Research Project Experience Report

Oscar Toomey, Supervised by Dr Ciarán Mc Goldrick

1 Introduction

This report details and reflects on my leadership development experience while working on the Research Project portion of the Laidlaw Leadership development programme. I will summarise my research, the challenges I faced, the skills I developed and reflect on how well my final project reflected the original goals I set out to achieve. Finally, I will detail what I learned about leadership and how my perspective on the subject has changed throughout the project.

2 Research Project Summary

One of the challenges faced by robots when trying to navigate and traverse unknown terrain is determining exactly where they are relative to their surroundings. *Structure-from-motion (SfM)* is the field of computer vision that attempts to solve this problem. The equivalent field in robotics is known as *Visual Simultaneous Localisation and Mapping (Visual SLAM)*. For my research project I set out to tackle a small subset of this problem: refining pose estimates using vision. This involves taking a pose estimate, captured using sensors; and then refining the pose using images captured from a camera.

My methodology had the following structure:

1. Prior to beginning the project, I identified a small area in the Structure-from-motion stack where I believed existing methods could be augmented to improve processing time.
2. The first task I completed once I started was to construct a test set of images and camera orientations with which to test any algorithms I developed in the project.
3. I formalised the problem statement that I was trying to solve.
4. I designed, developed, tested and analysed 4 algorithms to tackle this problem.

3 Algorithms

This section details the four algorithms designed and written for this project. The input to each algorithm is two image frames, an initial estimate for the frame positions and orientations, and a set of image features that have been matched between the two (Figure 2). These image features were extracted using the *Oriented FAST and Rotated BRIEF (ORB)* algorithm (Rublee, 2011) and then matched using the distance ratio test described in (Lowe, 2004).

The overall design is depicted in Figure 3, where for the purposes of testing the plant (or physical device environment) has been replaced with the test datasets generated for this project.

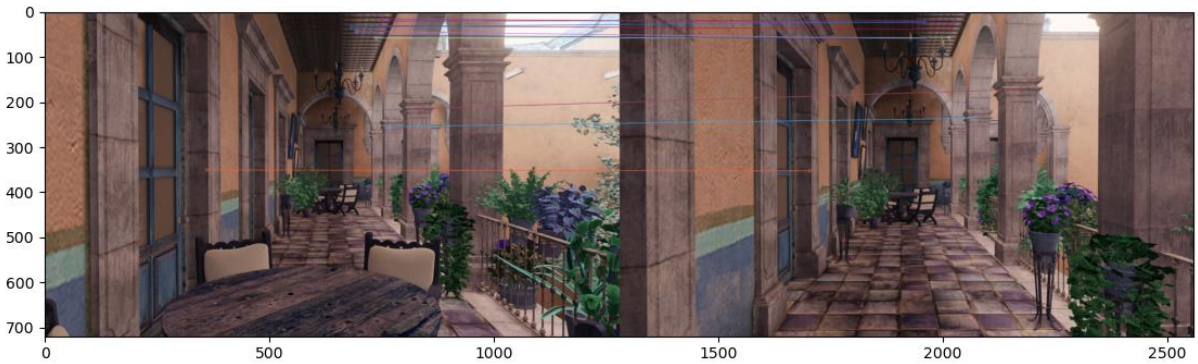

Figure 2: Features matched between two frames in the San Miguel dataset.

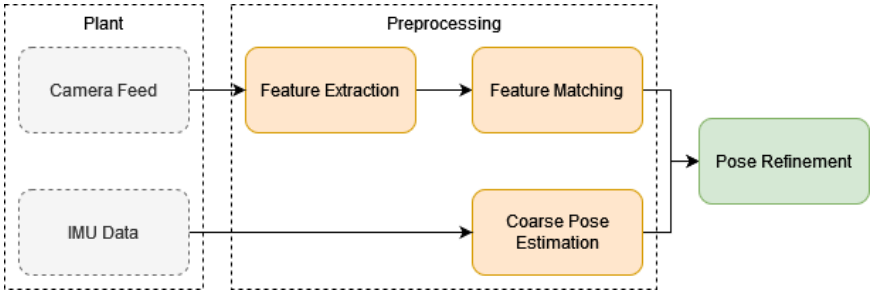

Figure 3: Flow diagram depicting how the algorithms in this work interact with the camera and inertial measurement data.

4 Results and Performance

The final algorithm developed in this project, adapted from *Chaumette, S 2006*, converges on low error solutions when the error in the angle and position is acceptably small, for angles less than 30 degrees and positions within 5 degrees relative to the feature positions in space:

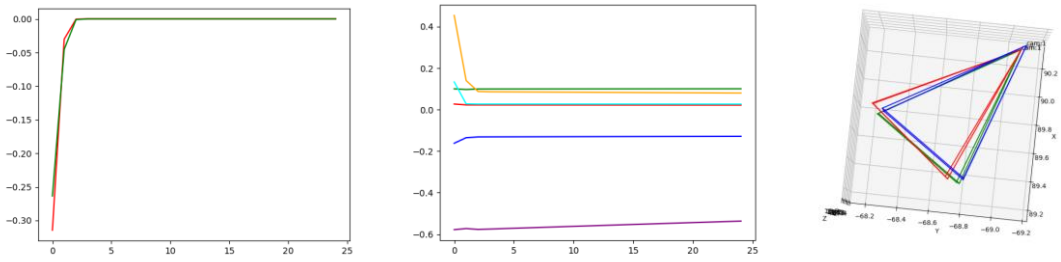

Figure 4: Left – Screen space conversion of the error in the X and Y axis to 0. Middle – Conversion of errors in the X, Y, Z and RX, RY, RZ axis. Right – Conversion of estimated pose (red) to final pose (green) and ground truth (blue)

As the second graph in Figure 4 shows, convergence in screen space does not necessarily mean that the camera has been reorientated to its true position. There are multiple reasons why this happens:

- Triangulation error: The depth of each point in the scene must be estimated and is less accurate the greater the initial error is between the camera's estimated position and actual position.
- Feature extraction error: The features matched by the feature matches often don't exactly line up between two image frames.
- Feature matching error: The feature matcher reports a match on two features that do not correspond to the same object. This can be seen for some of the matches in Figure 2.
- Quantization error: For features that are further away, the maximum precision they can triangulated to is limited by the size of the pixels on the image frame.
- Positions close to the true value often result in a small screen space error (Figure 5).

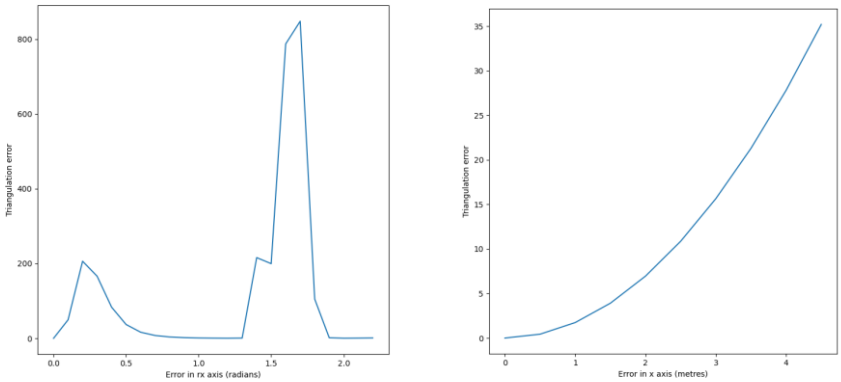

Figure 5: Triangulation error measured for 19 features resulting from error in rotation (left) and error in translation (right)

5 Findings

A key observation made during this work is that the performance of these pose recovery algorithms is entirely dependant on the quality of the features recovered by the feature matcher and extractor. Garbage in, garbage out: if the input features are poor, it is impossible to recover the pose accurately. Additionally, while pose recovery is a solved problem, feature extraction and matching has received a lot of attention recently with the resurgence of machine learning. Models like *LoFTR* (Sun, 2021) and *SuperGlue* (Sarlin, 2019) use deep learning to find hundreds of matches between image pairs compared to the two or three dozen found by the *ORB* matcher used in this work. Given more time, I would look at how initial pose estimates from sensor data can be used to increase the performance of these feature algorithms, rather than just in pose refinement.

6 Achievements & Challenges

In my view, my greatest achievement working on this project was persevering until I managed to build an algorithm that worked; a process that took five weeks and three failed attempts. This required dealing with a great deal of frustration and loss of motivation, and the result doesn't work as well as I had hoped, but I managed to get there in the end. I believe that I managed to demonstrate the ability to turn ideas in action as well as the capacity for resilience. I also count learning a huge amount about the field of study, as well as experience with computation frameworks and 3D linear algebra, as one of my personal achievements during this project.

On the technical side, a constant source of issues in this project was coordinate space conversion between different frameworks. *OpenGL*, *OpenCV* and *Pytorch3D*, the frameworks used in this project, all have subtle differences in how they treat translations and rotations. For instance, the camera Z and Y axis are flipped when converting between *OpenGL* and *OpenCV*, while *PyTorch3D* applies translations before rotations when constructing the extrinsic matrices for its cameras. It is therefore highly advisable to spend the time looking for the framework that best suits your needs and sticking to it; ideally never converting between different coordinate systems. A significant portion of time spent working on this project was burnt resolving these types of issues.

Additionally, it was difficult to establish a performance baseline with which to compare the algorithms developed in this work. This was primarily due to the lack of implementations of existing perspective and pose algorithms, such as Lambda Twist (Persson, 2018), in the Python programming language. Doing a fair comparison would have meant having to rewrite the algorithm in this work in C++, which was unfortunately outside the 6-week scope of this project.

7 Lessons Learned

The result of this work aligns with that which I set out to achieve in my proposal, albeit at a reduced scale. I had originally intended to build an end-to-end system in the 6 weeks, running on a drone or a phone, but it took far longer than I had anticipated to build a working algorithm. This is partly because I had no prior experience with SfM or SLAM systems.

I learned several things about working on a project by myself:

- It is extremely demotivating and frustrating when things consistently do not work. Working with others, who have fresh ideas to offer, can help greatly to minimize this.
- Having colleagues also helps during the initial brainstorming phase of a project. In my case, I became hyper-focused on one area too early and failed to properly identify the best issue to spend my time tackling.

- As I was nervous about hitting the deadline on time, I had dived in as fast as possible before identifying all the tools that I could use to reduce the amount of work I needed to do.

The content of my research project was not related to leadership, however I did learn just how much discipline is required to make a project successful, to constantly second guess your decisions and consult with others to determine the best path forward.

I interacted with my supervisor only a handful of times during my project. With hindsight, I would recommend to future scholars to have more meetings with their supervisor and experts in the field of study. Their knowledge and experience are invaluable to avoiding hidden pitfalls when exploring different ideas and implementing them. Those meetings also provide good checkpoints to work toward, rather than simply stumbling towards some end goal.

8 Leadership Development

The motivation for undertaking this research project was to personally develop better leadership skills and to gain insights into challenges amounting from the human element in large projects. While our leadership training sessions focussed on the leader-group dynamic, one of my biggest realisations working on the project element is that human capacity for motivation and engagement should be viewed as a critical project resource. In the same way that we fashion office environments that allow us to do work, we require spaces and channels that allow us to vent our frustration and regain some energy. As part of effective communication, a leader should always ensure that every member of the team understands the necessity and impact of their work, because when obstacles appear insurmountable, those are the type of questions we tend to ask ourselves. What can also help is knowing that other members of the group are in the same position, but this only works if everyone feels comfortable sharing their difficulties. From working in group projects in the past, many people feel insecure about asking for help, especially when they are a new member in the team. Something that might be worth exploring is designating someone in the team as the go-to problem person, or proactively organising a tutorial or help session with new members that allows them to interactively ask questions about aspects of a project or problem.

I also learned a great deal about myself as a researcher while working on this project. From my experience, I now believe that many leadership traits are also required to be a successful solo researcher. From being able to lead without authority, to being able to effectively communicate with experts to critical and creative thinking for solving problems at hand. However, this project also demonstrated to me that it is impossible to maintain all attributes of a leader all the time. Specially, being ever resilient and always determined to accomplish a goal is unsustainable and perhaps

unhealthy. It can be important to be able to take a step back and re-evaluate, to determine whether the current path really is the best one forward.

Before joining the Laidlaw programme, I realised that I did not have a good definition for what constitutes leadership, but through experience with this project I have found one that I am content with – leadership is specifically the ability and skills required to maximise human utility, both long and short term, for maximum positive impact. And I believe this project showed how important the process of managing human resources and channels is for achieving this positive impact; and that one cannot ignore one's own personal capacity for engagement and enthusiasm in that process. In this sense, the self-knowledge and awareness traits of leadership aren't limited to technical skills or domain knowledge. We should always challenge ourselves; life is too short not to and there is often little to gain in things that aren't difficult – but we must respect our limits.

10 References

1. Chaumette, S. Hutchinson, *Visual servo control, Part I: Basic approaches*, HAL Open Science, Institute of Electrical and Electronics Engineers, 2006
2. M. Leal Llaguno, McGuire, Cardenas, Mara, Hull, San Miguel 2.0, CC BY 3.0, available at https://casual-effects.com/data/ (last accessed September 2022)
3. Persson, Nordberg, *Lambda Twist: An Accurate Fast Robust Perspective Three Point (P3P) Solver*. European Conference on Computer Vision (ECCV) 2018
4. Rublee, Rabaud, Konolige, R. Bradski, *ORB: An efficient alternative to SIFT or SURF*. International Conference on Computer Vision (ICCV) 2011: 2564-2571.
5. Sarlin, DeTone, Malisiewicz, Rabinovich, *SuperGlue: Learning Feature Matching with Graph Neural Networks*, Conference on Computer Vision and Pattern Recognition (CVPR) 2019
6. Sun, Jiaming and Shen, Zehong and Wang, Yuang and Bao, Hujun and Zhou, Xiaowei, *LoFTR: Detector-Free Local Feature Matching with Transformers*, Conference on Computer Vision and Pattern Recognition (CVPR) 2021