the subject with no harsh or extraneous shadowing. Therefore, lighting is another constraint for Metaporter since image processing plays a vital role in an accurate 3D reconstruction. To guarantee a more precise 3D model, the subject should be well-lit with minimal interfering light sources that may botch the rendering. Metaporter is designed to create 3D reconstructions of only still targets.

5.0 Sources Cited:

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

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

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

[4] Velodyne Lidar. (2022). [Online]. Available: <https://velodynelidar.com/what-is-lidar/>

[5] Y. Li *κ.ά.*, ‘DeepFusion: Lidar-Camera Deep Fusion for Multi-Modal 3D Object Detection’, στο *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, σσ. 17182–17191.

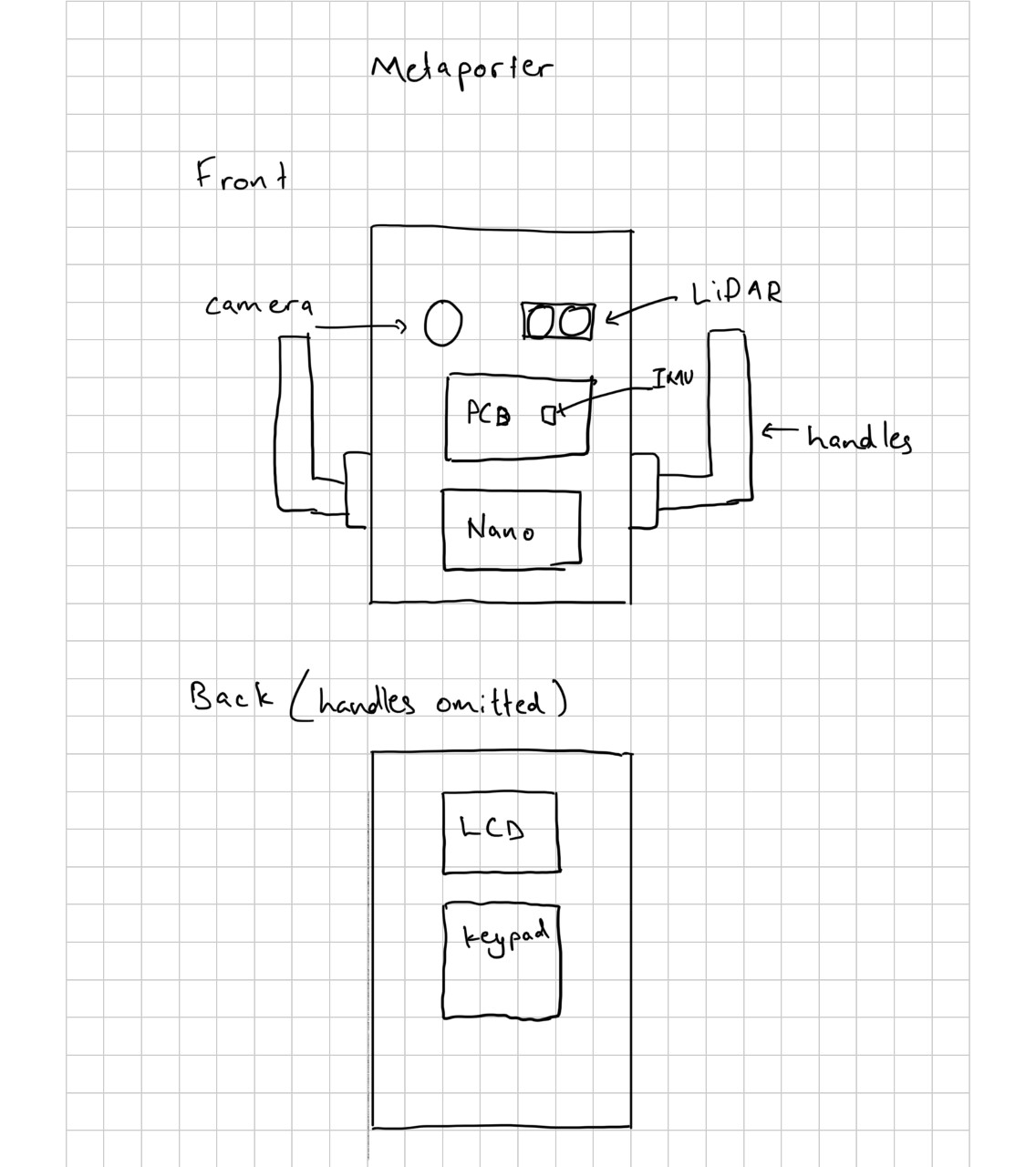
[6] T. Müller, A. Evans, C. Schied, and A. Keller, “Instant neural graphics primitives with a multiresolution hash encoding,” *ACM Transactions on Graphics*, vol. 41, no. 4, pp. 1–15, 2022.

[7] IEC 62368-1:2018 RLV Redline version. [Online]. Available <https://webstore.iec.ch/publication/63964>

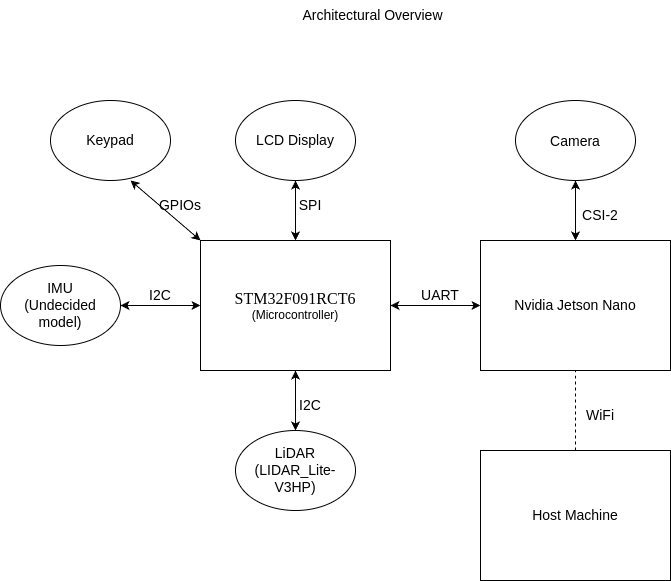
[8] MATLAB, “Understanding sensor fusion and tracking, part 2: Fusing a mag, accel, & gyro estimate,” *YouTube*, 22-Oct-2019. [Online]. Available: <https://www.youtube.com/watch?v=0rlvvYgmTvI&list=PLn8PRpmsu08ryYoBpEKzoMOveSTyS-h4a&index=4>. [Accessed: 03-Sep-2022].

[9] “Tutorial,” *Tutorial - COLMAP 3.8 documentation*. [Online]. Available: <https://colmap.github.io/tutorial.html>#. [Accessed: 03-Sep-2022].

Appendix 1: Sketch of Project Prototype



Appendix 2: Architectural Overview



Appendix 3: Functional Block Diagram

