Automatic Scaling in Structure-from-Motion Photogrammetry for 3D Fragmentation Size Distribution Measurement

2021

Zedrick Paul Lizardo Tungol

Akita University

Doctoral Thesis 2021

Automatic Scaling in Structure-from-Motion Photogrammetry for 3D Fragmentation Size Distribution Measurement

Department of Earth Resource Engineering and Environmental Science

Graduate School of International Resource Sciences, Akita University

6519107 Zedrick Paul Lizardo Tungol

Supervisor:

Professor Tsuyoshi Adachi

Acknowledgements

<To be added later>

Akita University

Table of Contents

Acknowledgements i

Table of Contents ii

List of Figures iv

List of Tables vi

Abstract vii

1 Introduction 1

1.1 Background and Overview of the Study 1

1.2 Research Objectives 3

1.3 Chapter Structure 4

2 Review of Related Literature 6

2.1 Fragmentation Size in Mining Operations 6

2.2 Previous Methods of Fragmentation Size Distribution Measurement 7

2.2.1 Visual estimation and Sieve Analysis 7

2.2.2 Optical Granulometry (WipFrag) 8

2.2.3 Fragmentation Modeling 10

3 Materials and Methods 13

3.1 Structure-from-Motion – Multi View Stereo (SfM-MVS) 13

3.1.1 Keypoint Detection 15

3.1.2 Keypoint Matching 16

3.1.3 Keypoint Filtering 17

3.1.4 Sparse Reconstruction: Structure-from-Motion 18

3.1.5 Dense Reconstruction: Multi View Stereo 19

3.2 Application in the Geosciences 20

3.3 SfM Experiment on Mining Muckpiles 22

3.3.1 OpenDroneMap (ODM) 23

3.3.2 AliceVision Meshroom 25

3.3.3 Reconstruction Analysis 27

3.3.4 Limitations in the Reconstruction Process 29

3.4 GNSS-aided Scaling in Photogrammetry 29

3.4.1 GNSS-aided scaling in Bundle Adjustment 31

3.4.2 Preliminary experiment for validating GNSS-aided scaling fundamentals 32

4 Experiments on GNSS-aided Photogrammetry 36

4.1 GNSS-constrained SfM on monuments of known dimensions 36

4.2 Experiment on a pseudo-muckpile 39

4.3 Experiments on a Muckpile in Mikurahana Quarry Site 46

5 Conclusion 52

6 References 54

List of Figures

Figure 1. Application of a 3D Fragmentation Measurement System. 2

Figure 2. Relationship between fragmentation size and operational costs in mining....6

Figure 3. Progression of WipFrag fragmentation analysis from base image (top) to net of rock edges (middle) and identified rock fragments (bottom) (Palangio et al., 1995). 9

Figure 4. Example particle size distribution output from WipFrag (Wipware Inc., 2017). 10

Figure 5. A sample fragmentation curve created using the Kuz-Ram Model (Cunningham, 2005). 12

Figure 6. An illustration of the coordinate systems. Described is the conversion of world point Pw to camera point Pc, to imaging plane coordinates (x, y) and finally pixel coordinates (u, v). (Madali, 2020) 14

Figure 7. SfM-MVS pipeline. 15

Figure 8. Keypoint detection on a pile of rocks. 16

Figure 9. Graphical representation of image gradients and keypoint descriptors (Lowe, 2004). 17

Figure 10. Keypoint matching tracks over two different views.....18

Figure 11. SfM example showing common points. 19

Figure 12. Sparse point cloud (Left) and dense point cloud (Right) of a scene. 20

Figure 13. 3d mapping of a landslide using SfM-MVS (Carrivick et al., 2016) 21

Figure 14. A single frame taken from the video of the muckpile. Presence of soccer ball for scaling purposes. 22

Figure 15. A sample set of images used in the reconstruction.....23

Figure 16. Workflow utilizing OpenDroneMap to create 3d models 24

Figure 17. Dense 3d point cloud from OpenDroneMap 24

Figure 18. Textured model of the underground muckpile, output from OpenDroneMap 25

Figure 19. Workflow utilizing Meshroom to create 3d models.....26

Figure 20. Dense reconstruction (uncolored) from Meshroom 26

Figure 21. Textured Model output from Meshroom.....27

Figure 22. Close-up of the Textured Mesh generated by Meshroom. Note that the soccer ball is almost unidentifiable and the presence of floating 'blobs' 28

Figure 23. Close-up of the Textured Mesh generated by ODM. The soccer ball has relatively been well reconstructed and no floating 'blobs' are present 28

Figure 24. GNSS-aided photogrammetry workflow.....31

Figure 25. Grid layout of the 1st preliminary experiment (left) and a toy as the object (right). Red dots indicate camera positions and the blue rectangle indicates the board that was inserted for improved feature detection. 33

Figure 26. Data input and output of the preliminary experiment. 34

Figure 27. Data input and output of the experiment using 200 images. 36

Figure 28. Mesh reconstructed 3D CG model of the monuments. 3rd picture (bottom) shows measured side of cube monument when using 100 images. 37

Figure 29. GNSS location of camera positions as logged by a smartphone (red dots) of the experiment on monuments of known dimensions.....38

Figure 30. Graph detailing the results of the experiment.....39

Figure 31. Data input (Set #1) and output of the experiment using 100 images. 40

Figure 32. GNSS location of camera positions as logged by a smartphone (red dots) of the experiment on a pseudo-muckpile. 40

Figure 33. Meshed 3D CG reconstruction of wooden box reference with measurement (at 100 images used). 41

Figure 34. Data input (Set #2) and output of the experiment using 100 images. 42

Figure 35. Close up of Set #2 meshed 3D CG reconstruction, with measurements on long rectangular prism-shaped rock that was used as reference. 43

Figure 36. 3D CG reconstruction results of the combined data sets. 45

Figure 37. Satellite view of the Mikurahana quarry site; the red box indicates location of the muckpile and its photographs. 47

Figure 38. Close-up of quarry and muckpile photograph locations (marked as red dots) 47

Figure 39. Sample images from the dataset. The yellow box is used as an absolute reference. 48

Figure 40. Sparse and dense reconstruction of limestone muckpile using 400 images, shown at different angles. 49

Figure 41. Close-up of mesh reconstruction, featuring reference yellow box and dimension of its longest side as measured. 49

Figure 42. Graph detailing the results of the Mikurahana experiment. 50

List of Tables

Table 1. Results of the Experiment on Monuments of Known Dimensions. 38

Table 2. Results of the experiment on a pseudo-muckpile 43

Table 3. Results of the experiment on the Mikurahana quarry muckpile 50

Akita University

Abstract

In any mining operation, there are a number of factors that can affect the efficiency of the day-to-day resource extraction process across all of its stages from mine (drilling, blasting, haulage) to mill (mineral processing). As such, it is up to the management, when possible, to monitor these various factors and act accordingly by making modifications to mine planning as well as tweaking its execution. One of such factors, specifically in mining operations that employ explosives and mineral processing, is the fragmentation size distribution of rock after it has been blasted.

Scaling is a critical component of fragmentation size distribution measurement using photogrammetry as it will directly determine the accuracy of the size estimation. In creating a 3D model, extrinsic data such as ground truths are needed to create a properly-scaled reconstruction of the scene. There are several methods that are used to resolve scale in photogrammetry. Most of these methods have the same basic idea in that once the exact distance between at least two different points in a scene is known, a scale factor can be applied to the 3D model. One way to do this is to include an object of known length such as scale bars in the scene. In larger applications such as aerial mapping, GCPs (Ground Control Points) are used, which are marked points of known absolute or relative coordinates.

The study aims to create a system for creating a scaled 3D model without the use of ground truth data such as GCPs (Ground Control Points) for the purpose of improving fragmentation size distribution measurement using positional data such as GNSS (Global Navigation Satellite System)-aided photogrammetry. To achieve this, the study firstly aimed to 1) investigate the effect of camera positional data and constraints on 3D model scaling accuracy; then 2) simultaneously collect image and

positional data (e.g., GNSS) from scenes that are to be reconstructed using photogrammetry; then 3) use the positional data in the photogrammetry workflow to scale the resulting 3D model; then 4) observe the effect of increasing the number of datapoints (image + positional data) on the scaling accuracy of the generated 3D model and finally 5) Determine the most effective configuration in data taking and data processing to achieve acceptable scale with the least number of datapoints needed.

A preliminary experiment that was done show results that constraining camera positions to locations, relative or otherwise improves the accuracy of the generated 3D model. With this fundamental idea in mind, the study moved on to larger scale experiments that involved the actual use of GNSS positional data in conjunction with image data in 3D photogrammetry. In these further experiments, results show that the scale error decreases when more images from the same dataset are used. In conclusion two observations have been drawn from the study: firstly, constraining cameras to accurate positions in SfM will result in a properly scaled 3D model; secondly, increasing the number of georeferenced images in SfM will incrementally improve the scaling error of the reconstruction. These results lend credence to the possibility of improving the scaling aspect of 3D fragmentation measurement systems without the use of GCP or manual scales, specifically in surface mines where GNSS data is generally readily available. This shows that monitoring the fragmentation distribution can potentially be performed using just a camera and a GNSS-enabled devices, such as smartphones.

Akita University

1 Introduction

1.1 Background and Overview of the Study

A mining operation contains a sizable number of factors that affect its efficient extraction of resources. One of such factors is fragmentation size, a key parameter across almost all of the stages of production from mine (drill, blasting, haulage, etc.) to mill (mineral processing). Several studies have explored the effect of fragmentation size on other factors such as drilling and blasting cost (Afum & Temeng, 2014) and the performance of crushing and grinding circuits (Kanchibotla et al., 1999; Valery et al., 2001). It can be therefore considered that it is vital for companies to continuously monitor fragmentation size and make necessary changes to mine planning and execution to keep the fragmentation size that is most beneficial to the operation as a whole. Traditionally, methods such as manual sieving, boulder counting and visual estimation have been used for fragmentation size measurement. However, due to the generally large amount of material being mined as well its innately heterogeneous nature, difficulties arise from using traditional methods. Limitations on sampling as well as bias make these methods relatively inefficient (Nefis & Talhi, 2016). As such, there exists a need for a quick and accessible method of rock fragmentation size distribution determination that can surmount the limitations of physical sampling and laboratory analysis. A currently used digital solution to this problem is to employ image-based particle size analysis software. Commercial products such as WipFrag (Palangio et al., 1995) make use of images of a muckpile or orthomosaics to measure fragmentation size distribution.

A 3-Dimensional Fragmentation Measurement (3DFM) system was developed that makes use of 3D Photogrammetry to measure particle size distribution at accuracies greater than that of conventional methods (Jang et al., 2020). A theoretical visualized workflow for this particular system when applied to a mining operation is shown in Figure 1.

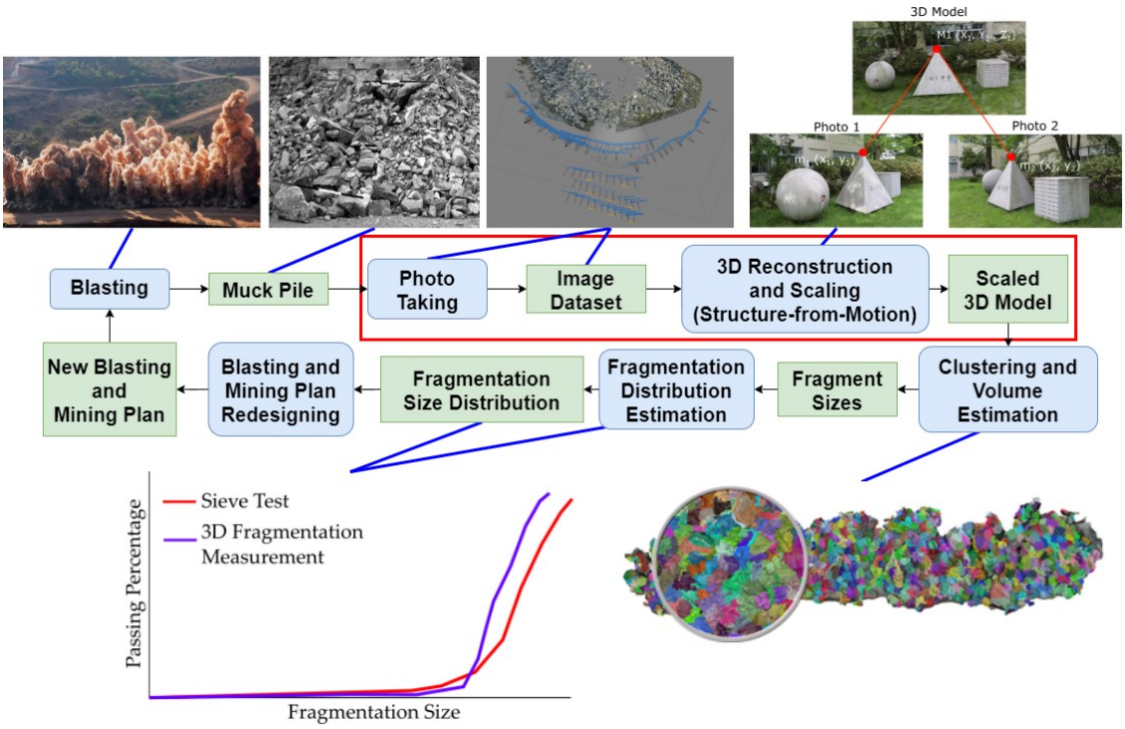

Figure 1. Application of a 3D Fragmentation Measurement System.

The developed system is divided into stages, utilizing multiple computational techniques in order to achieve its purpose. In a hypothetical application of the system, pictures of the muckpile from the products of blasting are taken. The sizes of muckpiles vary greatly depend on the specifications of the hauling equipment as well as the mine plan that the operation employs. In situations where the muckpile is too large or have parts that are inaccessible to photo-taking, it is possible for the system to reconstruct only a representative “slice” of the muckpile. The images are then processed in a high-power computer by a sequence of 3D imaging techniques that will ultimately output a scaled 3D model of the muckpile in the form of a point cloud. A technique known as

supervoxel clustering is then performed on the 3D model then undergoes supervoxel clustering in order to divide the individual fragments into segments whose dimensions have been calculated. The dimensional data can then be used in the computation of the fragment size distribution of the muckpile. Using this information, the blasting product can be judged if it is up to the expected specification. Adjustments are then made the blasting design such as the amount and type of explosive and blasting patterns in order to achieve the required distribution.

Scaling is a critical component of fragmentation size distribution measurement using photogrammetry as it will directly determine the accuracy of the size estimation. In creating a 3D model, extrinsic data such as ground truths are needed to create a properly-scaled reconstruction of the scene. There are several methods that are used to resolve scale in photogrammetry. Most of these methods have the same basic idea in that once the exact distance between at least two different points in a scene is known, a scale factor can be applied to the 3D model. One way to do this is to include an object of known length such as scale bars in the scene. In larger applications such as aerial mapping, GCPs (Ground Control Points) are used, which are marked points of known absolute or relative coordinates.

1.2 Research Objectives

This study focuses on the 3D model scaling aspect of this system, as highlighted with a red box in Figure 1. To this end, the study aims to investigate the utilization of other more readily available extrinsic data such as GNSS positional data in order to achieve proper scaling of 3D models generated through photogrammetry without the need of physical manipulation of the scene with things such as scale bars and GCPs. Specifically, this study aimed to achieve these objectives:

- Investigate the effect of constrained camera positional data on 3D model accuracy.
- Simultaneously collect image and positional data (e.g., GNSS) from scenes that are to be reconstructed using photogrammetry.
- Use the positional data in the photogrammetry workflow to scale the resulting 3D model.
- Observe the effect of increasing the number of datapoints (image + positional data) on the scaling accuracy of the generated 3D model.
- Determine the most effective configuration in data taking and data processing to achieve acceptable scale with the least number of datapoints needed.

1.3 Chapter Structure

Chapter 2 discusses previous literature related to the study. Specifically, it discusses the importance of fragmentation size in mining operations, as well as traditional methods of fragmentation size distribution measurement. Chapter 3 discusses the various materials and methods that are involved in the creation of a scaled 3D model using photogrammetry. Chapter 3.1 discusses the general SfM-MVS (Structure-from-Motion – Multi View Stereo) workflow and Chapter 3.3 includes a previous preliminary study done by the author that involves the creation of a 3D model of a muckpile found in an underground mine using two different software. Chapter 3.4 focuses on the utilization of GNSS data in the SfM-MVS workflow as well as a preliminary experiment that was used to validate fundamentals regarding the effect of positional data and scaling accuracy. Chapter 0 discusses the two experiments that were performed to observe the effectivity of GNSS-aided photogrammetry when it comes to

model scaling, including the results and the discussion. Chapter 5 discusses the conclusions that were made based on the results of the experiments as well as recommendations for further steps and future studies.

2 Review of Related Literature

2.1 Fragmentation Size in Mining Operations

The mine and the mill usually have an agreed-upon average fragmentation size from which various elements of the operation will be based on, such as drilling equipment, explosives, and mineral processing machinery. The importance of this specification is due to how it generally affects mining and milling costs differently. When a relatively high fragmentation size is set, the cost of drilling and blasting is proportionately lower (Afum & Temeng, 2014), but will require more mineral processing in order to achieve the final product. Inversely, a relatively lower fragmentation size means less expense is needed for milling, but the cost of drilling and blasting would increase (BME South Africa, 2016). Considering these effects, a middle ground must be reached in order to minimize the total cost of the operation, as illustrated by a theoretical graph in Figure 2.

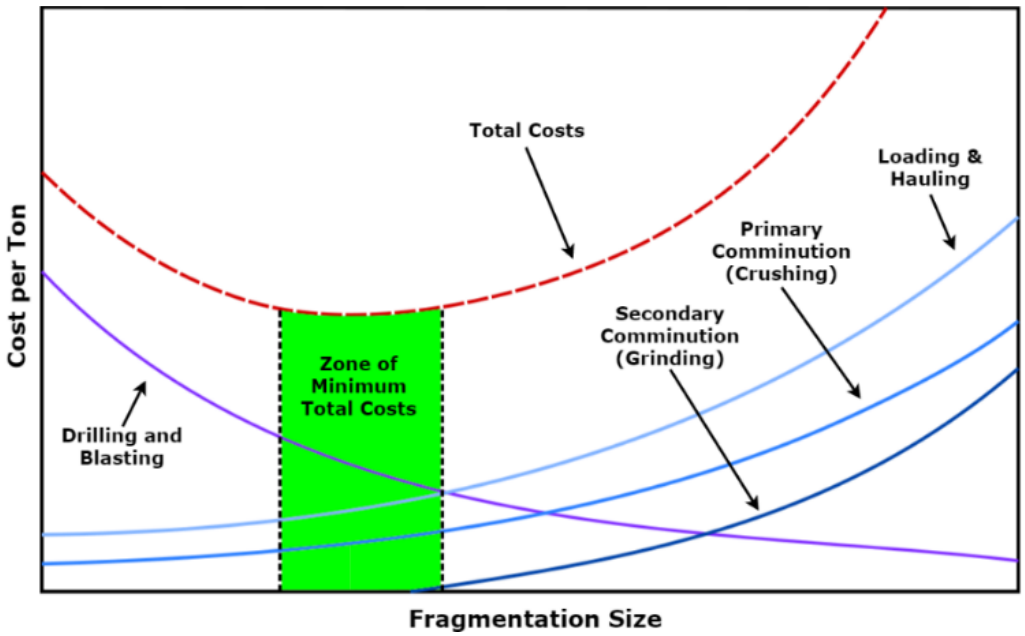

Figure 2. Relationship between fragmentation size and operational costs in mining

2.2 Previous Methods of Fragmentation Size Distribution Measurement

The importance of measuring and monitoring fragmentation size distribution manifests itself in the development of various methods and techniques. Research on this particular topic has attracted experts from a variety of fields such as computer vision, machine learning and even simulation through mathematical models.

2.2.1 Visual estimation and Sieve Analysis

One of the earliest and easiest methods of evaluating fragmentation size was through visual estimation, which was done by individually counting rocks of predefined sizes in a muck pile, sometimes through the use of photographs (Grant & Dutton, 1983). While this method proved to be quick and inexpensive, it was a highly subjective method and was only accurate up to a certain level (Aler et al., 1996).

Another earlier method which was borne out of the need for greater accuracy and consistency was manual sieving. Some applications would include using samples taken from scaled down blasting tests (Bergmann et al., 1973). However, the method was traditionally used on a representative sample directly from the muck pile or at times, the crusher product. Using this process provided highly accurate results and is actually usually used as a reference for determining the performance of new fragmentation size distribution measurement methods, including this study. Despite this, however, regularly performing this in an actively producing operation has been widely known to be very costly and time consuming (Kanchibotla et al., 1999).

2.2.2 Optical Granulometry (WipFrag)

The next significant step in fragmentation size measurement for mining applications was the development of photographic analysis methods. These applications came to be known as Optical Granulometry. It is a method in which an image of the product is digitally analyzed by a series of algorithms which essentially detect the edges of individual fragments in the muckpile. Using this data, a size distribution can be made for that particular photo. This technique has proven to be an accurate method that is sufficiently quick, inexpensive, doesn't interfere with production, and is not limited by the volume of the product (Palangio et al., 1995).

One of the more popular applications of this technique is the WipFrag Image Based Granulometry System, which was developed by WipWare Inc. Some testing has shown that it is relatively more accurate than some of the other image-based fragment size analysis suites such as FragScan and Split (Liu & Tran, 2018).

WipFrag utilizes images taken from various sources and of various formats. Using automatic algorithms, a "net" is created that outlines the edges of the fragments. Measuring this net, WipFrag reconstruct an imaginary 3-d distribution using geometric probability along with some calibrations. From this 3-d distribution a fragmentation size distribution can be produced. Practical tests comparing WipFrag measurements to sieving tests during its introduction have yielded errors ranging from 2 to 20% (Palangio et al., 1995)

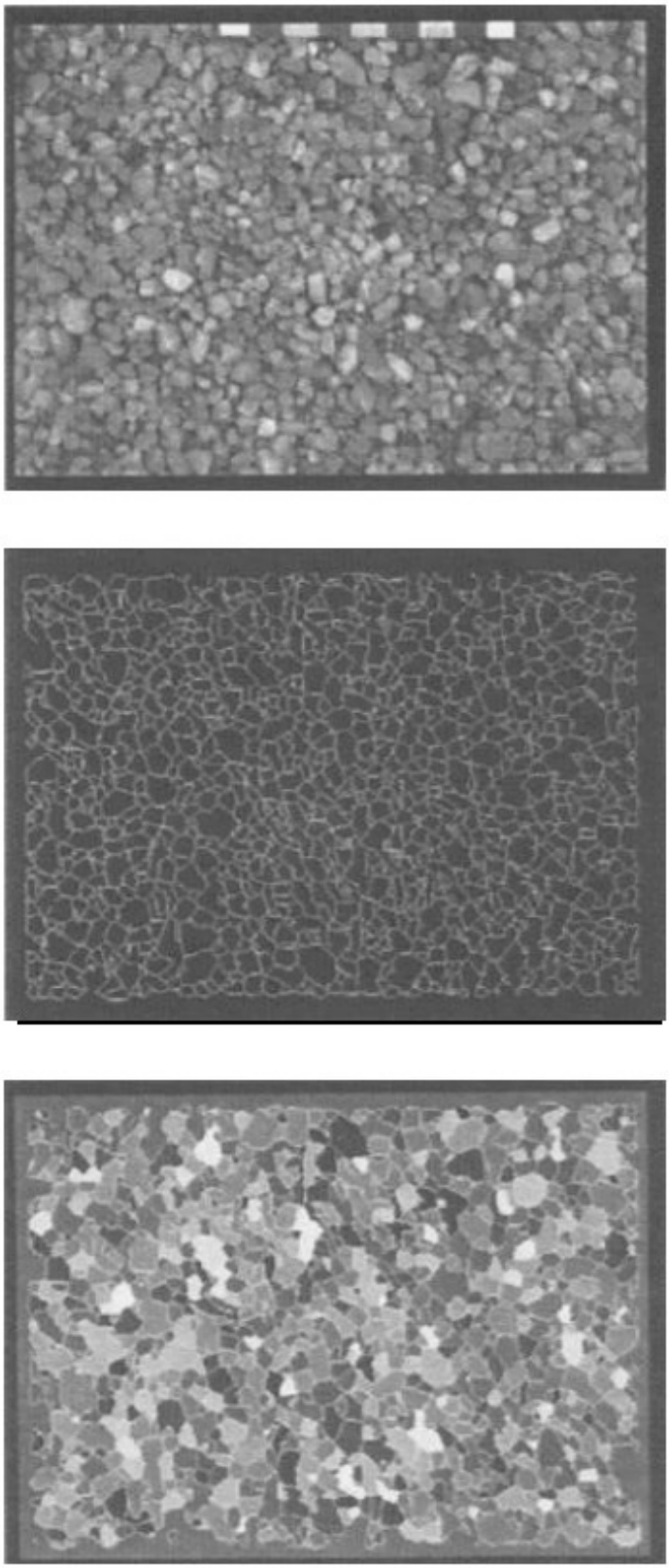

Figure 3. Progression of WipFrag fragmentation analysis from base image (top) to net of rock edges (middle) and identified rock fragments (bottom) (Palangio et al., 1995).

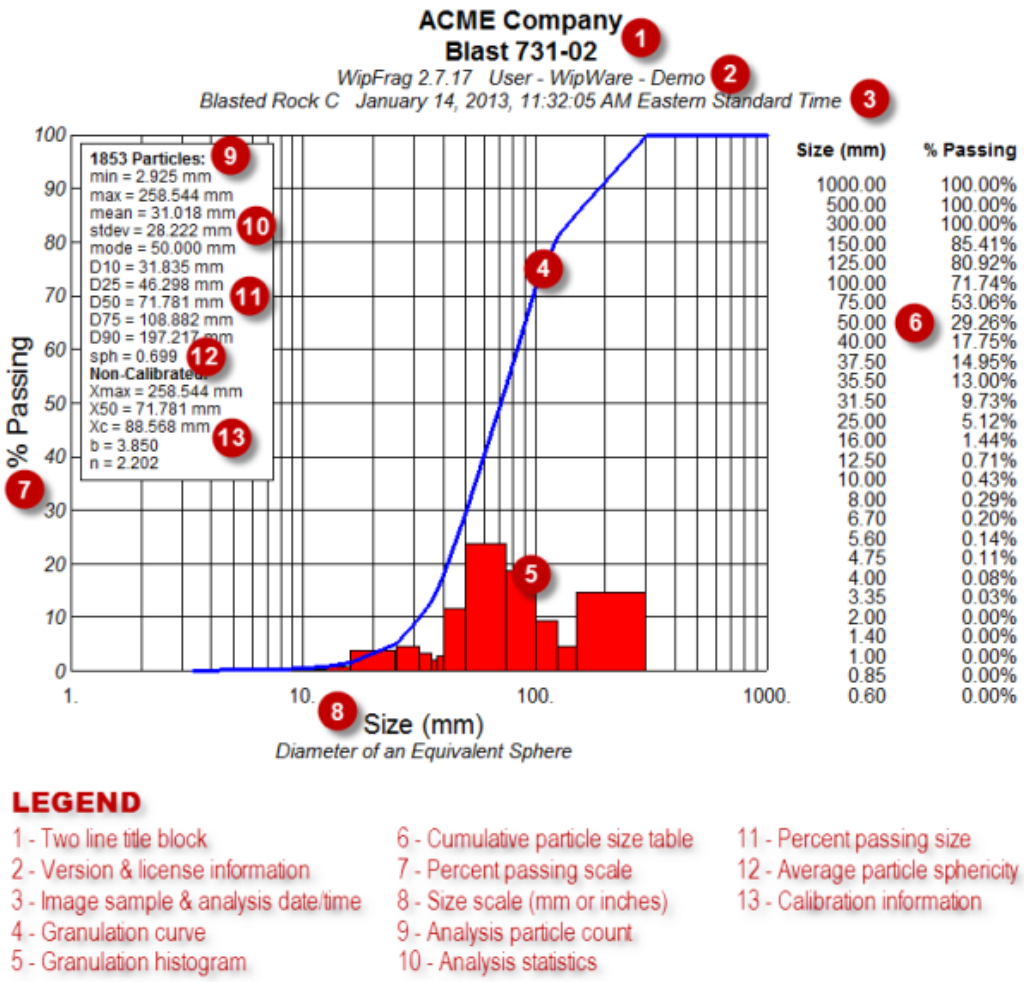

Figure 4. Example particle size distribution output from WipFrag (Wipware Inc., 2017).

2.2.3 Fragmentation Modeling

Another alternative approach that was taken in this line of research was to predict the fragmentation size distribution before blasting was performed. A number of these models have been developed throughout the years, in an attempt to simulate size distribution from certain blast designs. The approaches fall mainly into two categories: empirical modelling, which assumes that the higher the energy input, the finer the fragmentation would be, and mechanistic modeling, which analyzes the physics of detonation as well as the energy transfer in well-defined rock for particular blasting layouts, in the process getting the whole range of blasting results. Despite mechanistic

modeling allowing for the illustration of the effect of individual mechanisms within the system, something that purely empirical models cannot achieve, it is quite difficult to apply on a daily basis due to its long run times and the requirement of collecting significantly more data from detonation to the end result, and in the end is not essentially more accurate. As such, empirical models have been considered to be the more practical tool for daily blast designing and amongst the various applications, the Kuz-Ram model was the most popular. It featured three key equations:

The adapted Kuznetsov equation,

$$x_m = AK^{-0.8}Q^{\frac{1}{6}}\left(\frac{115}{RWS}\right)^{19/20}, \qquad (1)$$

where x_m = mean particle size, cm; A = rock factor (varying between 0.8 and 22, depending on hardness and structure) K = powder factor, kg explosive per cubic meter of rock; Q = mass of explosive in the hole, kg; RWS = weight strength relative to ANFO.

The adapted Rosin-Rammler equation,

$$R_x = \exp\left[-0.693\left(\frac{x}{x_m}\right)^n\right], \qquad (2)$$

where R_x = mass fraction retained on screen opening x; n = uniformity index, usually between 0.7 and 2.

And the uniformity equation,

$$n = \left(2.2 - \frac{14B}{d}\right)\sqrt{\left(\frac{1+S/B}{2}\right)}\left(1 - \frac{W}{N}\right)\left(abs\left(\frac{BCL-CCL}{L}\right) + 0.1\right)^{0.1}\frac{L}{H}, \qquad (3)$$

where B = burden, m; S = spacing, m; d = hole diameter, mm; W = standard deviation of drilling precision, m; L = charge length, m; BCL = bottom charge length, m; CCL = column charge length, m; H = bench height, m.

Used together, these equations were able to ultimately output the mass fraction of a particular size and has been widely used since its inception in the 1980s. However,

deficiencies such as certain parameters (rock properties, blasting dimensions, timing, etc.), difficulty in scaling blasting effects and its limited ability to measure fragmentation were present in the initial version of the model. As such, changes were made to the model to address these deficiencies and the model has been evolving since then (Cunningham, 2005).

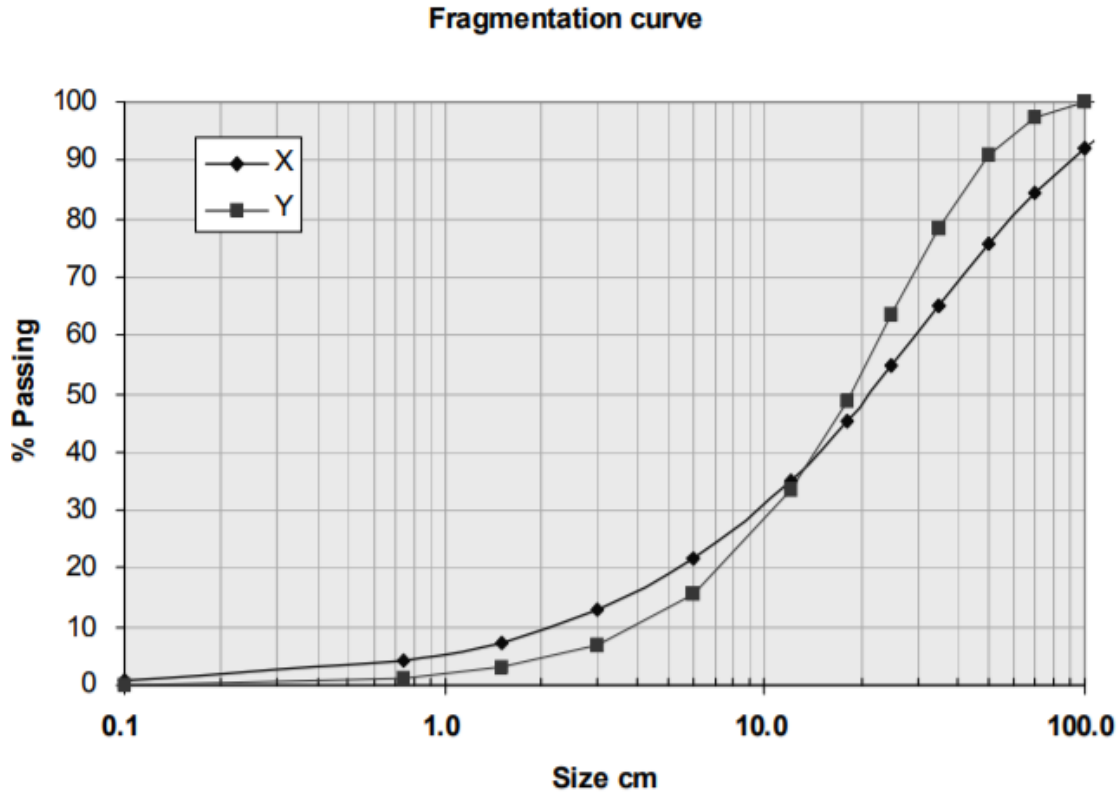

Figure 5. A sample fragmentation curve created using the Kuz-Ram Model (Cunningham, 2005).

3 Materials and Methods

3.1 Structure-from-Motion – Multi View Stereo (SfM-MVS)

Structure from Motion (SfM) has been collectively defined as the photogrammetric technique (Westoby et al., 2012), the process (Schönberger & Frahm, 2016a), as well as the tools (Carrivick et al., 2016) used to generate 3D models from 2D images taken at different angles. It has been developed since the 1980s, resulting in various applications such as photogrammetric surveys, virtual reality model reconstruction, determination of camera motion (Cipolla & Robertson, 2009), and odometric scale estimation (Gräter et al., 2015). Compared to traditional photogrammetry where the calculation is more direct, SfM makes use of repeating algorithms to identify matching features in a set of overlapping images, and use these matched features to calculate camera location and orientation as well (Carrivick et al., 2016). SfM can be computed in several ways, depending on numerous factors such as camera type, image ordering, capture format, and more.

Mathematically speaking, SfM can be described as the conversion of four coordinate systems, illustrated in Figure 6:

1. Image pixel coordinate system, which concerns the pixels on the 2D image.
2. Imaging plane coordinate system, which lies on the same plane of the previous system, but whose origin is the plane's intersection with the camera's optical axis.
3. Camera coordinate system, which concerns a pinhole camera's point of view of the image.
4. World coordinate system, which is a reference system to describe the position of the camera and the objects being taken pictures of.

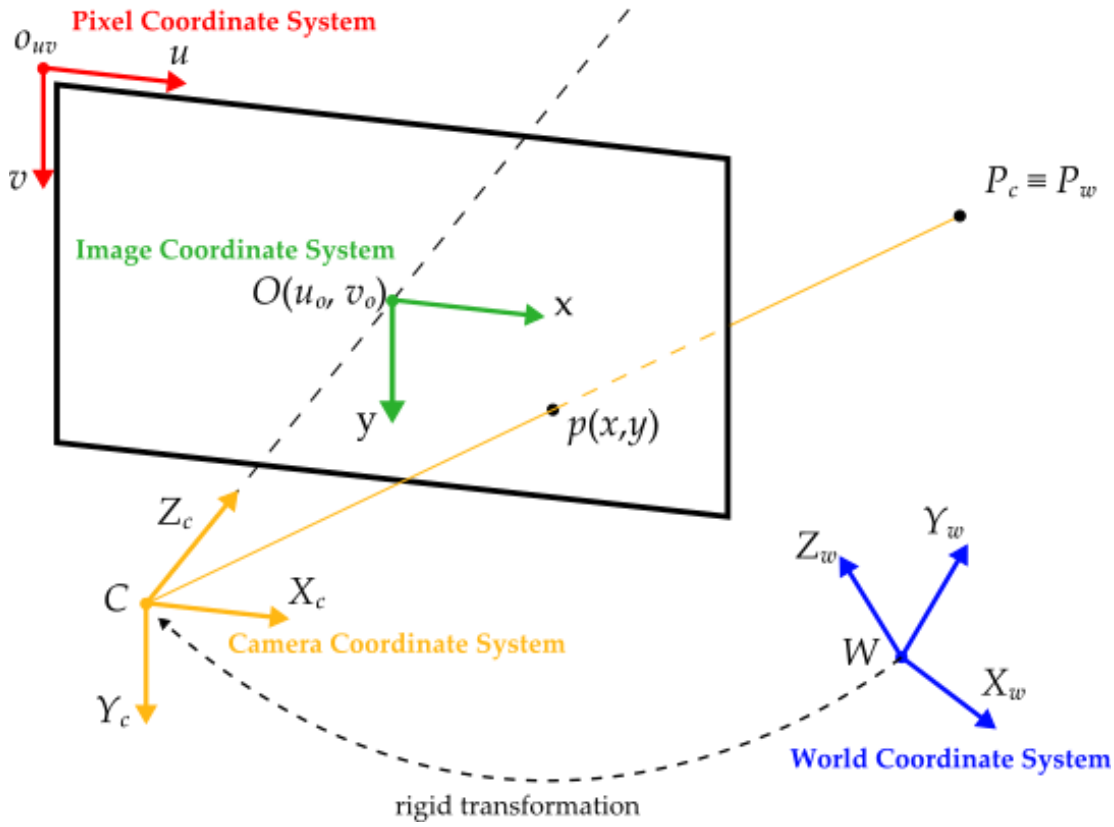

Figure 6. An illustration of the coordinate systems. Described is the conversion of world point Pw to camera point Pc, to imaging plane coordinates (x, y) and finally pixel coordinates (u, v). (Madali, 2020)

The conversion of these 4 coordinates systems can be described by equation (4). u and v describe the axes in the imaging planes. u_0 and v_0 are the coordinates of the origins of the imaging plane in the pixel coordinate system. δ_x and δ_y represent the physical size of each pixel in the image in the imaging plane (zoom ratio). f describe the focal length, which is the distance from the optical center of the camera to the pixel plane. $\boldsymbol{R} \in \mathbb{R}^{3\times3}$ and $\boldsymbol{t} \in \mathbb{R}^3$ describes the rotational and translational vectors that relate the camera and the world coordinate systems. X_w, Y_w, Z_w are the actual coordinates of a point in the world coordinate system (Madali, 2020).

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\delta_x} & 0 & u_o \\ 0 & \frac{1}{\delta_y} & v_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} [\boldsymbol{R} \quad \boldsymbol{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\boldsymbol{R} \quad \boldsymbol{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (4)$$

Equation (4) represents the fact that in order to estimate the position of a point in the real world, the external parameter matrix of the camera (i.e., R and t) needs to be measured first. Once R is known, the relative position of the object in the world coordinate system can be estimated, and once t is known, the absolute position can also be acquired as well. Basic SfM can relatively estimate R and t.

It can be inferred that there is no single 'correct' workflow or process in the conversion of 2D images into models. However, there are key processes that are present in almost all applications of the method, as shown in the steps are briefly described in the following sections.

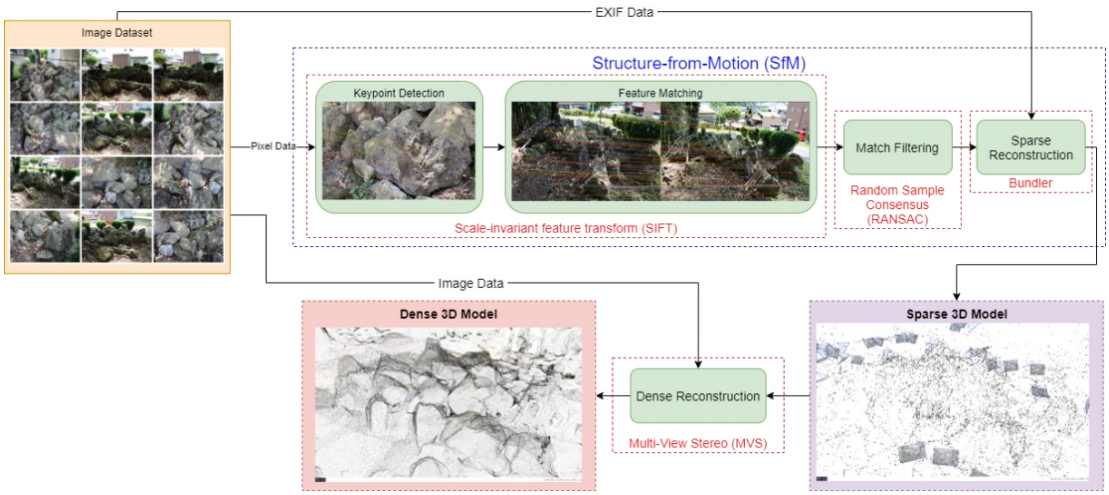

Figure 7. SfM-MVS pipeline.

3.1.1 Keypoint Detection

The initial processing step after acquiring the images is feature detection, or extraction, where possible common features (keypoints) in the individual images are

identified as shown in Figure 8. It is by these features that allow the different images in the dataset to be matched at the next stage. There are several techniques that have been developed for the solution of this step (Carrivick et al., 2016) but the most widely used amongst modern SfM applications is the scale-invariant feature transform (SIFT) (Lowe, 2004). The system recognizes feature points in the image set which are uniform in scaling and rotation and relatively uniform to changes in lighting and 3D camera view angles. The number of keypoints that are extracted in an image relies heavily on the resolution and texture of the images themselves, with high-quality, original-resolution pictures returning the most results (Westoby et al., 2012).

Figure 8. Keypoint detection on a pile of rocks.

3.1.2 Keypoint Matching

The next step is to match the keypoints and identify the correspondences between them. Matches are found by identifying a keypoint's nearest neighbor in the database. The nearest neighbor is defined as the keypoint with the least Euclidean distance for its descriptor vector, as shown in Figure 9. (Lowe, 2004). It is also important to note at this point that not all keypoints are guaranteed to have a good match in the dataset. It is therefore necessary to discard these unmatched keypoints, making use of the ratio be-tween the Euclidean distance of the nearest neighbor with that of the second nearest, at a certain minimum value as a criterion for discarding false keypoint matches (Carrivick et al., 2016). The inherent complexity of the keypoint descriptors

gives rise to the need of an efficient solution to the search process, as brute-force searching for nearest neighbors proves to be computationally difficult and time-consuming. Several solutions such as k-dimensional trees (k-d trees), best-bin first (BBF) and approximate nearest neighbor (ANN) searching are used to solve this problem of efficiency by partitioning the data into bins which are prioritized for match searching, decreasing the number of recursions needed to go through all the keypoints (Carrivick et al., 2016).

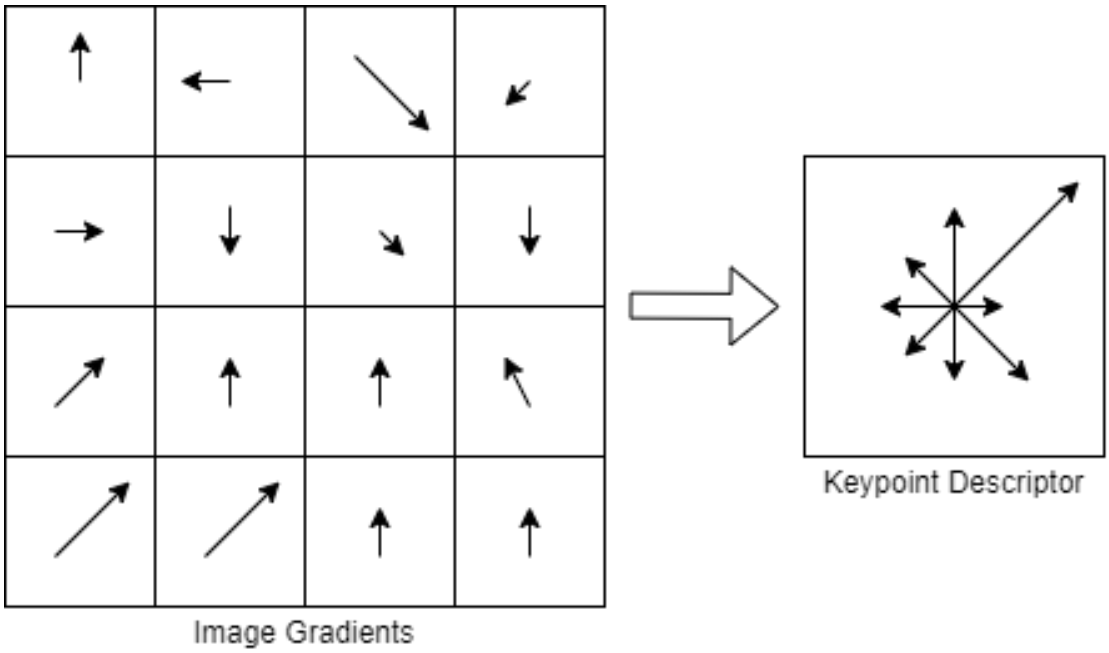

Figure 9. Graphical representation of image gradients and keypoint descriptors (Lowe, 2004).

3.1.3 Keypoint Filtering

The third stage, also known as geometric verification or match filtering, is done to further eliminate erroneous matches. Since the initial matching is solely based on ap-appearance, it cannot be guaranteed that the matched keypoints refer to the same point in an image (e.g., images with symmetrical or similar features) (Schönberger & Frahm, 2016b). SfM then needs to verify matches by mapping keypoints across images using projective imagery. An example of this step can be illustrated by the image pair in

Figure 10. The two images are of the same scene, taken at two different angles, and the keypoints found in both images are matched, as shown by colored matching tracks.

Figure 10. Keypoint matching tracks over two different views.

3.1.4 Sparse Reconstruction: Structure-from-Motion

The fourth step, which also by itself is sometimes called SfM (Structure-from-Motion), is to reconstruct the scene that was taken using 2D images into an initial sparse 3D structure. Using the verified matched keypoints, SfM aims to simultaneously reconstruct the: (a) 3D scene structure, (b) camera position and orientation (extrinsic parameters), and (c) intrinsic camera calibration parameters. The intrinsic camera parameters are defined by a camera calibration matrix that includes image scale, skew, and the principal point that is defined as the location on the image plane which intersects the optical axis. Further intrinsic parameters are also required to resolve additional internal aberrations such as distortion on non-pre-calibrated cameras. These intrinsic parameters are either included in the camera's image file format (e.g., EXIF) or will be resolved in additional intermediate steps. After this, a process known as bundle adjustment is used to produce sparse point-clouds (Schönberger & Frahm, 2016b). This process will be described further in another part of the study below as it is in this step that GNSS constraints will come to play in scaling the produced 3D model. A simplified illustration of this process is described by the illustration in Figure 11.

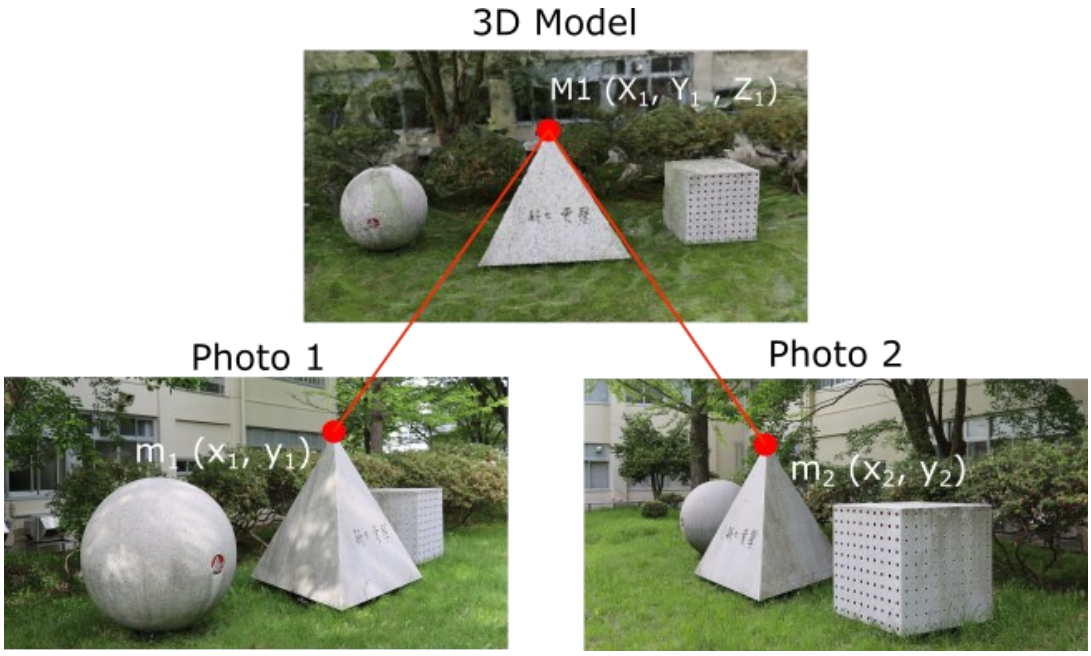

Figure 11. SfM example showing common points.

3.1.5 Dense Reconstruction: Multi View Stereo

An additional, post-processing method known as MVS (Multi-View Stereo) can be applied to the sparse 3D model from SfM in order to generate an enhanced "dense" 3D model. The final output of MVS is a complete 3D scene reconstruction from a collection of images of known intrinsic and extrinsic parameters, which is already resolved through SfM. A variety of MVS algorithms are available but recent variants called clustering views for MVS (CMVS) and patch-based MVS (PMVS) has been observed to perform well against other algorithms (Carrivick et al., 2016). CMVS decomposes the camera poses from bundle adjustment into manageable clusters and PMVS is used to independently reconstruct the 3-dimensional model from these clusters (Westoby et al., 2012). Most modern MVS pipelines, including the one in the software used for this study includes features from both these variants of MVS. A

comparison between point density between sparse and dense reconstruction is illustrated in Figure 12.

Figure 12. Sparse point cloud (Left) and dense point cloud (Right) of a scene.

3.2 Application in the Geosciences

In recent times, SfM has found various applications in the field of geoscience. Even though most of the research currently exist as proof-of-concept studies, SfM becomes closer to mainstream use in the geosciences as it is subjected to the rigors of the scientific method, improving the analysis of error and uncertainty. Presently, SfM-MVS finds its use commonly in land surveys, specifically in the creation of digital elevation models (DEM), which are gridded from 3d point cloud for use in mapping, landform geometry analysis, and even for input in numerical models. DEMs have the potential to be used to detect changes in topography and subsequently deduce dynamic processes in varying geological environments.

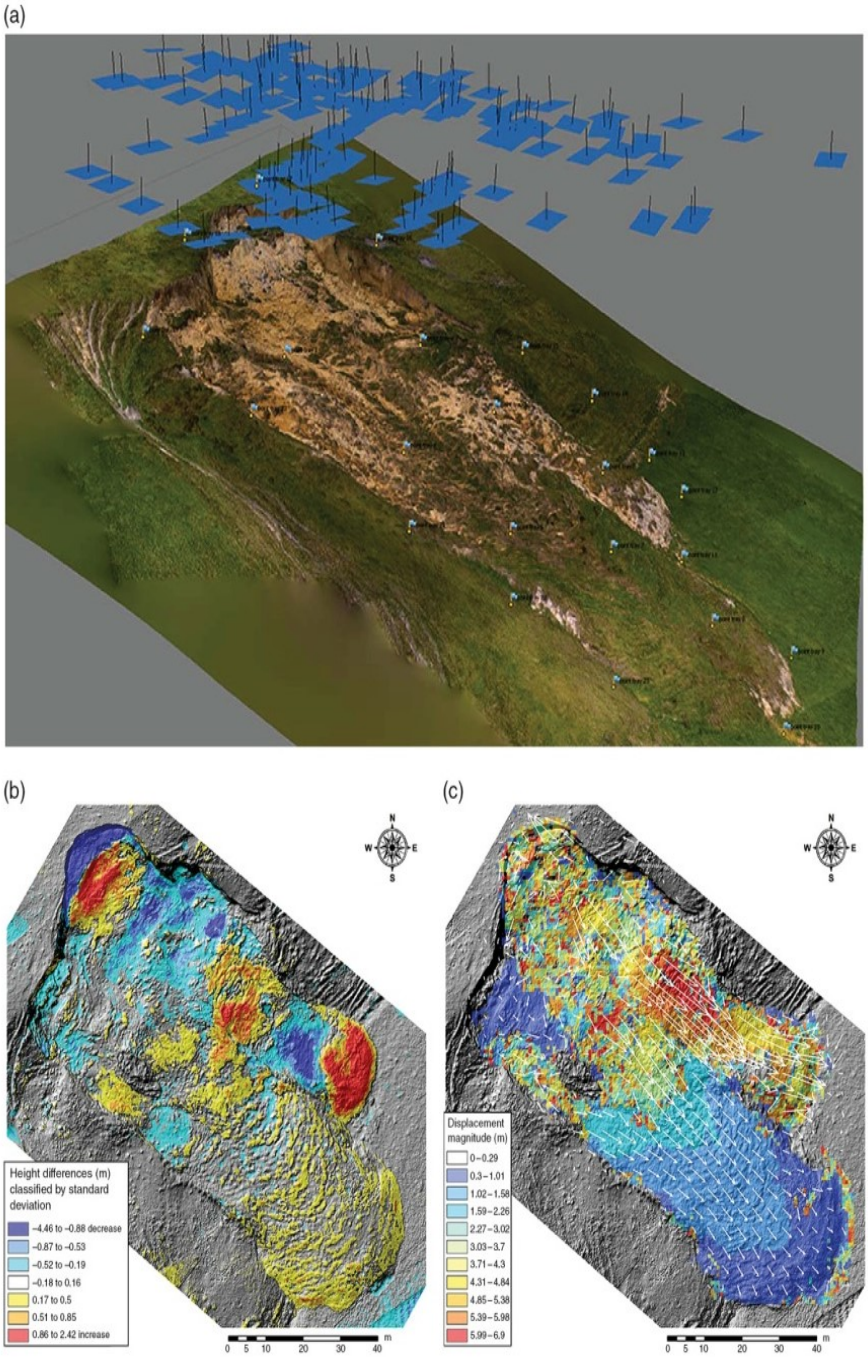

Figure 13. 3d mapping of a landslide using SfM-MVS (Carrivick et al., 2016)

Inspiration for this research has been drawn from one of the more novel applications of SfM-MVS in the geosciences, which was to ascertain the volume of various geological features such as stalagmites and boulders by using 3d point cloud data.

3.3 SfM Experiment on Mining Muckpiles

In a previous study (Tungol, 2019), a video recording of a muckpile in a tunnel in an underground mine operated by Kalgoorlie Consolidated Gold Mines (KCGM) in Kalgoorlie, Australia was used as the test data for reconstruction using SfM-MVS. Individual images were extracted from the video for use in the reconstruction process. A total of 218 images were extracted from the video. A sample set of images used in the reconstruction. OpenDroneMap and Meshroom were used to create 3d reconstructions of the muckpile, and the results are viewed on MeshLab, an open-source 3d model mesh processing tool (Cignoni et al., 2008).

Figure 14. A single frame taken from the video of the muckpile. Presence of soccer ball for scaling purposes.

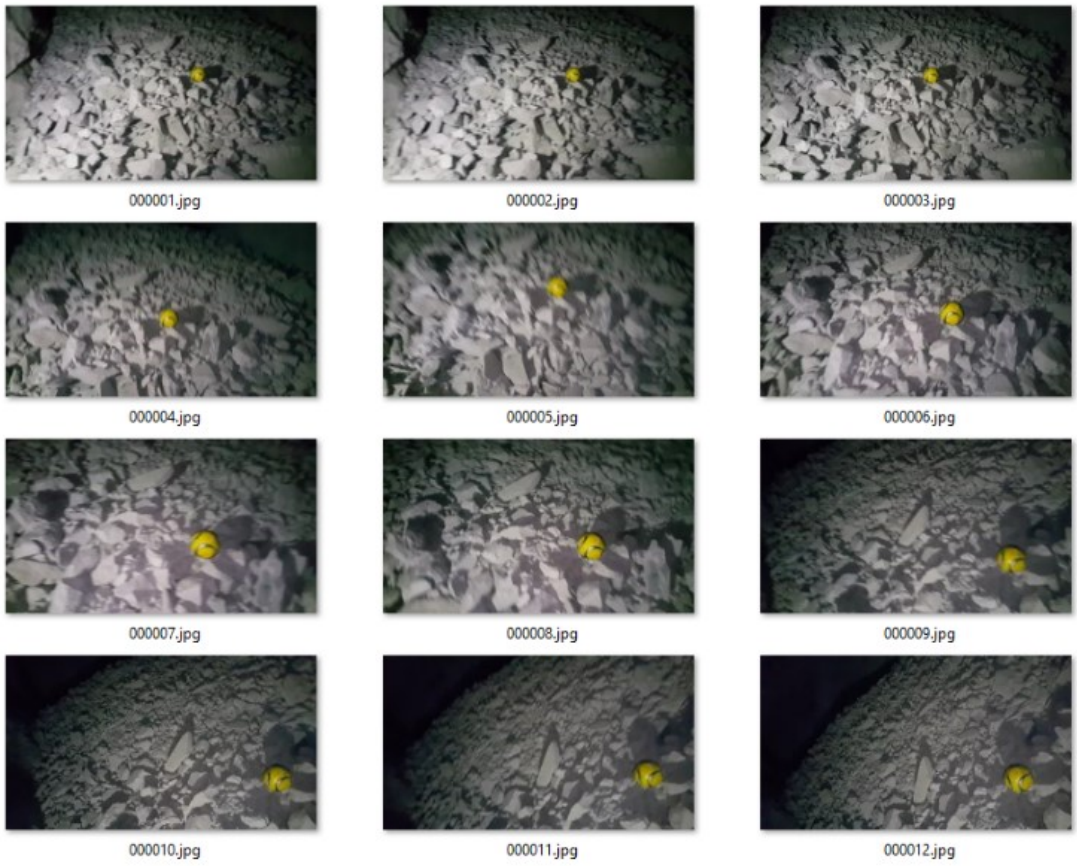

Figure 15. A sample set of images used in the reconstruction

3.3.1 OpenDroneMap (ODM)

ODM is an open-source, Linux-based (written on Ubuntu) SfM-MVS program that was mainly developed to process aerial imagery to create a variety of georeferenced outputs such as maps and 3d models (OpenDroneMap Authors, 2020). At its core, it uses OpenSfM, an open-source SfM library (OpenSfM Authors, 2020). Despite being intended to be used for aerial imagery, it still has sufficient capability to create muckpile 3d models.

ODM utilizes a command-line based reconstruction workflow housed in a Docker container. As such, the user can manipulate the command input to tweak the process by taking in and out certain steps, inputting auxiliary data such as georeferenced information, outputting intermediate results, and more. A basic illustration of the

workflow can be seen in Figure 16, and reconstruction results can be seen in Figure 17 and Figure 18.

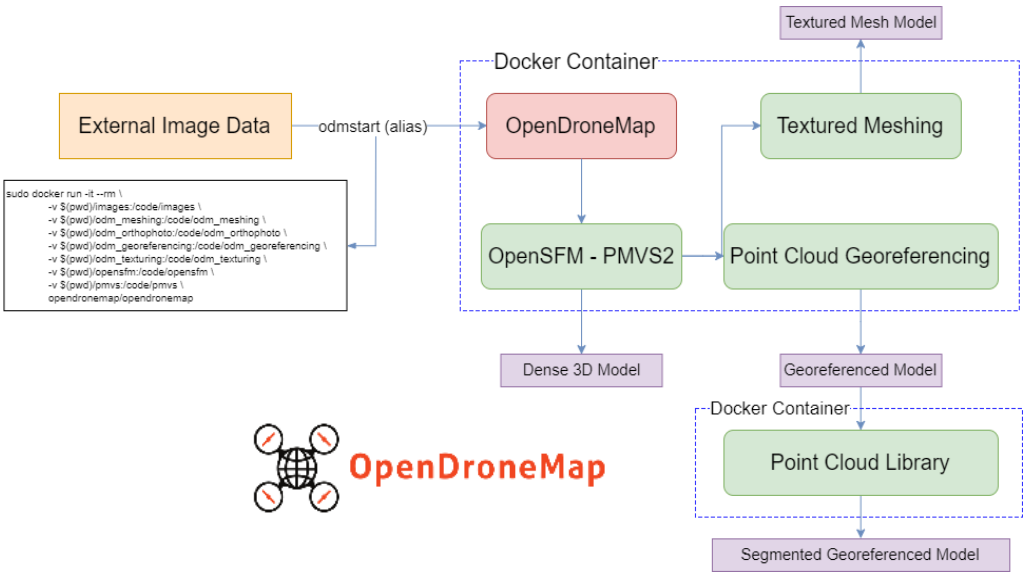

Figure 16. Workflow utilizing OpenDroneMap to create 3d models

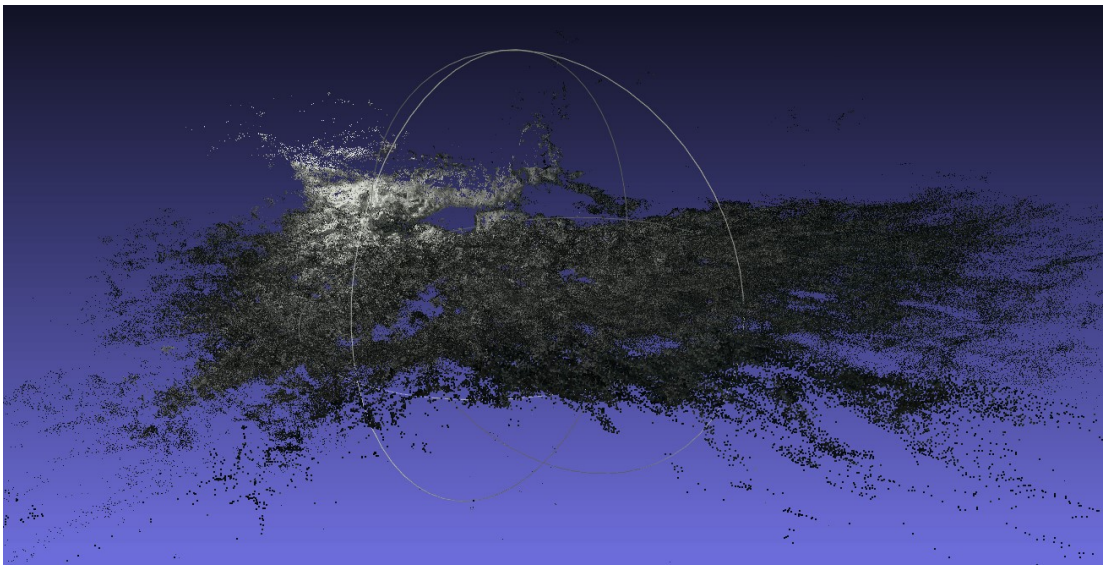

Figure 17. Dense 3d point cloud from OpenDroneMap

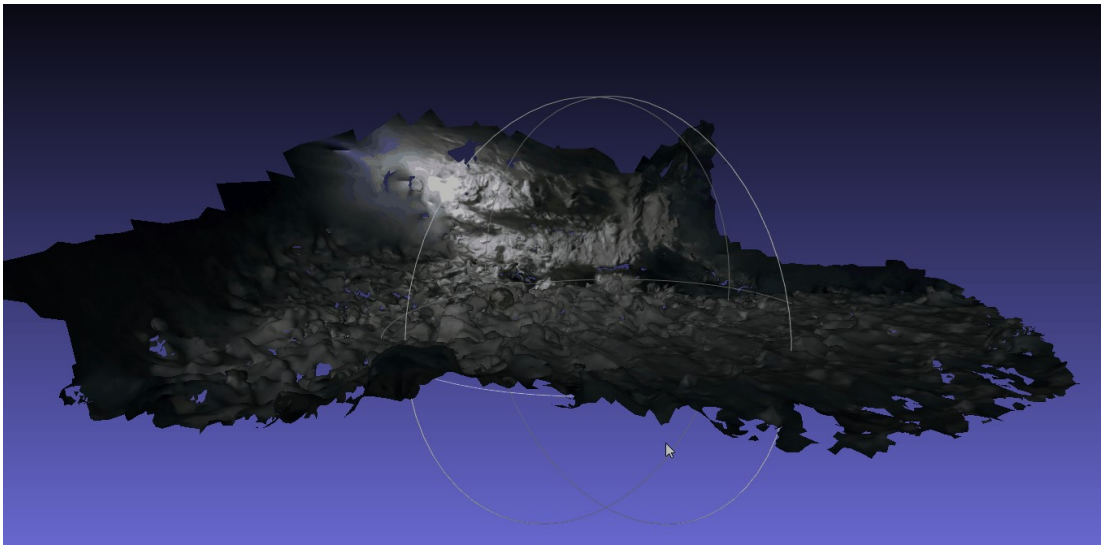

Figure 18. Textured model of the underground muckpile, output from OpenDroneMap

3.3.2 AliceVision Meshroom

Meshroom is another open-sourced 3D reconstruction software, based on its own AliceVision framework (Griwodz et al., 2021). It is available for use in both Windows and Linux operating systems, and features a modular reconstruction pipeline that can be fine-tuned by the user as is appropriate for the output that is needed. The general workflow of Meshroom is described in Figure 19, and reconstruction results are found in Figure 20 and Figure 21.

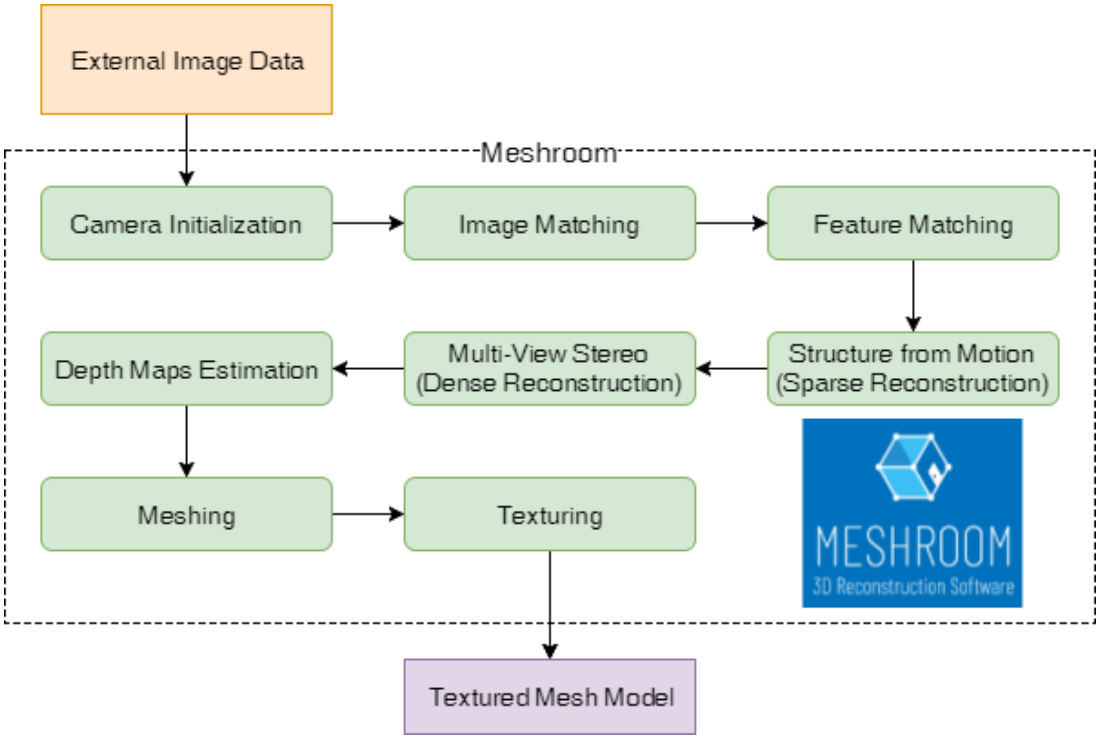

Figure 19. Workflow utilizing Meshroom to create 3d models

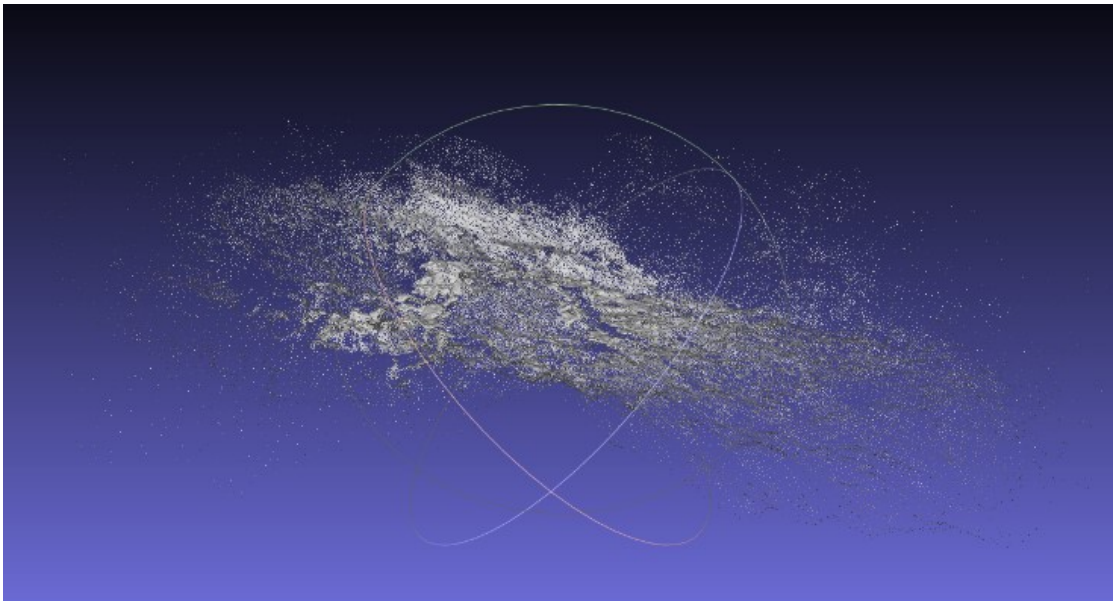

Figure 20. Dense reconstruction (uncolored) from Meshroom

Figure 21. Textured Model output from Meshroom

3.3.3 Reconstruction Analysis

Comparing the final textured output from the two workflows, output from the Meshroom produced a model with a greater number of vertices (points) and faces at 172297 points and 344496 faces, while the OpenDroneMap reconstruction generated 100108 vertices and 143361 faces. However, upon visual inspection of the final textured models, ODM's output is subjectively more accurate than Meshroom's, judging by the reconstruction of the soccer ball and the general form of fragments.

Figure 22. Close-up of the Textured Mesh generated by Meshroom. Note that the soccer ball is almost unidentifiable and the presence of floating 'blobs'

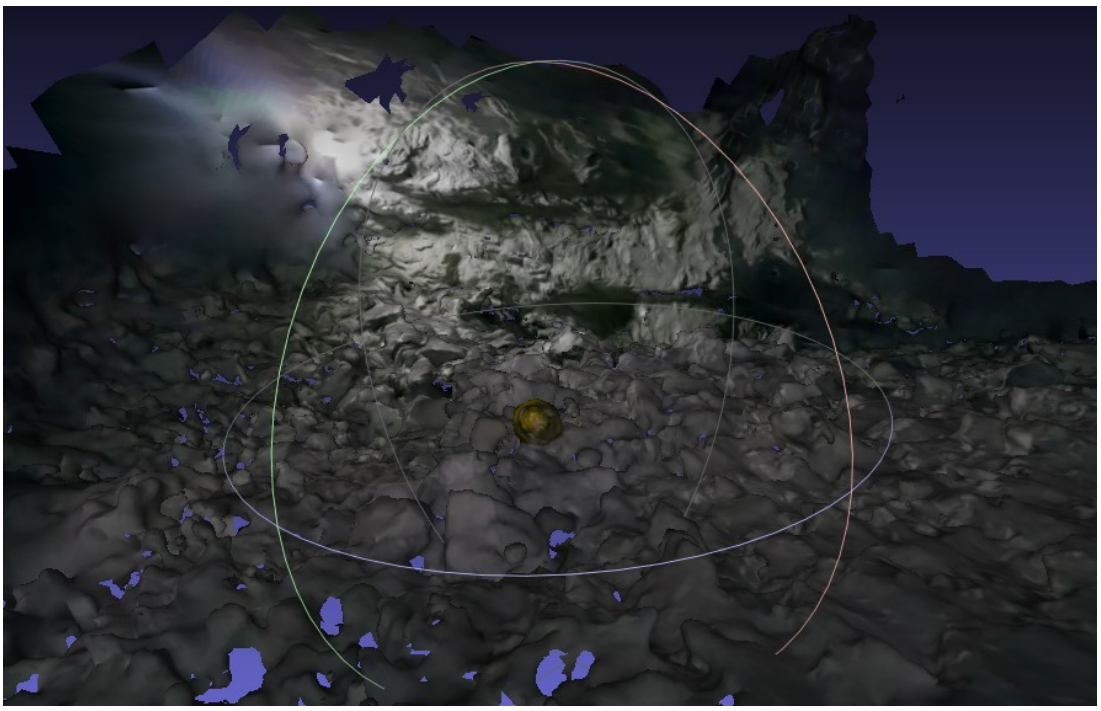

Figure 23. Close-up of the Textured Mesh generated by ODM. The soccer ball has relatively been well reconstructed and no floating 'blobs' are present

Suffice to say, this suggests that the number of vertices and faces generated isn't an appropriate metric for determining the accuracy of a reconstruction.

In addition, upon using MeshLab's measuring tool to measure the soccer ball's diameter in the scene. It has been observed that ODM's reconstruction is fairly close to real scale, with the ball, being a regulation FIFA ball with a computed average circumference of 22cm (The International Football Association Board, 2020). Measuring the ball in the scene shows a rough average diameter of 21cm. It should be noted that due to the nature of the reconstruction some difficulty was present in using the measuring tool.

3.3.4 Limitations in the Reconstruction Process

Difficulties in properly reconstructing the scene was present in both workflows, requiring multiple runs and tweaking to achieve what was shown previously. These issues are possibly due to a variety of reasons. One of these reasons could be video quality in the form of lighting, as video was taken using a smartphone camera in a low-light environment and the fact that a large percentage of the images extracted from the video contains a significant amount of blur. Additionally, there is a substantial amount of tweaking that can be done when using the reconstruction software, requiring some finesse that will come as more reconstructions are performed. Structure-from-Motion was first and foremost a tool for use in computer generated imagery (CGI) and animation and could be considered an art.

3.4 GNSS-aided Scaling in Photogrammetry

The study proposes a method that makes use of GNSS (Global Navigation Satellite System) data to create scaled 3D models without the need for post-reconstruction rescaling. GNSS positional data and its sub-systems such as GPS,

Beidou, GLONASS, and Japan's own QZSS can be utilized. A previous study has been with regards to using GPS in reconstruction, but mostly in the context of UAV (Unmanned Aerial Vehicle) Mapping (Fonstad et al., 2013). This study aims to create a system that doesn't need ground truth data such as GCPs to create a properly scaled 3D model of a muckpile. This would aid greatly in the fragmentation size distribution measurement of muckpiles using photogrammetry.

It is a known fact that an inherent error exists within GNSS and its subsets, and even high-end geodetic GNSS receivers have errors in the centimeter range (Khomsin et al., 2019). For this study, a smartphone is used as a GNSS receiver for the digital camera. This decision is due to the end-goal of this research which is to be able use both image data and GNSS data from a smartphone, as this practicality can be important in a mining operation environment. This comes at a drawback to the GNSS accuracy, as recreational grade GNSS chips like those found in smartphones typically have errors in the meter range (Merry & Bettinger, 2019). To overcome this error, the study proposes to make use of an increasing number of georeferenced images to statistically decrease the scaling error of the constructed 3D model. Figure 24 shows a general overview of the proposed system for this study. Utilizing a smartphone's built-in GNSS receiver, GNSS data can be logged and sent to a camera. At the moment an image is taken, GNSS data can be embedded into the image's metadata (EXIF).

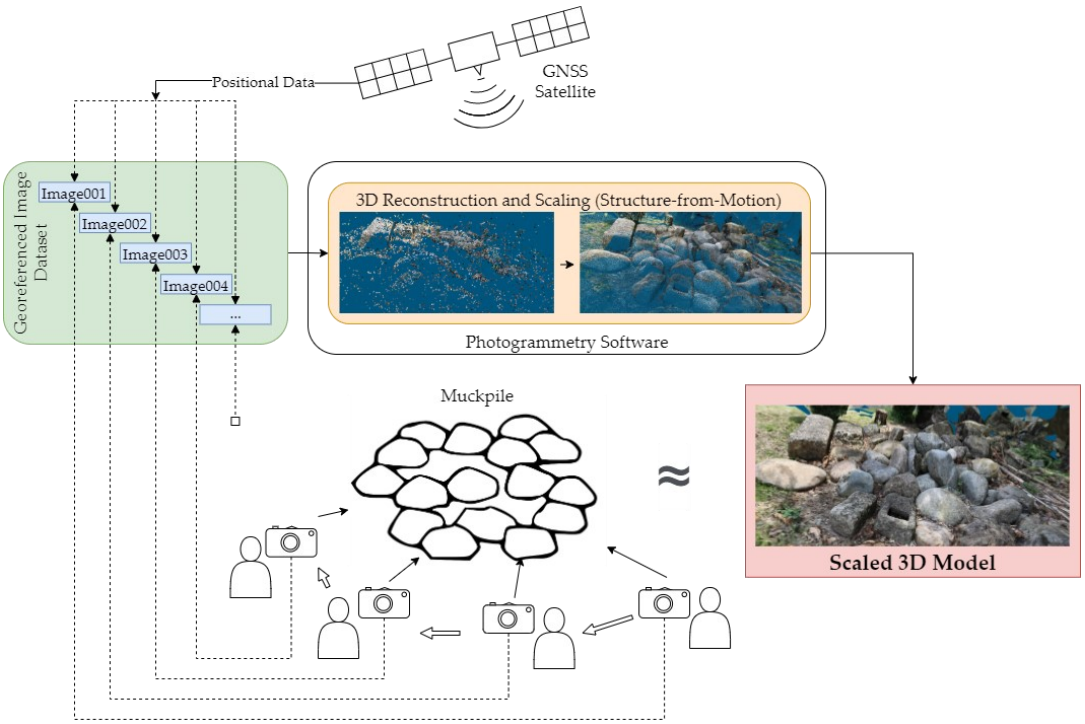

Figure 24. GNSS-aided photogrammetry workflow.

3.4.1 GNSS-aided scaling in Bundle Adjustment

In the bundle adjustment phase of SfM, the previous and imperfect solutions regarding camera positions and 3D features of the scene are refined (Zhang et al., 2006). More specifically, bundle adjustment is a non-linear minimization procedure that jointly optimizes camera parameters and point position by minimizing the reprojection error between the image locations of observed and predicted image points. This minimization is done using nonlinear least-squares algorithms (Lourakis & Argyros, 2009). Numerous studies have been done since its inception in the 90's regarding bundle adjustment, with most of the research going into reducing its computational burden and accelerating the problem-solving process (Triggs et al., 1999). One of such propositions is the fusion of positional data and bundle adjustment, with GNSS data being used as constraints for solving reprojection errors (Kume et al., 2010). This concept is what this research aims to produce accurately-scaled 3D reconstructions of

muckpiles. GNSS data is used to provide position and covariance estimates for the bundle adjustment process. The nominal form of these solutions is:

$$r_{GNSS}^{M}(t) = r_{c}^{M}(t) + R_{c}^{M}(t) r_{GNSS}^{c} + \left(b_{GNSS}^{M} + d_{GNSS}^{M}(t - t_0)\right), \quad (5)$$

where $r_{GNSS}^{M}(t)$ is the position of the GNSS receiver, $r_{c}^{M}(t)$ denotes the camera position, and $R_{c}^{M}(t)$ is the rotational matrix that aligns camera and mapping space axes, and r_{GNSS}^{c} is the difference between the GNSS receiver and camera position. b_{GNSS}^{M} and d_{GNSS}^{M} denote bias and drift terms and are included to account for data inconsistencies and the inherent errors that exist within GNSS. A previous study (Jaud et al., 2020) applied GNSS-assisted terrestrial photogrammetry to model coastal areas without the use of GCPs. With bundle adjustment being an error minimization problem with multiple factors, weights can be assigned to them, as is the case in the study's SfM workflow. The software that was used for this study was developed with the aim of being able to perform SfM without the need for any additional intrinsic or extrinsic data (such as GNSS) aside from the image themselves. (Toldo, 2013). However, the software itself still allows for the importation of GNSS data from images for the purpose of constraining camera positions. Weights are assigned to this GNSS data and is used in the bundle adjustment step (3Dflow, 2019). As such, the study deems it necessary to initially prove if properly constraining camera positions will help in creating a properly scaled 3D model. A preliminary experiment was designed to test this theory, which is described in the proceeding section.

3.4.2 Preliminary experiment for validating GNSS-aided scaling fundamentals

A preliminary photogrammetry experiment was performed before the main experiment to test some core concepts regarding the study, specifically the effects of

known and constrained camera positions on scaling error and reconstruction quality. This small-scale experiment involved taking photos of a scene that was set-up indoors in the laboratory that consisted of a stuffed dog plushie that was placed on the floor in such a way that it was in the middle of a grid of 9 carpet panels. The panels are 50cm by 50cm in size and form a 3 x 3 grid measuring 150cm by 150cm in total. Figure 25 shows the general layout of the scene. The purpose of this grid is to provide a spatial reference for the camera positions when taking pictures of the scene. A detailed board was put on the middle of the scene to provide enough feature points for SfM, as initial reconstructions without the board resulted in distorted point clouds with missing parts.

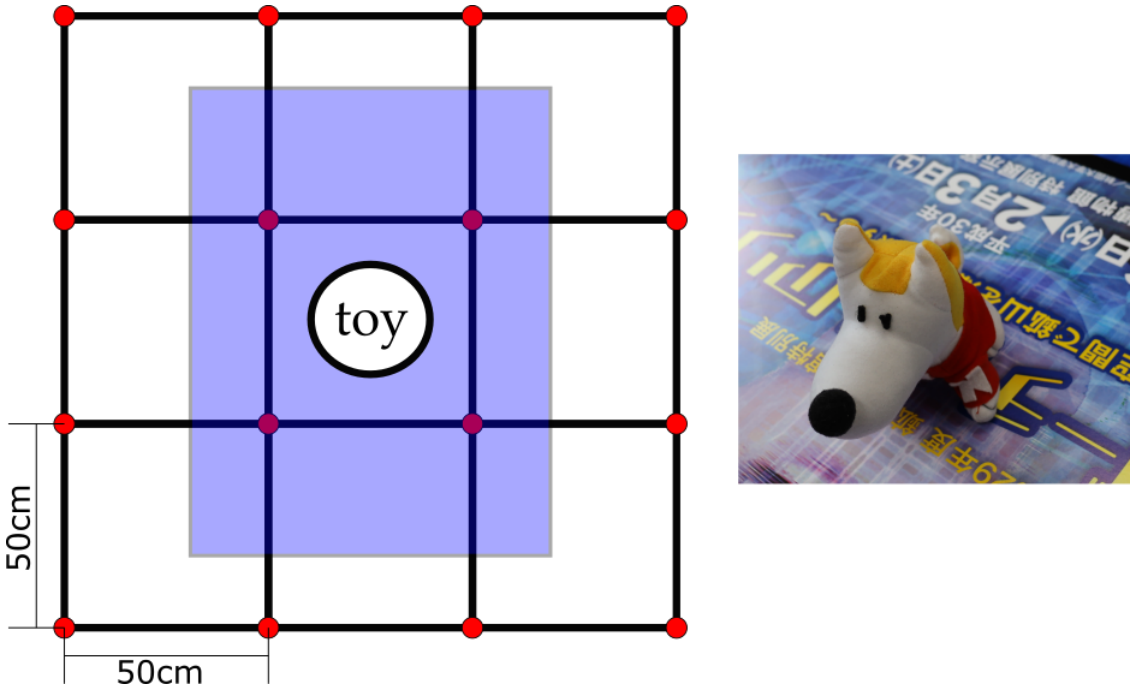

Figure 25. Grid layout of the 1st preliminary experiment (left) and a toy as the object (right). Red dots indicate camera positions and the blue rectangle indicates the board that was inserted for improved feature detection.

The camera used for this preliminary and succeeding experiments is a Canon EOS R equipped with a Canon 24-105mm lens. The f-stop is set at 4 with variable exposure times, automatic white balancing enabled, and the zoom was kept at a minimum to provide a fixed focal length which is required for SfM. A total of 32 photos

were taken, two at each of the intersection points of the grid at different heights- 45 and 60cm, with sample images shown in Figure 26. The height was maintained by mounting the camera on an adjustable tripod. Along with the 50cm spacing, this provided known relative camera positions. The captured images were then processed, with a workflow that consisted of making the sparse point cloud and a dense point cloud using photogrammetry software. Creating a textured mesh was deemed unnecessary as it meant a longer processing time and larger project file size and ultimately did not contribute to analyzing the results of the experiment. After which, the camera positions were constrained to their known locations. The scaling error of the reconstructed model was then analyzed.

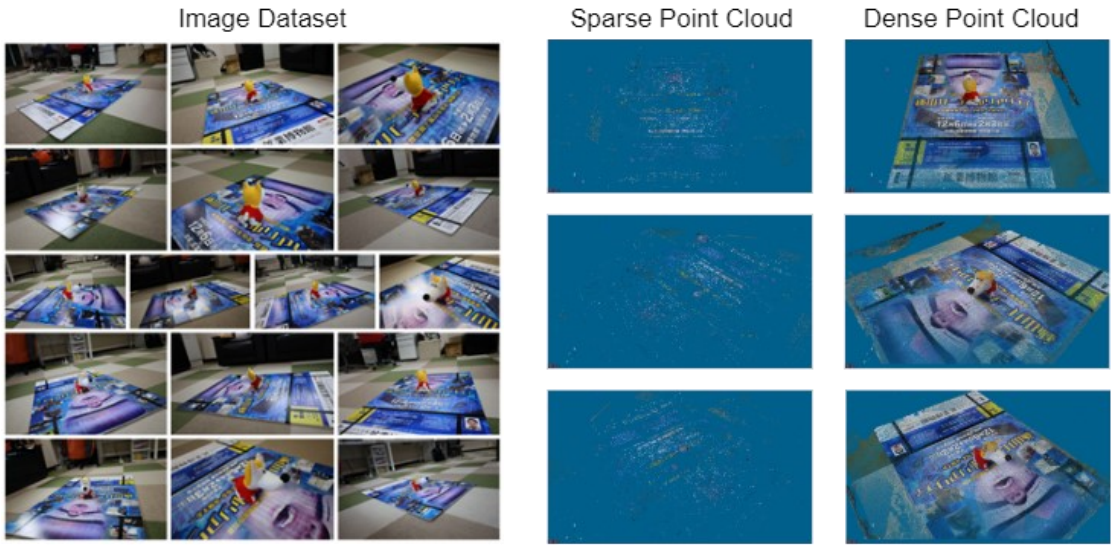

Figure 26. Data input and output of the preliminary experiment.

Without any camera constraints, the software arbitrarily designated a scale, rotation and translation for the model, and the measurement of the dimension of the carpet panel measured at about 10 units (since there are no constraints applied, this value cannot be assigned a specific unit of measurement). However, upon adding constraints to the camera (at the centimeter level) by importing a file describing each image's distance from each other, the same dimension now measures at 51.2719cm,

with a difference of 1.2719cm from the real measurement of 50cm. This difference was attributed to human error during the shooting process. A possible specific example is that the center of the tripod (and by extension the center of the camera) was what was used to align the camera to the grid, instead of the nodal point of the camera lens. This means that the images were offset from the actual intended grid position depending on the orientation of the camera and the tripod. Despite this difference, the study still proves fundamentally that accurately constraining the camera positions to their real-world values improves the scale accuracy of the constructed 3D model.

4 Experiments on GNSS-aided Photogrammetry

4.1 GNSS-constrained SfM on monuments of known dimensions

To perform quantitative evaluation of the effects of GNSS constraints on the scaling error of 3D reconstruction using SfM, an analysis using monuments of known dimensions outside Akita University was done. The experiment aims to correlate the scaling error to the number of images used in SfM. The hypothesis of this experiment is that as more images are used, the scaling error due to GNSS error will decrease. In this scene, the cube-shaped monument has sides measuring approximately 1 meter. This dimension is used to compute scaling error. This particular scene was chosen for this reason, in addition to the monuments being of simple 3D shapes, as demonstrated in Figure 28, making analysis of measurements more accurate for the purpose of quantitative evaluation.

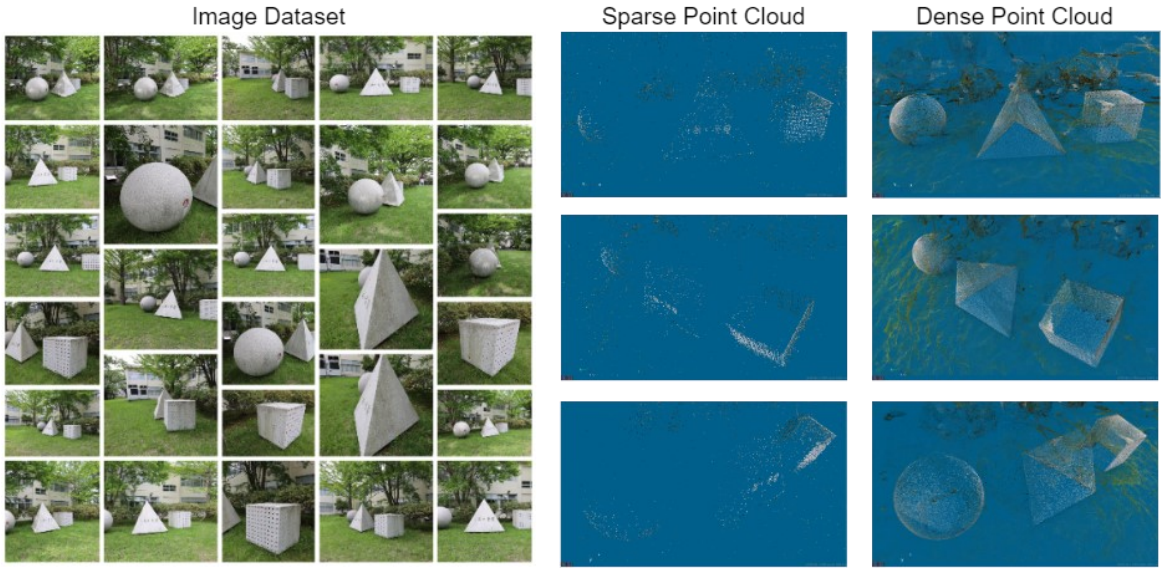

Figure 27. Data input and output of the experiment using 200 images.

Figure 28. Mesh reconstructed 3D CG model of the monuments. 3rd picture (bottom) shows measured side of cube monument when using 100 images.

For this experiment, around 200 images of the scene were taken across 2 days at roughly the same time of the day, with sample images shown in Figure 27 and a map of the depicted photo taking area and the recorded camera positions in Figure 29. For this and the proceeding experiment, the camera was used freehanded without a tripod, with a Xiaomi Mi 9T Pro smartphone placed close to the camera sending GNSS data to it via Bluetooth. The dataset, as with the previous experiment, was used to create 3D reconstructions at different image numbers. The scaling error when varying number of images are used was noted and compared. For reconstruction purposes in this and the

following experiments, 3DF Zephyr was used, with a setting of 50% GNSS data weight, as specified in the software's manual (3Dflow, 2019).

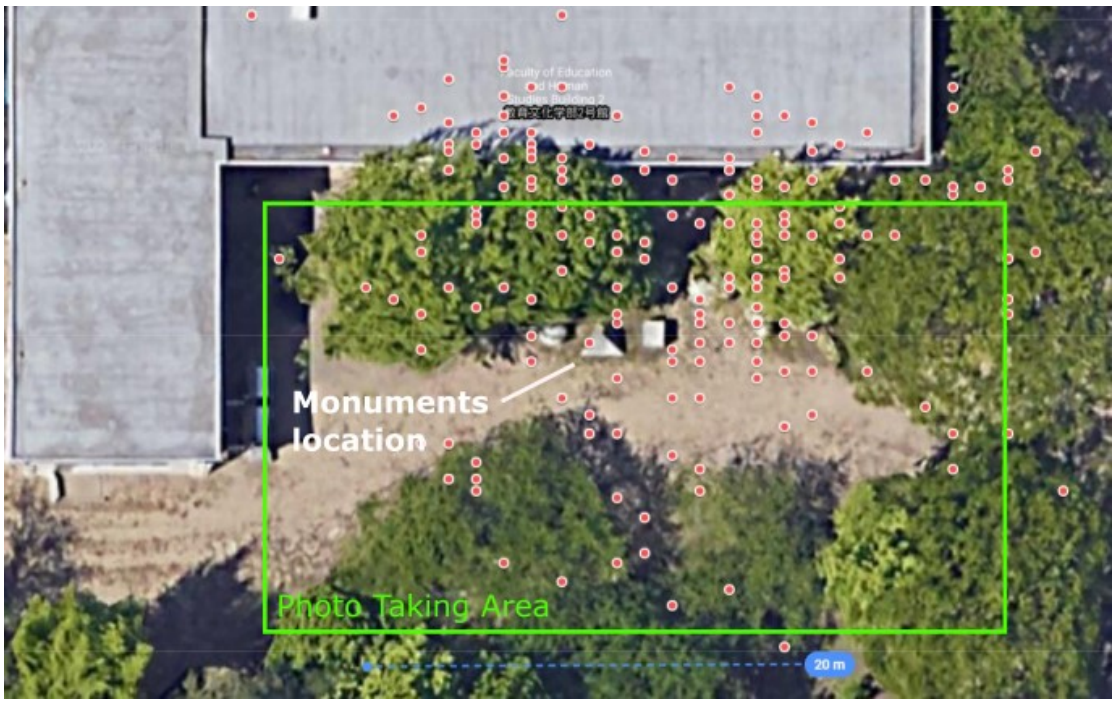

Figure 29. GNSS location of camera positions as logged by a smartphone (red dots) of the experiment on monuments of known dimensions.

Table 1. Results of the Experiment on Monuments of Known Dimensions.

Data	Measured (m)	Real Measurement (m)	Difference from Real Measurement (m)
50 images	2.00	1	1.00
100 images	1.97	1	0.97
150 images	1.74	1	0.74
200 images	1.61	1	0.61

As shown in Table 1 and Figure 30, there is a trend that at increasing number of images used in reconstruction, the difference from the real measurement decreases. This increase in accuracy lends credence to the hypothesis that using more images for reconstruction has the tendency to lessen scale error in 3D models. Using the trendline of the data, a model with a difference from real measurement of 0.1m (10% scaling error) can be hypothetically created if 386 (385.93) images are used.

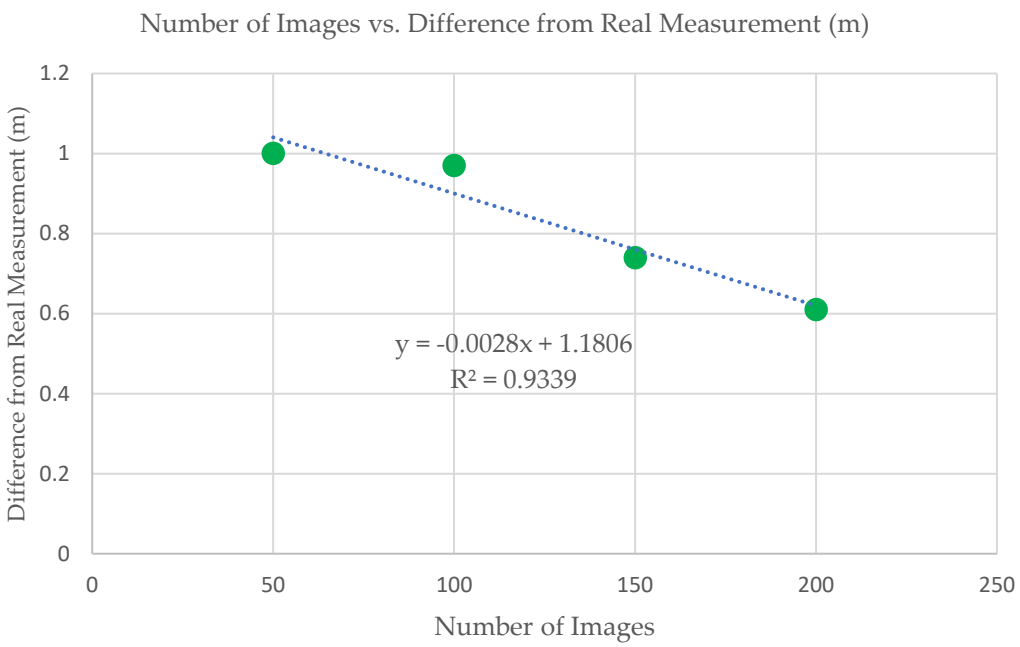

Figure 30. Graph detailing the results of the experiment.

4.2 Experiment on a pseudo-muckpile

For this test, the goal was to recreate a scene of a collection of boulder-sized rocks found at a temple site near the university, shown in Figure 31. The aim of this case study is to provide both quantitative and qualitative evaluation of the effects of GNSS constraints on the scaling error of 3D reconstruction with a subject that is a close simulation of an actual muckpile in a mining environment. The study conducted an experiment using a rock pile located near Akita University. These rocks are similar in

size and shape to a muckpile, and if the effectiveness of the method on this dataset can be confirmed, it can be assumed that the method will be equally effective on an actual muckpile in a mine site.

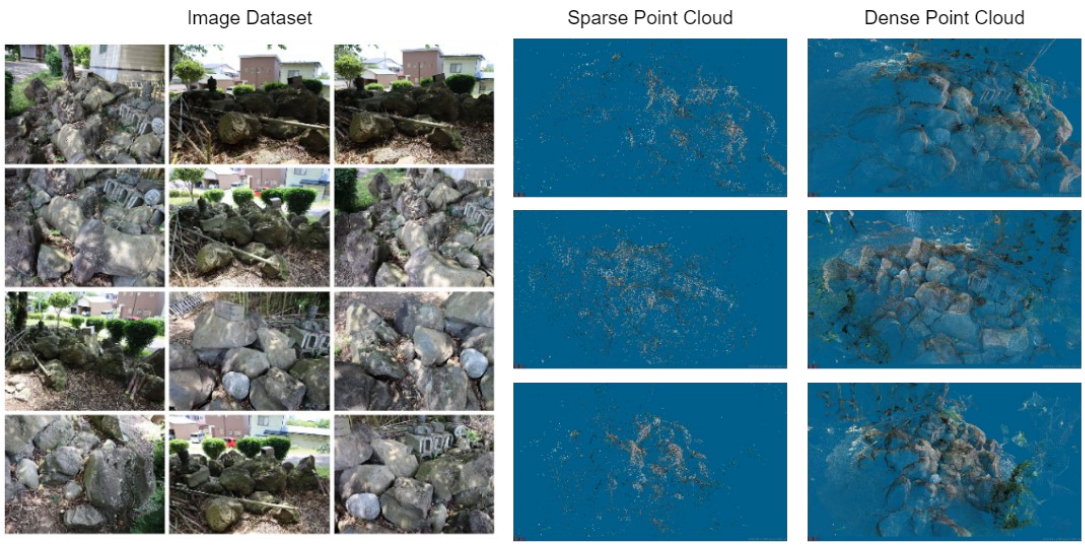

Figure 31. Data input (Set #1) and output of the experiment using 100 images.

Figure 32. GNSS location of camera positions as logged by a smartphone (red dots) of the experiment on a pseudo-muckpile.

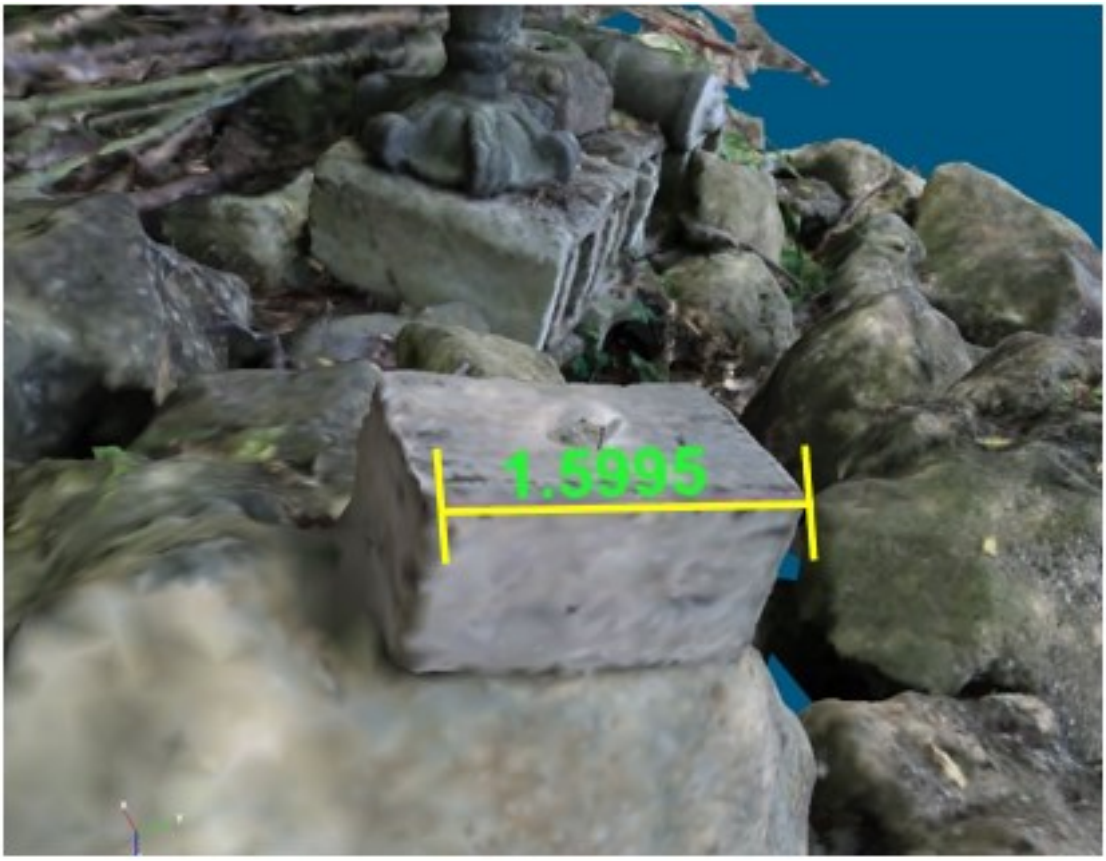

Figure 33. Meshed 3D CG reconstruction of wooden box reference with measurement (at 100 images used).

A total of 200 photos were taken, split into two datasets (Set #1 and #2) as shown in Figure 31 and Figure 34, with a map depicting the photo taking area and the recorded camera positions found in Figure 32. The rockpile was divided into two parts, one with bigger, angular rocks and another with smaller, rounded rocks. Both piles are around 4m wide on their longest side and are less than a meter long. A wooden box measuring 30 x 30 x 17 cm was placed in the scene for reference, as shown in its reconstructed form in Figure 33. In addition, measurement of the big, rectangular prism-shaped rock with dimensions of 35 x 40 x 30 cm were taken for reference as well, which can be seen in its reconstructed form in Figure 35. After the photos were taken, they were once more processed to produce several 3D models at different image numbers. The scaling error and the reconstruction quality is then observed in a similar fashion to

the previous experiments. Since two sets of data were used for this experiment, scaling error between using 50 images (chosen at random) and 100 images for each set are used. An additional exploratory test using 200 images using both sets are added for testing. The study's initial hypothesis, however, was that this will introduce some reconstruction errors as there are not enough images that are similar between these two scenes.

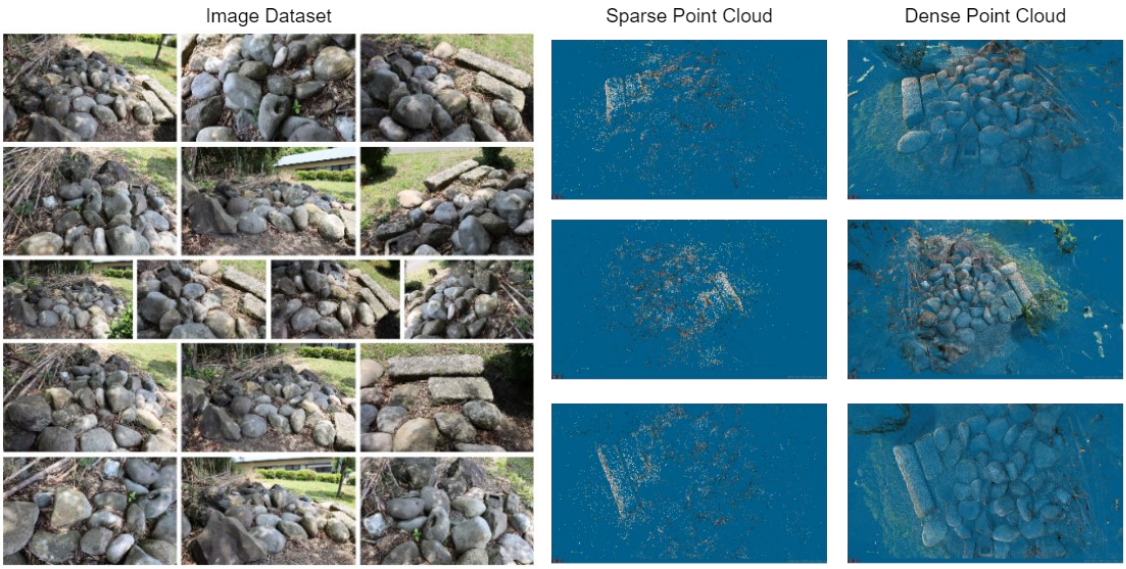

Figure 34. Data input (Set #2) and output of the experiment using 100 images.

Figure 35. Close up of Set #2 meshed 3D CG reconstruction, with measurements on long rectangular prism-shaped rock that was used as reference.

Table 2. Results of the experiment on a pseudo-muckpile

Data Set and Image Count	Measured (m)	Real Measurement (m)	Difference from Real Measurement (m)
Set #1 (50 images)	2.98	0.30	2.68
Set #1 (100 images)	1.60	0.30	1.30
Set #2 (50 images)	6.11	1.40	4.71
Set #2 (100 images)	6.09	1.40	4.69
Combined Set (200 images)	0.16	0.17	0.01

The measurement comparison is shown in Table 2. For Set #1, at 50 images used, the difference from the real measurement of the width of the box (0.3m) is 2.6m. At 100 images used, the difference is 1.3m. This leads to a decrease in 1.3m in the scaling error when using 50 more images. For Set #2, at 50 images used, the difference from the real measurement of the width of the rectangular rock (1.4) is 4.8m. At 100 images used, the difference is 4.6m. This leads to a decrease in 0.2m in the scaling error when using 50 more images. For a final, investigational set of using 200 images combining both previous sets, a surprising result has been observed – even though there were significantly more reconstruction errors (missing parts, duplicating parts, etc.) in this particular reconstruction, the wooden box width in this reconstruction was measured at 0.167m, with a difference of 0.01m from the real measurement. Reconstruction is shown below in Figure 36. The study considers that this increase in accuracy can be attributed to not just the number of images increasing, but also the general area of the scene becoming larger as it includes both pseudo-muckpiles (the effect of model size on GNSS error is discussed in a latter part of this section). However, combining the datasets also means that the scene being reconstructed is contextually different as it now includes both parts of the pseudo-muckpile.

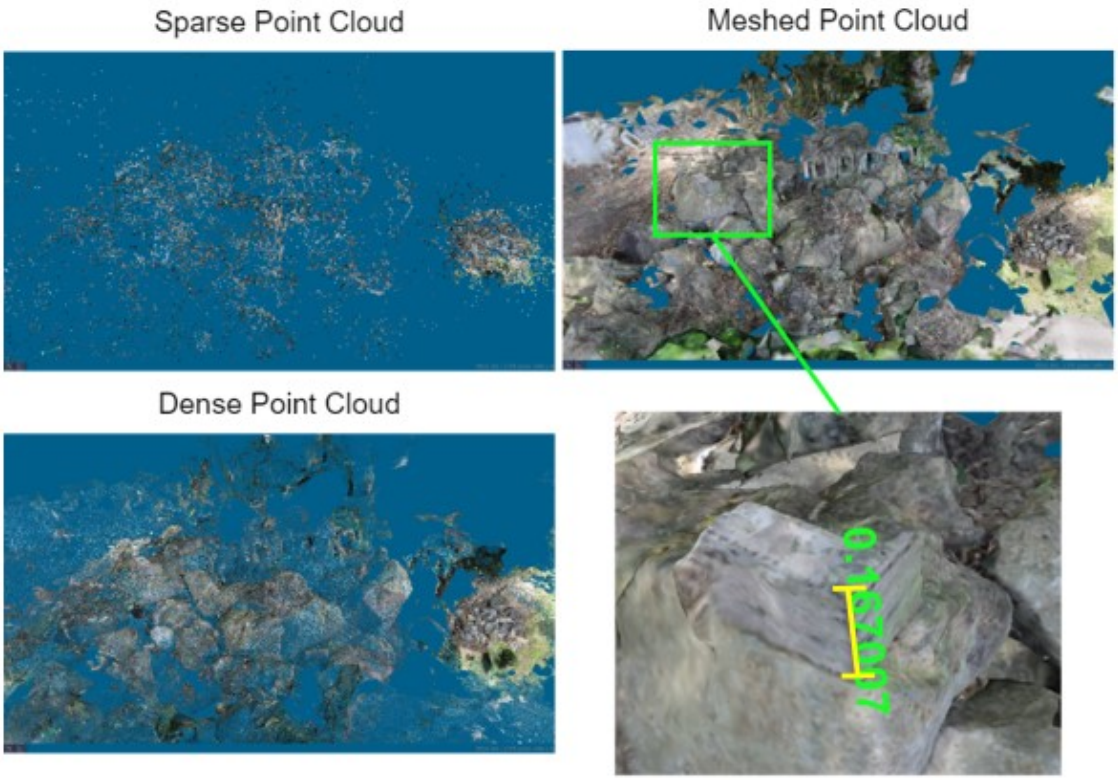

Figure 36. 3D CG reconstruction results of the combined data sets.

It can be seen from both experiments through the maps in Figure 29 and Figure 32 the apparent GNSS drift that occurs during the photo taking. Some of the recorded camera positions are either outside the photo taking area or in spots that are obstructed. The study recognizes that these changing boundary conditions have an effect on the results and a separate investigation on this can provide insight for GNSS-aided photogrammetry. Aside from the inaccuracies found in GNSS, several additional factors have been considered to contribute to the drift. One of such is the effect of the partial tree cover in some of the camera positions. A previous study (Uzodinma & Nwafor, 2018) in a similar setting (university campus) that analyzed the effect of not only partial tree cover, but also nearby infrastructure on GNSS accuracy by comparing GNSS data to total station survey data. Results show that some points were no longer suitable for GNSS positioning due to high GDOP (geometric dilution of precision) and where it was suitable, the GNSS recorded position differed by as much as 5.7m from

the total station data. This difference is consistent with what transpired in this study's experiments, as can be seen from the maps. In a mining site, where there is usually less vegetation and obstruction, this effect should be diminished, however.

Another factor that can be considered is the overall scale of the pseudo-muckpile. A large majority of GNSS-aided photogrammetry applications are usually in the form of aerial imagery and mapping, with a scope and scale larger than both of the terrestrial photogrammetry experiments performed in this study. A study (Fleming Zachariah & Pavlis Terry, 2018) investigating the application of terrestrial photogrammetry in field geology by using SfM-MVS aided by GPS to model an outcrop that long observed scaling and rotational errors in their reconstruction. Aside from concluding that GNSS contributed highly to these model errors, they suggested that at a larger scale, the error would be less of an issue. In parallel to this, the study observes that the relatively small scale of the experiment area affects the data; particularly the pseudo-muckpile whose size is smaller than a muckpile that one would normally find in a mining operation. Ultimately however, the results show that even at this scale, incremental improvements to 3D model scaling have been made as shown in the data.

4.3 Experiments on a Muckpile in Mikurahana Quarry Site

For this experiment, an actual muckpile found in an active mining (quarry) site is recreated in 3D space. The aim of this experiment is to assess the performance of GNSS-aided photogrammetry in a practical application in an open mining environment. The site chosen for this experiment is the Mikurahana Quarry Site, found in the town of Hachirogata in Akita Prefecture, Japan. The muckpile is composed mainly of limestone and fragments are generally large, with fragments as large as 2 meters wide

in its largest dimension. Satellite imagery is shown in Figure 37 and Figure 38 to provide locational context for the quarry and the muckpile in question.

Figure 37. Satellite view of the Mikurahana quarry site; the red box indicates location of the muckpile and its photographs.

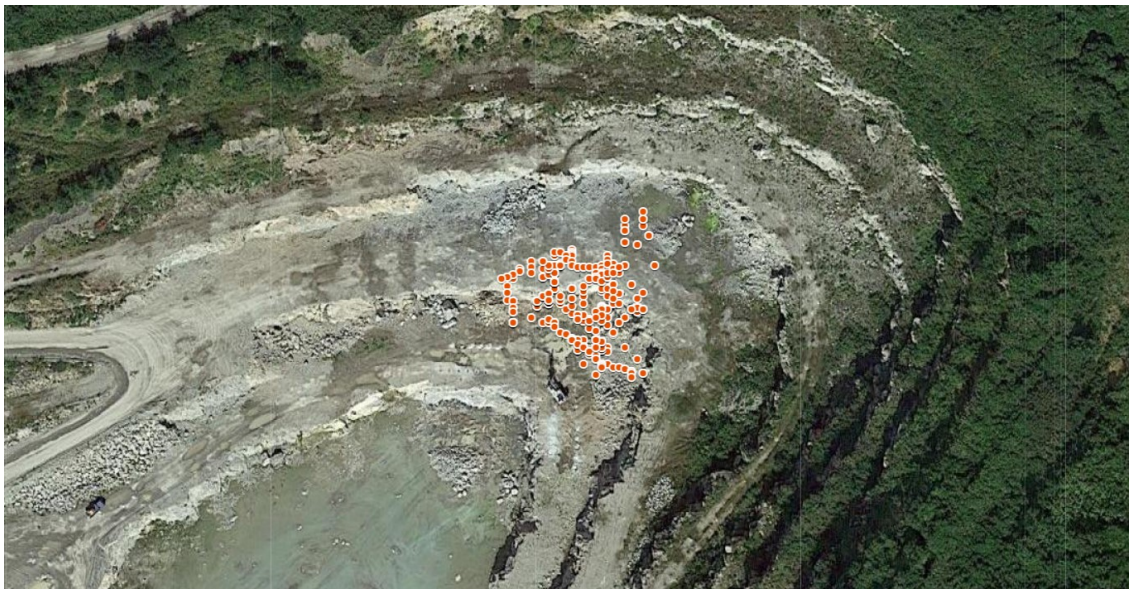

Figure 38. Close-up of quarry and muckpile photograph locations (marked as red dots)

A total of 400 images of the muckpile were taken, taken from varying angles facing the muckpile using the same camera-smartphone GNSS set-up from the two previous experiments. Sample images are shown in Figure 39, which also feature the yellow cardboard box that was used as an absolute reference for scaling error. The box measures 15 x 15 x 60 cm, with the 60 cm-long side used for measuring scaling error.

Figure 39. Sample images from the dataset. The yellow box is used as an absolute reference.

Using 3DF Zephyr, the images are used to create 3D models at different image numbers, namely 50, 100, 200 and 400 images. The scaling error of each model with respect to the yellow box is measured and plotted in a graph to observe the relationship between scaling error and image number. The sparse and dense reconstruction of the muckpile at 400 images used is shown in Figure 40 at different angles for reference. A mesh reconstruction featuring measurement of the yellow box in 3DF zephyr is also shown in Figure 41.

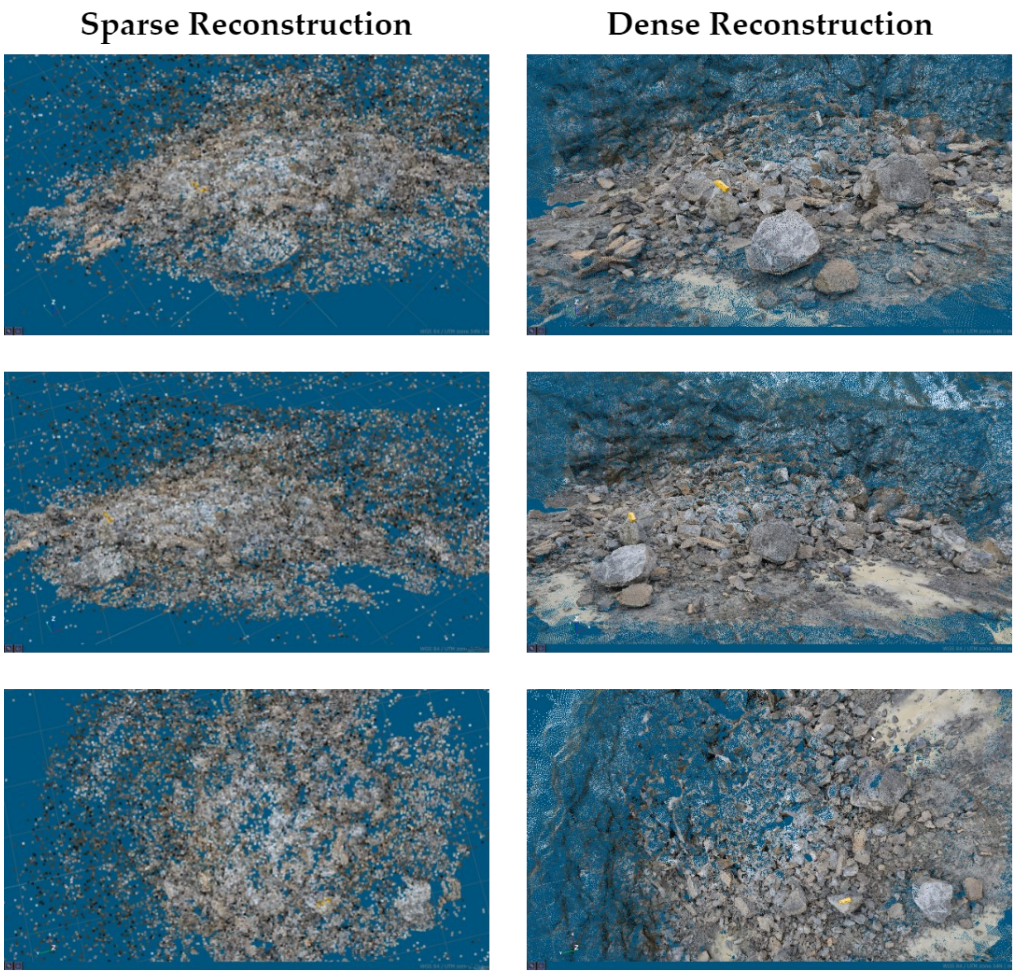

Figure 40. Sparse and dense reconstruction of limestone muckpile using 400 images, shown at different angles.

Figure 41. Close-up of mesh reconstruction, featuring reference yellow box and dimension of its longest side as measured.

Table 3. Results of the experiment on the Mikurahana quarry muckpile

Number of Images	Measured (m)	Real Measurement (m)	Difference from Real Measurement (m)	Percent Error (%)
50	0.82	0.6	0.22	36.67
100	0.76	0.6	0.16	26.67
200	0.74	0.6	0.14	23.33
400	0.67	0.6	0.07	11.67

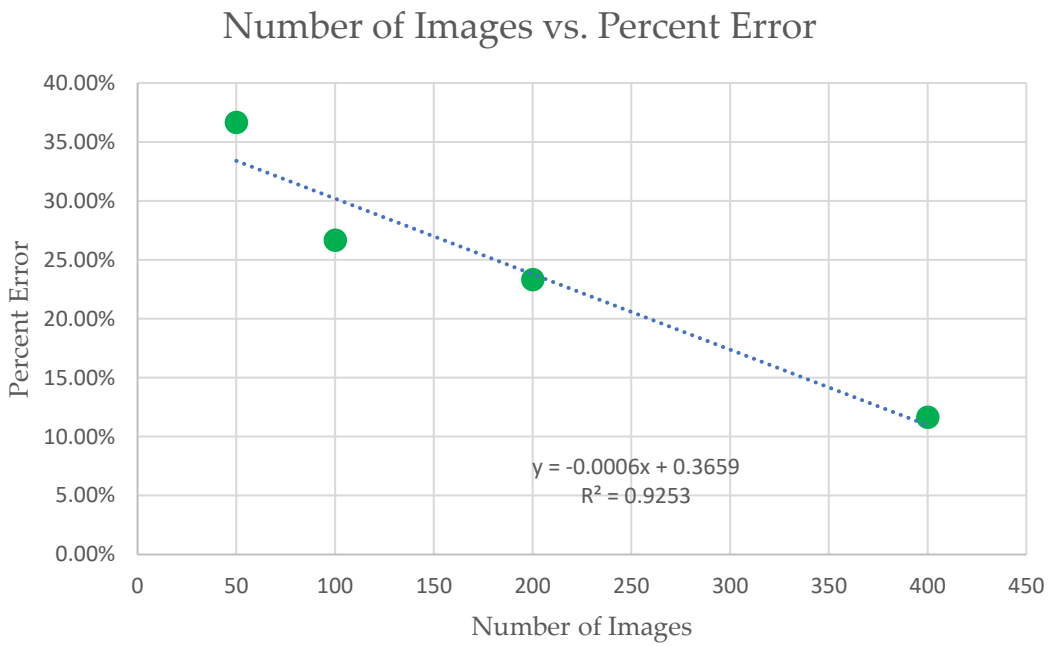

Figure 42. Graph detailing the results of the Mikurahana experiment.

As shown in Table 1Table 1 and Figure 1, there is a trend that at increasing number of images used in reconstruction, the difference from the real measurement decreases. This result, as with the previous experiments, lends more credence to the hypothesis that using more images for reconstruction has the tendency to lessen scale error in 3D models, with a relatively linear relationship between image number and

scaling error as shown by the trendline with an R-value of 0.93. It is noted, however, that the improvement in scaling error by increasing the number of images is lessened as the scaling error at using only 50 images is already 36.67%, which is much higher than when using 50 images in the other experiments. The study hypothesizes that this is due to the lack of vegetation in the area, which improved the GNSS accuracy. The study achieved 11.67% scaling error at 400 images, which is close to the extrapolated 10% scaling error at 400 images in the previous experiment on monuments.

5 Conclusion

In this study, the study proposes a method of creating an accurately scaled 3D model by constraining camera positions through the use of georeferenced images as input for SfM. Monitoring fragmentation size is an important procedure in optimizing mining operations that perform blasting. In recent years, a new method that involves using 3D photogrammetry to measure fragment sizes has been developed that has the potential to surpass traditional techniques. For this particular process to be accurate, a method for properly scaling 3D model with georeferenced images using GNSS is investigated. To validate the method, several experiments were performed. As an initial test to prove fundamentals, an indoor scene involving a small object was recreated in 3D space using SfM with photos of known relative positions for constraining camera location, and good results that show that the created 3D model has a scaling error of 1.27cm were achieved. For the next experiment, the study took georeferenced photos of an outdoor scene with a monument of known dimensions and made several reconstructions at increasing number of images used (50, 100, 150 and 200 images respectively). The results show a linear pattern with an R-squared value of 0.93 in which the scaling error decreases as the number of images used increases. Additionally, an experiment was done to verify the study's hypothesis further, using a scene that includes a pseudo-muckpile to simulate the usage of the proposed system for a mining operation. In a similar fashion, the results show increasing scale accuracy with increasing number of images used in reconstructions. Finally, an experiment was performed on an actual muckpile in Mikurahana quarry to test the system's accuracy in a practical application. 3D reconstructions were created at image numbers of 50, 100, 200 and 400 of a limestone muckpile and the scaling error was measured and graphed against the image number. It also showed a linear pattern with an R-squared value of

0.93 in which the scaling error decreases with increasing image number, albeit at a lower ratio that has been hypothesized to be due to the lack of interference from vegetation and buildings. Two observations can be drawn from the experimental results:

1) constraining cameras to accurate positions in SfM will result in a properly scaled 3D model and
2) increasing the number of georeferenced images in SfM will incrementally improve the scaling error of the reconstruction. These observations can help improve scale accuracy in GNSS-aided 3D fragmentation measurement.

These results lend credence to the possibility of improving the scaling aspect of 3D fragmentation measurement systems without the use of GCP or manual scales, specifically in surface mines where GNSS data is generally readily available. This shows that monitoring the fragmentation distribution can potentially be performed using just a camera and a GNSS-enabled devices, such as smartphones.

6 References

3Dflow. (2019). *3DF Zephyr Manual*. http://3dflow.net/zephyr-doc/3DF%20Zephyr%20Manual%204.500%20English.pdf

Afum, B., & Temeng, V. (2014). Reducing Drill and Blast Cost through Blast Optimisation - A Case Study. *3rd UMaT Biennial International Mining and Mineral Conference*, 137–145. https://www.researchgate.net/publication/269165866

Aler, J., du Mouza, J., & Arnould, M. (1996). Measurement of the fragmentation efficiency of rock mass blasting and its mining applications. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics*, *33*(2), 125–139. https://doi.org/10.1016/0148-9062(95)00054-2

Bergmann, O. R., Riggle, J. W., & Wu, F. C. (1973). Model rock blasting—effect of explosives properties and other variables on blasting results. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *10*(6), 585–612. https://doi.org/10.1016/0148-9062(73)90007-7

BME South Africa. (2016, May 4). *Better blasting can save on power costs*. https://bme.co.za/better-blasting-can-save-on-power-costs/

Carrivick, J., Smith, M., & Quincey, D. (2016). *Structure from motion in the geosciences*. Wiley, Blackwell.

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. In V. Scarano, R. de Chiara, & U. Erra (Eds.), *Eurographics Italian Chapter Conference*. The Eurographics Association.

https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136

Cipolla, R., & Robertson, D. P. (2009). Chapter 13: Structure From Motion. *Practical Image Processing and Computer Vision: Wiley*.

Cunningham, C. V. B. (2005). The Kuz-Ram fragmentation model–20 years on. *Brighton Conference Proceedings*, *4*, 201–210.

Fleming Zachariah, D., & Pavlis Terry, L. (2018). An orientation based correction method for SfM-MVS point clouds—Implications for field geology. *Journal of Structural Geology*, *113*, 76–89. https://doi.org/10.1016/J.JSG.2018.05.014

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L., & Carbonneau, P. E. (2013). Topographic structure from motion: A new development in photogrammetric measurement. In *Earth Surface Processes and Landforms* (Vol. 38, Issue 4, pp. 421–430). https://doi.org/10.1002/esp.3366

Grant, J. R., & Dutton, A. J. (1983). Development of a fragmentation monitoring system for evaluating open slope blast performance at Mount Isa Mines. *Proceedings of the 1st International Symposium on Rock Fragmentation by Blasting, Lulea, Sweden*, 637–652.

Gräter, J., Schwarze, T., & Lauer, M. (2015). Robust scale estimation for monocular visual odometry using structure from motion and vanishing points. *2015 IEEE Intelligent Vehicles Symposium (IV)*, 475–480.

Griwodz, C., Gasparini, S., Calvet, L., Gurdjos, P., Castan, F., Maujean, B., de Lillo, G., & Lanthony, Y. (2021). *AliceVision Meshroom*. 241–247. https://doi.org/10.1145/3458305.3478443

Jang, H., Kitahara, I., Kawamura, Y., Endo, Y., Topal, E., Degawa, R., & Mazara, S. (2020). Development of 3D rock fragmentation measurement system using photogrammetry. *International Journal of Mining, Reclamation and Environment*, *34*(4), 294–305. https://doi.org/10.1080/17480930.2019.1585597

Jaud, M., Bertin, S., Beauverger, M., Augereau, E., & Delacourt, C. (2020). RTK GNSS-Assisted Terrestrial SfM Photogrammetry without GCP: Application to Coastal Morphodynamics Monitoring. *Remote Sensing 2020, Vol. 12, Page 1889*, *12*(11), 1889. https://doi.org/10.3390/RS12111889

Kanchibotla, S., Valery, W., & Morrell, S. (1999). *Modeling fines in blast fragmentation and its impact on crushing and grinding*.

Khomsin, Mutiara Anjasmara, I., Guruh Pratomo, D., & Ristanto, W. (2019). Accuracy Analysis of GNSS (GPS, GLONASS and BEIDOU) Obsevation for Positioning. *E3S Web of Conferences*, *94*. https://doi.org/10.1051/e3sconf/20199401019

Kume, H., Taketomi, T., Sato, T., & Yokoya, N. (2010). Extrinsic camera parameter estimation using video images and GPS considering GPS positioning accuracy. *Proceedings - International Conference on Pattern Recognition*, 3923–3926. https://doi.org/10.1109/ICPR.2010.954

Liu, Q., & Tran, H. (2018). Comparing systems—validation of Fragscan, Wipfrag and Split. In *Measurement of Blast Fragmentation* (pp. 151–155). Routledge.

Lourakis, M. I. A., & Argyros, A. A. (2009). SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software*, *36*(1), 1–30. https://doi.org/10.1145/1486525.1486527

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, *60*(2), 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Madali, N. (2020, July 30). *Structure from Motion*. https://towardsdatascience.com/structure-from-motion-311c0cb50e8d

Merry, K., & Bettinger, P. (2019). Smartphone GPS accuracy study in an urban environment. *PLoS ONE*, *14*(7), e0219890. https://doi.org/10.1371/journal.pone.0219890

Nefis, M., & Talhi, K. (2016). A model study to measure fragmentation by blasting. *Mining Science*, *23*, 91–104. https://doi.org/10.5277/msc162308

OpenDroneMap Authors. (2020). *Open Drone Map - A command line toolkit to generate maps, point clouds, 3D models and DEMs from drone, balloon or kite images*. https://github.com/OpenDroneMap/ODM

OpenSfM Authors. (2020). *OpenSfM*. https://github.com/mapillary/OpenSfM

Palangio, T. C., Franklin, J. A., & Maerz, N. H. (1995, January 1). WipFrag -- A Breakthrough in Fragmentation Measurement. *6th High-Tech Seminar on State of the Art Blasting Technology, Instrumentation, and Explosives Applications*.

Schönberger, J. L., & Frahm, J. (2016a). Structure-from-Motion Revisited. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Schönberger, J. L., & Frahm, J. (2016b). Structure-from-Motion Revisited. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

The International Football Association Board. (2020). *Laws of the Game*. www.theifab.com

Toldo, R. (2013). *Towards automatic acquisition of high-level 3D models from images*.

Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W. (1999). Bundle adjustment—a modern synthesis. *International Workshop on Vision Algorithms*, 298–372.

Tungol, Z. P. L. (2019). *A Study on the Use of 3D Photogrammetry Software for Rock Fragmentation Size Distribution Measurement*.

Uzodinma, V. N., & Nwafor, U. (2018). Degradation of GNSS Accuracy by Multipath and Tree Canopy Distortions in a School Environment. *Asian Journal of Applied Sciences*, *6*(4). https://doi.org/10.24203/AJAS.V6I4.5458

Valery, W., Morrell, S., Kojovic, T., Kanchibotla, S., & Thornton, D. (2001). *Modelling and Simulation Techniques Applied for Optimisation of Mine to Mill Operations and Case Studies*.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300–314. https://doi.org/10.1016/J.GEOMORPH.2012.08.021

Wipware Inc. (2017, January 13). *WipWare ® Sampling and Analysis Guide*. https://wipware.com/wp-content/uploads/2018/01/Sampling-and-Analysis-Guide.pdf.

Zhang, J., Boutin, M., & Aliaga, D. G. (2006). Robust bundle adjustment for structure from motion. *Proceedings - International Conference on Image Processing, ICIP*, 2185–2188. https://doi.org/10.1109/ICIP.2006.312973