EC601 Placement Exam Paper

Zachary Halvorson

August 29, 2020

3D imaging and reconstruction has a long-standing history, with applications in topography, medicine, static object reconstruction, gesture tracking, and many other fields. The specific user story of interest for this paper details the use of a “drone capturing images of an archeological site, construction site or a 3D real object”, and the subsequent challenges in providing “an interactive virtual experience where the user can do measurements or analyze details of a small area”. Throughout this paper I’ll first describe my limited prior experience with 3D reconstructions, followed by a thorough literature and industry analysis, including applications and demonstrations where open source code and data is available, concluding with a proposal for future development in the field.

**Personal Experience and Prior Knowledge**

Related to my personal interest in the specialty coffee industry, I set out in June of 2020 to attempt to [reconstruct the flow of espresso](https://zachhalvorson.com/3d-espresso-reconstruction/) during extraction, which can be modeled in the shape of a hyperbolic cone pointing downwards from the bottom of the espresso portafilter screen. After quickly researching the topic and receiving recommendations from a colleague in the GIS and ecology field, I developed a plan to use a set of seven mirrors and a single camera to capture six different angles simultaneously of the espresso liquid flowing out of the machine into the cup. Next, I would capture frames for a set of timepoints in the video, and auto-crop the six different views form the mirrors out of each captured frame. These extracted images would then be imported into either Meshroom or Metashape for reconstruction. I quickly realized after recording the videos that I had nowhere near the resolution required to perform an accurate reconstruction, as well as recognizing that my light source would likely have introduced many errors as well due to shadows and bright glares and reflections. This project was put on further hold but served to be a great demonstration of the challenges with 3D reconstruction, namely resolution, lighting, and overlapping views for static objects.

As part of this research and planning I reviewed some academic papers to help determine my mirror arrangement and placing, mostly focusing on the use of mirrors with single camera reconstructions[[1]](#footnote-1),[[2]](#footnote-2),[[3]](#footnote-3),[[4]](#footnote-4), but also including time-of-flight sensors[[5]](#footnote-5), as well as an interesting application of a moving mirror[[6]](#footnote-6). My motivation for this research was to determine the required mirror arrangement for my experiment, so after briefly reviewing these papers and other online sources, I decided to collect at least 5 overlapping angles of the hyperbolic cone of the espresso fluid. Interestingly, I found a paper in which the researchers designed a single camera, five mirror system to image and reconstruct cavitation bubbles in a vertical column of water[[7]](#footnote-7), which is a very similar application to reconstructing and tracking the flow of liquid espresso, which contains small bubbles and color variations throughout the fluid.

Additionally, I attempted to perform a 3D reconstruction of my espresso maker with photos taken from a Sony RX100M5, and while I had some success, I learned of the difficulty in reconstructing reflective objects, as well as the importance of imaging with a non-textured background, as this introduced significant error in my reconstruction. I also experienced first-hand the extensive time it takes for a moderately powered laptop to perform 3D reconstructions from large sets of images.

**Drone-Produced Data and 3D Reconstructions**

More directly related to the drone user story, I conducted a literature and industry review to understand the history, current status, and future challenges of this area. I began by reviewing the relevant terminology of different methods, techniques, and technologies that are used in the field.

While “3D reconstruction” is a broad term, there are many approaches to this technique, which can be described as either active or passive. Passive methods include techniques like photogrammetry, in which multiple overlapping photos are taken of the same target object, to encompass all 360° surrounding the object. The images are then processed, where common features from each image corresponding to the overlapping sections and matched, extracted, and used to produce a geometric mapping of the target object that can be combined with the captured images to create an interactive 3D model. There are many variations to this technique, as this can be achieved with multiple cameras capturing at the same time, a single camera capturing successive images while moving around the target, or a combination of a single or multiple cameras with mirrors, to create additional simulated perspectives. Additionally, these software programs can operate with input images only, however they also are able to incorporate positioning data of each image or camera, often through GPS, to augment the 3D reconstruction and mapping process. Passive methods also include monocular and binocular methods, which rely on the effect of shadows, lighting changes, as well as a fixed position for the camera or multiple cameras in use. Active methods, however, are based on emitting or projecting a known signal, which can include infrared light, ultrasound or radio waves, and/or colored light, and measuring the subsequent reflection of that signal, both in the time to reflect as well as the intensity.

While both active and passive methods can be effective provided clear and accurate input data, a brief search for industrial products resulted in most static 3D object scanners relying on active methods, and the inverse for drone-based mapping. While some drone systems use LIDAR and precise GPS systems to map geographic areas, these systems are often far more expensive than camera-based models. However, it is worth noting that LIDAR models can create accurate maps during the day as well as during the night, as the active method does not need an illuminated scene, whereas the passive camera-based systems must be operated during the day, and are thus also impacted by changes in cloud cover, sunlight intensity, or other interference for large captures that take long periods of time.

The static 3D object scanner market is very well developed, with products ranging from open-source 3D printed projects[[8]](#footnote-8), to kits which only require assembly[[9]](#footnote-9), to accessories that can be attached to smartphones and other mobile devices[[10]](#footnote-10), and ultimately to industrial and professional equipment[[11]](#footnote-11). Some of these devices are handheld and portable, with other models designed for desktop use as fixed position scanners in which the target object rotates on a turntable with a fixed sensor positioning.

Focusing instead on the drone-based systems and software, one immediately becomes aware of OpenDroneMap, ODM, an open source program for aerial based photogrammetry. ODM was created in 2014 and has had an immense impact on the field. Incredibly, while researching ODM, I was stunned to learn that the colleague that I referenced earlier for my espresso reconstruction project, is actually the co-founder of ODM, Stephen Mather. I was stunned to say the least.

Aerial photogrammetry is an effective tool for generating 3D models, interactive maps and building models[[12]](#footnote-12),[[13]](#footnote-13),[[14]](#footnote-14), as well as extracting relevant data for many applications, from farm management and agriculture, to utility line management[[15]](#footnote-15), forestry tools[[16]](#footnote-16), among others. There are many options for maneuvering the drone, which can be described as either automatic, semi-automatic, or fully manual modes. Software programs like Litchi allow for fully autonomous flight planning, as well as providing many additional settings for object tracking, and image capture modes with partial control to produce smooth panning shots.

Returning to the original user story of a drone surveying an archaeological site, and producing 3D reconstructions from the acquired data from the drone (images, time-of-flight data, GPS measurements, ground control points, among others), we are presented with many challenges. For raw image input data, challenges include consistent lighting, reflectivity and/or glare from the imaged objects, trees and other dynamic objects that move slightly throughout multiple captures, the need to capture photos during the day, preferably with some cloud cover so as to diffuse the sunlight and prevent sharp shadows, as well as prevent locals and tourists from visiting the area during the capture period. Many of these challenges restrict the available time periods in which optimal conditions exist, as well as make the facilitation of capturing the area more expensive as the project may take longer and interrupt normal operations in the area for a longer period of time.

However, it is also possible to combine aerial photography with ground-based photography, as well as LIDAR, or other depth-based measurements to effectively reproduce the scene[[17]](#footnote-17). Combining multiple types of data is an effective way to ensure redundancy as well as increase the accuracy and visual appearance of the final reconstruction. Further, to provide dimensionally accurate 3D reconstructions[[18]](#footnote-18),[[19]](#footnote-19), researchers can incorporate additional sensors and physical calibration objects in order to ensure that the reconstructions are accurate and able to be utilized for continued research and experimentation. By capturing within the environmental images, a reference object of a known size, the dimensions of other features can be extracted given the specifications of the camera sensor in use, namely the pixel size and resolution of the camera, as well as the effective magnification additional settings. Depending on the scale of the measurements of interest, GPS points acquired from the drone while capturing each image could be sufficient. For power lines or other large objects where a relevant metric could be the amount that the power line sags between to towers, GPS combined with the output of a sensitive altimeter, could effectively provide utility line engineers with valuable data regarding the tension on the cable, and how this may change over time due to weather conditions (ice on the line, expansion and contraction of the cable due to temperature). Utilizing a drone to collect these measurements could likely prove to be safer and more efficient than requiring personnel to manually inspect the lines. In general, the smaller and more accurate the desired measurement, the more difficult in ensuring dimensional accuracy in a 3D reconstruction. It should be noted that although some GPS units are very accurate, variation due to satellite positioning could introduce significant error in the collected data.

Companies like CyArk[[20]](#footnote-20), among others[[21]](#footnote-21), have been working towards the goal of digitally documenting and producing interactive 3D reconstructions of world heritage sites, with incredible results. In a similar context, some museums, specifically the Cleveland Museum of Art[[22]](#footnote-22), provide interactive 3D scans of their physical artifacts, with other projects aiming to reproduce entire museums virtually[[23]](#footnote-23). CyArk describes for their project MasterWorks[[24]](#footnote-24), that they incorporated LIDAR, ground-based photogrammetry (some of which being dual camera capture units for stereo capture), and drone or aerial photogrammetry to reconstruct their target sites. They partnered with FarBridge to release these reconstructions for the Oculus Rift, Samsung Gear VR, and later the HTC Vive and other devices, allowing the wide public and many educational institutions to access and interact with these virtual reconstructions.

Notably, one of the main limitations, and the point on which I will propose a path forward, is the difficulty or inability to capture data and reconstruct archaeological sites during non-ideal weather conditions. These conditions could include rain, snow, hail, strong winds, or dust storms. Weather and seasonal conditions are an integral part of the human experience, and being able to provide and simulate 3D reconstructions of sites during all seasons would provide a new level of immersion for all users. Being able to interact with a virtual site only during a clear and sunny day does not provide a true understanding of the diverse conditions and environments in the area. Mesa Verde National Park[[25]](#footnote-25) in Colorado, which has been virtually reconstructed in MasterWorks, serves as a valid example. The Mesa Verde region experiences[[26]](#footnote-26) an average snowfall of 15 inches per month in December, January, and February, with corresponding maximum and minimum temperatures during the winter between 18°F and 40°F. Juxtaposing these conditions with the currently available reconstruction, likely during a summer month, of approximately 75°F air temperatures, with less than an inch of rain over an entire month.

To achieve this goal of providing seasonal and more diverse weather reconstructions[[27]](#footnote-27) that users can interact with, a few conditions must be achieved. First, there must be access to, or collection of, accurate weather data over the course of the year, as well as data or simulations of more severe weather events. If a region’s history has been greatly impacted by the effect of seasonal flooding and storms, it would be incredibly valuable to be able to simulate those experiences in order to preserve additional aspects of our cultural heritage. Second, there must be the capability, both technologically through the reconstruction itself, and externally through the modification of the user’s environment, to simulate these alternative weather conditions. The context of the following examples would be an individual in a museum, interacting with a 3D reconstruction through a VR headset system. For instance, a late-fall morning in the Pacific Northwest would contain a set of unique experiences, which could be simulated through the use of mist systems, fans, AC units set to cooler temperatures, and perhaps even specific scents (pine trees), to incorporate additional senses into the experience. The individual could then step into a sequential room in the museum, simulating the winter season, where the room temperature is much colder, perhaps with some simulation of light snow, and a different set of background noises and scents. The individual could then step into another room simulating a summer month, with much warmer temperatures and humidity levels, among other changes.

Within the digital reconstruction, these seasonal changes could be achieved through the acquisition of new images taken during winter months, that are then applied as a texture onto the previously generated mesh, or depth map, depending on the available data. However, it could be possible, and likely more cost effective, to utilize video game engines to edit and adjust the base model taken during ideal weather, and apply animations and effects to that scene to simulate different conditions, without the need to capture more data from the physical scene itself. If a group were inclined to collect new data, they could include the use of thermal cameras[[28]](#footnote-28) to provide additional layers of data on top of the previously generated models. Collecting all of the data from scratch in a new season is entirely unnecessary, and due to the challenges with capturing images during rainfall or snowfall, the collected data may not be usable. Further, in high wind, snow, and rain conditions, many commercial drones are unable to be operated, save for some specifically designed drones[[29]](#footnote-29). Further, LIDAR would result in extremely noisy data during rainy conditions, or dusty conditions, as the particles in the air would interact with (scattering or reflection) the laser light. In summary, augmenting ideal weather condition reconstructions with seasonal and storm variations, would provide those that interact with the models a far more immersive and realistic experience of the region, enhancing the educational and experiential benefit of documenting archeological and historical heritage sites.

**Demonstrations**

To experiment with open source aerial photogrammetry of heritage sites, I downloaded and installed WebODM[[30]](#footnote-30), and accessed data from OpenHeritage[[31]](#footnote-31), and additionally tested some static objection reconstruction with my own personal camera. Results and discussion are available within the repository here - <https://github.com/halveez/3Dscene2Dimage>.

**GitHub Supplementation**

Due to the majority of my work experience existing in private repositories, I wanted to provide some higher level overviews on three specific projects to supplement my resume.

First, I’ll describe my work during the summer of 2019 at the Salk Institute in the Busch Lab, where I worked with Matthieu Platre (postdoctoral fellow) to develop a root growth rate tracking algorithm, to improve his process in analyzing his experiments. Previously, Matt would use a macro that he designed in ImageJ to process a series of time-lapse images, which required him to manually crop images, select input and output folders, and generally supervise the process for the entire duration (multiple hours). Ultimately, I was able to develop a MATLAB program that is to this day being used by four researches in the lab, as well as separately at UCSD, and has saved hundreds of hours of time for the researchers, and allowed them to focus on other tasks. The algorithm was named RootWalker, as the fundamental principle of the program was to determine the starting position of the roots at the top of the image and sequentially “walk” along each one until finding the tip of the root, repeating this process for each root in the image, for each image in the panorama, and for each panorama in the time series. Fortunately, the raw image data is very clean, and the only major artifacts are the presence of bubbles which only exist in the initial images due to having transferred the roots from a larger growth plate to a smaller plate for imaging. The general process of the program was to determine the structure of the input images, extract and group them per timestep and sample slide, followed by a thresholding, denoising, and smoothing process to remove jagged edges that would otherwise overestimate the root length. The images were converted from their matrix form to a representative dataset, where each column corresponded to a single root, and the value at that position was the x-coordinate of the center of that root. This smaller dataset would then be processed to determine the length of each root, detecting overlaps and jumps, and automatically filtering and notifying the researcher of any roots that were disqualified from the final output. See “root\_walker\_video.mp4” for a demonstration.

Second, I’ll describe my work on a pressure gauge computer vision algorithm that I first worked on in the summer of 2019, before continuing for a class at BU in the fall of 2019. As an espresso enthusiast myself, the specific pressure that I apply while making espresso has a major impact on the taste and quality of the espresso being made. Since my specific espresso machine has an analog gauge, in order to track and record the pressure values for each espresso I make, I would have to purchase a $600 digital pressure transducer, which is approximately twice the cost of my espresso. However, I had seen earlier in the spring of 2019 some research papers indicating that it was possible to extract the pressure values using a computer vision approach. There were generally two approaches to this, one being a general feature detection and geometric approach, and the other being a ML approach, in which a dataset of pressure gauges is generated with various pressures, tagged with a corresponding value for each image, and then finding the closest match in that database. At the time, I had no experience in AI or ML, so I chose to develop an MVP in R, using some computer vision packages and libraries. I was able to successfully extract pressure values from a pre-recorded video only, as I did not attempt to run R code on a mobile device or other system to operate in real time. This was achieved in the fall of 2019, and I continued development into the spring of 2020 on an R-based web app, that I eventually scrapped. I then during this summer converted the R code into C++ using OpenCV (which the R libraries were simply wrapped around), and built a single page iOS app to run the computer vision code directly on an iPhone. As I continue to utilize a geometric approach, the general process of the algorithm is quite straightforward. First, the gauge itself must be detected on the screen, followed by a refinement and detection of the center point of the gauge. Next, the direction of the needle must be extracted, and the angle between the needle and some reference point or line must be calculated, and then converted into a pressure value from a set of known values. This project has been put on hold, in favor of working on the final project I’ll describe.

Third, I’ll describe a coffee roast level sensor that I’ve also designed and developed through my company. This device, among approximately 10-15 others in the market, serves to quantitatively determine the roast level of coffee beans. Typically used for quality control purposes, in which a coffee roaster aims to roast the same coffee beans consistently day in and day out so that their customers can expect a consistent product. Additionally, when comparing the quality of green (unroasted) coffee beans, it is important to keep the roast level consistent, so as to not introduce variation in taste as a result of the roasting process. In order to do this, the Specialty Coffee Association has released guidelines such that coffee roasted for evaluation and grading purposes must be roasted to a certain degree, plus or minus a small variation, on these roast level devices. The involvement of these devices in the grading and QC system has large downstream impacts on farmers, that take prices based on the quality of their crop, as assessed during these tastings. Further, these devices are typically $1500 or more, with some in the +$10,000 range. The cheapest current device costs $800, still out of reach for many small-scale roasters in the US, let alone small farms and co-ops in other countries. I focused on designing the most affordable device so that more individuals and groups can access these tools and empower themselves in their marketing and pricing options. Developing this device began in the summer of 2019, when I experimented[[32]](#footnote-32) with some sensors taped into paper cups, and found surprisingly good results. I continued working on this project over the last school year, intending to compete in the Imagineering Competition, which was cancelled due to the pandemic. While home for the remote semester, I purchased a 3D printer and began a long process of designing, testing, validating, and eventually filing a provisional patent for, and preparing for sales of my roast level sensor. I built automated testing systems that would measure 20 different samples of coffee with 6 sensors mounted to a rotating arm on a servo, which I could place in a fridge or oven to investigate the effect of temperature, as well as increase the speed and efficiency of my measurements. I compared multiple sensor to sample arrangements, multiple sensors and light sources, as well as ways to calibrate the sensors to a known standard, and focused on minimizing the size of the device in order to make it portable, as well as faster to 3D print the housing.

Please feel free to ask about these projects and I can provide more information.

1. Lu, Thomas & Chao, tien-hsin. (2006). A single-camera system captures high-resolution 3D images in one shot. Spie Newsroom. 10.1117/2.1200611.0303. [↑](#footnote-ref-1)
2. <https://www.lockhaven.edu/~dsimanek/3d/stereo/3dgallery16.htm> [↑](#footnote-ref-2)
3. Z. Y. Zhang and H. T. Tsui, "3D reconstruction from a single view of an object and its image in a plane mirror," Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170), Brisbane, Queensland, Australia, 1998, pp. 1174-1176 vol.2, doi: 10.1109/ICPR.1998.711905. [↑](#footnote-ref-3)
4. Nguyen, Trong-Nguyen & Huynh, Huu & Meunier, Jean. (2019). Matching-based Depth Camera and Mirrors for 3D Reconstruction. [↑](#footnote-ref-4)
5. Heide, Felix & Xiao, Lei & Heidrich, Wolfgang & Hullin, Matthias. (2014). Diffuse Mirrors: 3D Reconstruction from Diffuse Indirect Illumination Using Inexpensive Time-of-Flight Sensors. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 10.1109/CVPR.2014.418. [↑](#footnote-ref-5)
6. Kyung, Ho & Jang, Young & Cha, Soon & Jung, Soon. (2003). 3D Reconstruction using Moving Planar Mirror. [↑](#footnote-ref-6)
7. Haihang, Wang & Xu, He & Pooneeth, Vishwanath & Gao, Xiao-Zhi. (2018). A Novel One-Camera-Five-Mirror Three-Dimensional Imaging Method for Reconstructing the Cavitation Bubble Cluster in a Water Hydraulic Valve. Applied Sciences. 8. 1783. 10.3390/app8101783. [↑](#footnote-ref-7)
8. <https://github.com/LibreScanner/ciclop> [↑](#footnote-ref-8)
9. <http://store.murobo.com/atlas-3d-kit/> [↑](#footnote-ref-9)
10. <https://structure.io/> [↑](#footnote-ref-10)
11. <https://www.einscan.com/>, <https://peel-3d.com/>, <https://www.metron3d.com/metron-e/>, <https://www.artec3d.com/portable-3d-scanners/artec-eva-lite-v2>, <https://www.creaform3d.com/en/handheld-portable-3d-scanner-goscan-3d>, <https://matterandform.net/scanner> [↑](#footnote-ref-11)
12. Pollefeys, Marc & Nistér et al., H.. (2008). Detailed Real-Time Urban 3D Reconstruction from Video. International Journal of Computer Vision. 78. 143-167. 10.1007/s11263-007-0086-4. [↑](#footnote-ref-12)
13. A. Akbarzadeh et al., "Towards Urban 3D Reconstruction from Video," Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06), Chapel Hill, NC, 2006, pp. 1-8, doi: 10.1109/3DPVT.2006.141. [↑](#footnote-ref-13)
14. Andras, Molnar. (2018). 3D Reconstruction of Monuments from Drone Photographs Based on The Spatial Reconstruction of The Photogrammetric Method. Advances in Science, Technology and Engineering Systems Journal. 3. 10.25046/aj030633. [↑](#footnote-ref-14)
15. <https://www.flytechuav.com/power-engineering.html> [↑](#footnote-ref-15)
16. Carr, Julia & Slyder, Jacob. (2018). Individual tree segmentation from a leaf-off photogrammetric point cloud. International Journal of Remote Sensing. 39. 1-16. 10.1080/01431161.2018.1434330. [↑](#footnote-ref-16)
17. <https://www.capturingreality.com/Article-CyArk-and-RealityCapture> [↑](#footnote-ref-17)
18. Lourakis, Manolis & Zabulis, Xenophon. (2013). Accurate Scale Factor Estimation in 3D Reconstruction. 8047. 498-506. 10.1007/978-3-642-40261-6\_60. [↑](#footnote-ref-18)
19. Kim, M., Huh, K., Yi, W., Heo, M., Lee, S., & Choi, S. (2012). Evaluation of accuracy of 3D reconstruction images using multi-detector CT and cone-beam CT. Imaging Science in Dentistry, 42, 25 - 33. [↑](#footnote-ref-19)
20. <https://cyark.org/ourMission/caseStudies> [↑](#footnote-ref-20)
21. A. Molnar "3D Reconstruction of Monuments from Drone Photographs Based on The Spatial Reconstruction of The Photogrammetric Method", Advances in Science, Technology and Engineering Systems Journal, vol. 3, no. 6, pp. 252-258 (2018). [↑](#footnote-ref-21)
22. <https://www.clevelandart.org/art/collection/search?i=2&only-open-access=1&only-in-3d=1> [↑](#footnote-ref-22)
23. <https://vision.princeton.edu/projects/2012/museum/paper.pdf> [↑](#footnote-ref-23)
24. <https://cyark.org/about/blog/?p=case-study-8> [↑](#footnote-ref-24)
25. <https://cyark.org/projects/mesa-verde/overview> [↑](#footnote-ref-25)
26. <https://wrcc.dri.edu/cgi-bin/cliMAIN.pl?comesa> [↑](#footnote-ref-26)
27. Ignacio Garcia-Dorado, Daniel G. Aliaga, Saiprasanth Bhalachandran, Paul Schmid, and Dev Niyogi. 2017. Fast Weather Simulation for Inverse Procedural Design of 3D Urban Models. *ACM Trans. Graph.* 36, 2, Article 21 (April 2017), 19 pages. DOI:https://doi.org/10.1145/2999534 [↑](#footnote-ref-27)
28. <https://www.flir.com/products/vue-pro/> [↑](#footnote-ref-28)
29. <https://www.quadh2o.com/hexh2o/sarah2o-ready-to-fly/> [↑](#footnote-ref-29)
30. <https://www.opendronemap.org/webodm/> [↑](#footnote-ref-30)
31. <https://openheritage3d.org/data> [↑](#footnote-ref-31)
32. <https://zachhalvorson.com/roast-level-verification/> [↑](#footnote-ref-32)