[DC] Exploring Virtual Photogrammetry Techniques and Applications For Advancement of Digital Twin Generation

Jacob Rubinstein*
University of Maryland, Baltimore County

Abstract

We explore challenges and opportunities in the use of photogrammetry for digital twin generation, namely we aim to improve our understanding of what makes good input data for photogrammetry and to quantify the different traits of various photogrammetry processes. We propose the use of virtual photogrammetry - utilizing synthetic 2D images from pre-existing 3D models as input - to aid in this goal. Our approach aims to create a pipeline for generating datasets of synthetic images which can be used to evaluate and improve camera pose/ intrisics estimation as well as to assess the impact of errors on 3D reconstruction accuracy. By leveraging the advantages of this synthetic data, we aim to evaluate the resilience and accuracy of photogrammetry systems, leading to the higher quality results from non-virtual photogrammetry in the future.

Index Terms: Photogrammetry, Digital Twins, Reconstruction Computing Methodologies, Computer Vision, Artificial Intelligence

1 Introduction

Advancements in digital twin generation technologies, particularly photogrammetry, have greatly enhanced both the quantity and quality of 3D models available online. These approaches offer faster, more cost effective results when compared to traditional 3D modeling methods. However, despite the power of these tools, two main challenges remain 1) the cost of training data, and 2) the tuning of the models themselves.

When creating digital twins through photogrammetry, several critical factors must be optimized for the best final results, including sensor type, photo coverage and count, object surface texture, background choice, and the choice of capture technology. For those lacking expertise, it can be difficult to prioritize these elements effectively. Moreover, inaccuracies in the reconstruction process, particularly in camera pose estimation, can lead to cascading errors that harm the quality of the final model. Given that most reconstruction software lacks precise camera pose input, errors introduced during this intermediate stage are difficult to quantify but have a significant impact on the accuracy of the resulting 3D model.

In the domain of input image generation, prior work has explored the idea of attempting photogrammetric reconstruction from virtual images taken of pre-existing 3D models [1, 4, 6]. These past works explore the question of whether it is possible to reconstruct existing 3D, with each paper having their on reasons for doing so, but each were differently and more narrowly focused. We aim to adopt the workflow of capturing virtual photos of existing 3D models and adapt that to address the challenges inherent to 3D reconstruction.

*e-mail: jrubins1@umbc.edu

2 Advantage of Virtual Photogrammetry

Once a real object is captured as a 3D model, it is straightforward to generate many virtual photos of that 3D model with exact information on the virtual camera's relative position and orientation. Each of these virtual photos are exact representations of the relationship between the virtual camera and the 3D model, and we have ground truth raw data describing this relationship that is unavailable for regular photos taken of real-world objects.

This coupling of images to ground-truth camera parameters allows for the quantitative evaluation of existing heuristics in ways not otherwise possible, and also allows for the generation of data which can be used to train AI models for a variety of purposes at the interface between 2D and 3D representations.

Converting 2D content from ubiquitous, traditional cameras into 3D information has a wide range of applications which have become increasingly important in light of rapidly emerging technologies. VR headsets scan their physical environment for obstacles and to determine a user's position and orientation. Similarly, robots, drones, and self-driving vehicles use simultaneous location and mapping (SLAM) to navigate their environments. In photogrammetry, captured 3D models based on 2D imaging of real-world objects are used in medicine, civil engineering, disaster response, game development, cinematography, and cultural heritage preservation. Each of these applications has its own sensitivities to performance in latency, framerate, accuracy, and resolution. Nonetheless, implementations for all these application domains tend to use a shared foundation of longstanding heuristic algorithms for point-based visual feature detection and feature matching.

3 Research Aims

3.1 Research Aim 1: Create a Pipeline to Generate Semi-Synthetic Photos

To enable our other aims, as well as work by other researchers, we will create a pipeline for the generation of large set of semi-synthetic data in the form of simulated 2D "photos" rendered from photogrammetrically-captured 3D models. This pipeline will store each image's ground truth camera information (position, orientation, and intrinsics) as image metadata. Using an already-published set of 3D models, We will release a sample set of images (with embedded metadata), as well as the source code developed for generating this sample dataset.

3.2 Research Aim 2: Semi-Synthetic Images for Evaluation of Pose Determination

We will use the ground truth camera parameters from Research Aim 1 to evaluate the relative performance in camera pose determination by the existing metric of SE(3) distance functions [5] which calculate distance between two 6DoF poses.

3.3 Research Aim 3: Synthetic Ground Truth and Introduced Errors to Probe Sensitivity to Error in Photogrammetric Reconstruction

We will quantify the level of resilience in photogrammetry to artificially introduced errors in camera parameters. We will quantitatively connect the errors in various camera parameters (position, pose, intrinsics) to the resultant loss in quality of a 3D reconstruction, as measured by reprojection error and by image similarity measures.

4 Technical Approach

4.1 Technical Approach for Research Aim 1

Research Aim 1 requires the development of a pipeline for the generation of semi-synthetic images. This will be built on Blender using the built-in BPY python library for scripting. Camera locations will be generated as points on a Fibonacci Sphere [2] which distributes points semi-evenly with a low computation expense. Blender's camera perameters will then be tuned to match those of a specific camera, allowing for a wide variety of parameters to be accounted for with minimal effort.

As part of Research Aim 1, we will be using a publicly available dataset of 3D models to generate a sample output set. Our preference is to use the dataset from "A Real World Dataset for Multi-view 3D Reconstruction" [7], as this dataset was created as a large, high-resolution set of 3D captures of real world objects to support photogrammetry research. However, that dataset is currently not available at the URL published in the paper. The first author currently has the dataset available at a private link, and intends for the published URL to work again soon. We have a copy of the dataset, but if the original URL is not active by the time our work would begin, we will choose another, in the interest of reproducibility.

4.2 Technical Approach for Research Aim 2

In Research Aim 2, we compare the performance of existing methods for camera parameter determination against each other. There are several motivations behind this research aim. In SLAM, the entire goal is to accurately determine the position of a camera in order to track the movement of the camera (and usually an attached vehicle or robot) over time. The ground truth afforded us by Research Aim 1 allows for a new, quantitative approach to testing the accuracy of different approaches to this problem (SIFT, SURF, CNN, etc).

Another motivating factor is that photogrammetry software often shows the derived camera positions and some camera positions are often visibly wrong, to varying degrees. In the worst case, error in camera positions may result failed reconstructions or in messy reconstructions which need to be cleaned up by hand.

Because applications of pose estimation vary, and each application may indicate different choices of trade-offs, our evaluation will be similarly broad. Our ability to use a synthetic ground truth for quantification will allow us to benchmark performance against many parameters, such as the amount of data available for the camera parameter estimate and how such differences in input affect each of the six degrees of pose freedom and the variable number of internal degrees of freedom. The number of internal degrees of freedom which depending on the camera model being used, which may or may not include intrinsics like focal length, aperature, field-of-view, resolution, and geometric distortion.

4.3 Technical Approach for Research Aim 3

Unlike the other aims, which consider photogrammetry to be one of several application domains of interest, Research Aim 3 is focused entirely on photogrammetry. In particular, we wish to determine the extent to which photogrammetric reconstructions are resilient to errors in camera parameterization.

By leveraging the ground truth data from Research Aim 1 and introducing intentional errors ranging from very small to very large, we can quantify the extent to which small errors have any impact on the quality of the final model. If the effect of small errors is negligible, then we can conclude that choosing less accurate algorithms (e.g., because their patents have expired, or they were always open source, or they are fast) is non-consequential. On the other hand, if we find that even small introduced errors in camera non-negligible impacts on photogrammetric reconstruction, we can conclude that some amount of trouble (or cost) may yield a return on that investment.

To determine accuracy of a 3D reconstruction which has been given erroneous camera parameters, we will render images of these "bad" reconstructions while we also use the same virtual views to render matching images of the original 3D models. We will compare the "bad" image with these reference images through the various OpenCV tools meant for image-to-image comparison, with an initial focus (for simplicity) on mean square error as the "distance" between a pair of images. We will explore how this distance varies as a function of the extent of the introduced error. We will conduct this virtual experiment for all camera parameters (position, orientation, and basic intrinsics) to see if reconstruction is particularly sensitive to a particular type of error.

Our initial intent is to perform these photogrammetry reconstruction experiments by writing scripts which will make use of the command line interface for AliceVision, which serves as the backend of the open-source photogrammetry software Meshroom [3]. We anticipate possible difficulties in making this work, and have a backup plan of using our existing academic license for Reality Capture, a state-of-the-art commercial, closed-source 3D reconstruction suite. As with AliceVision, there is a command line interface available for Reality Capture, were performed using Reality Capture and its mobile companion app.

5 Conclusions

In this work, we present our plan to utilize virtual photogrammetry in various ways to allow for more informed data collection/ generation and model analysis beyond what is possible with current technologies. By completing this project, we hope to increase our top-to-bottom understanding of photogrammetric processes in a way which will enable the creation of cheaper, higher quality digital twins, and more robust photogrammetry pipelines to be developed in the future.

References

[1] H. Esmaeili and H. Thwaites. Virtual photogrammetry. In *2016 22nd International Conference on Virtual System Multimedia (VSMM)*, pp. 1–6, 2016. doi: 10.1109/VSMM.2016.7863153 1

[2] Á. González. Measurement of areas on a sphere using fibonacci and latitude–longitude lattices. *Mathematical geosciences*, 42:49–64, 2010. 2

[3] C. Griwodz, S. Gasparini, L. Calvet, P. Gurdjos, F. Castan, B. Maujean, G. D. Lillo, and Y. Lanthony. Alicevision Meshroom: An open-source 3D reconstruction pipeline. In *Proceedings of the 12th ACM Multimedia Systems Conference - MMSys '21*. ACM Press, 2021. doi: 10.1145/3458305.3478443 2

[4] G. Kortaberria, E. Gomez-Acedo, J. Molina, A. Tellaeche, and R. Minguez. Theoretical accuracy assessment of model-based photogrammetric approach for pose estimation of cylindrical elements. *Measurement Science and Technology*, 30(5):055003, 2019. 1

[5] F. C. Park and J. E. Bobrow. Efficient geometric algorithms for robot kinematic design. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, vol. 2, pp. 2132–2137. IEEE, 1995. 1

[6] E. Piatti and J. Lerma. A virtual simulator for photogrammetry. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5):4, 2006. 1

[7] R. Shrestha, S. Hu, M. Gou, Z. Liu, and P. Tan. A real world dataset for multi-view 3d reconstruction. In *European Conference on Computer Vision*, pp. 56–73. Springer, 2022. 2