3DCV Final Project

UAV motion compensation using Structure from Motion and Optical Flow

Group 6

R10525028 蔡承恩 R11922029 吳泓毅 R11525028 江昱翰

A. Introduction

In recent years, computer vision based methods have become a viable alternative for measuring structural displacements in structural health monitoring(SHM) systems. However, many methods require stationary cameras, which may be difficult to deploy in the field with appropriate sight. An unmanned aerial vehicle(UAV) can provide the opportunity to overcome the limitation of camera deployment location.

B. Encountered Problem

The camera motion, subjected to non stationary motion of UAV typically has 6 degrees of freedom(DOF). Due to the 6 DOF non stationary camera motion, the relative displacement of the structure will result in difference compared to the absolute displacement, the non stationary motion of the camera needs to be calculated and removed from the video. In this project, we reference the result from Ribeiro's proposed approach[1]. The camera motion from UAV can be divided into two parts. One for the image processing of video frames and the other for the numerical integration of the signals from the inertial system.

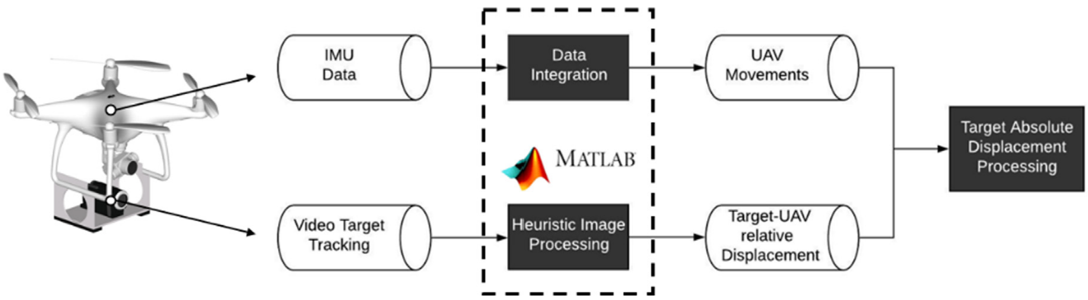

Figure 1. Workflow of Ribeiro's proposed approach [1]

They performed the field test to measure the displacement of a target fixed to a massive Reinforced Concrete(RC) wall with the UAV motion. In this circumstances, the target does not move, thus, the absolute displacement of the structure will be

$$0=\delta_v+\delta_{UAV}$$

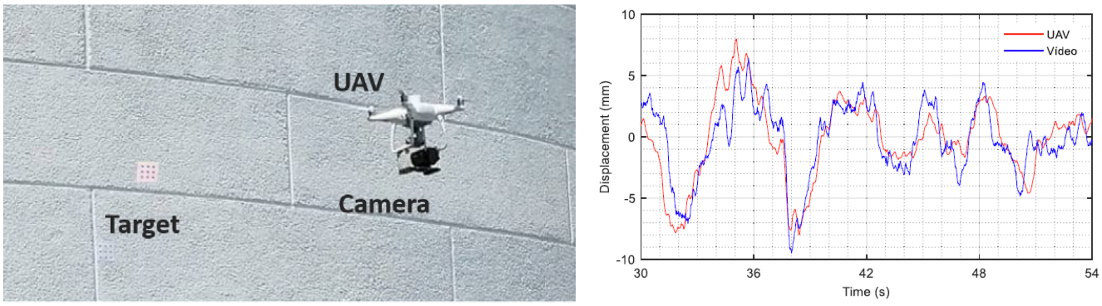

Figure 2. Field test and comparison of displacement records obtained through the video system and UAV inertial system,[1]

The purpose of the test is to validate the inertial system of the UAV, particularly the numerical integration tool, by comparing the displacements obtained through the inertial and video systems. If the inertial system works properly, both displacement values tend to be equal unless the sign. Therefore, the virtual displacement of the target estimated by the video system is only due to the movements of the UAV.

$$\delta_v=-\delta_{UAV}$$

However, It is hard to get the IMU data from UAV. In this report, we estimate UAV 6DOF by structure from motion. The signal will be similar to the red line in Figure 2. The blue line in Figure 2 will be computed by optical flow.

C. Methodology

In this report, we imitated the result from Ribeiro's proposed approach, which is shown in Figure 2. The red line is the IMU sensor after signal processing and we used SfM to recover the camera pose. The blue line is obtained by the image processing for target tracking and we used optical flow to get the target displacement.

a. Structure from Motion:

We followed the steps that we learned in the lecture. However, the camera trajectory is scale ambiguous, the appropriate scale is extracted by using the known dimensions in the environment(e.g. the size of the cylinder). Following are the steps that we estimate the camera pose.

1. Estimate essential matrix
2. Estimate the relative pose from the essential matrix
3. Sparse 3D reconstruction from 2D correspondences
4. Scale recovery

b. Optical flow:

We used a deep learning model called RAFT to do displacement estimation. RAFT is a dense optical flow method, we chose checkboard to be our ROI. Input two frames to get the displacement of each pixel in ROI, and then average them to get the average displacement.

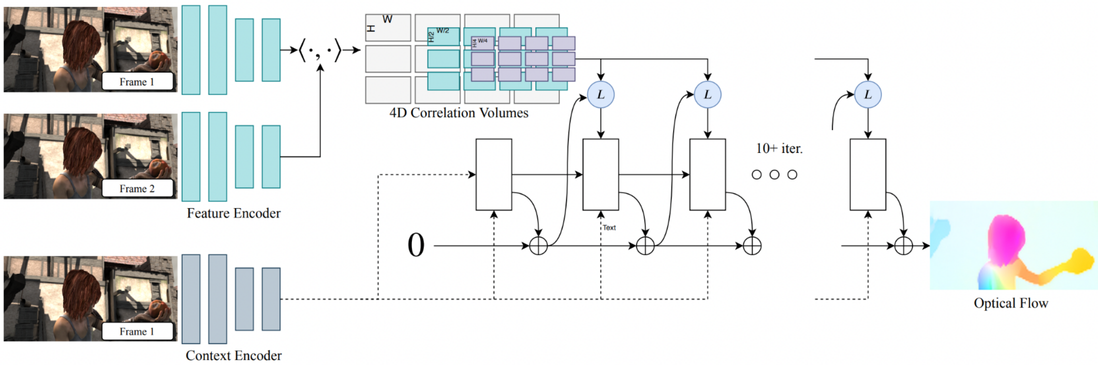

Figure 3.Workflow of RAFT.[2]

1. Get the scale factor between real world distance and pixel level distance.
2. Input two neighbor frames to RAFT model
3. Get the displacement of pixels in ROI
4. Average displacement of pixels in ROI
5. Multiply scale factor to the average displacement to change pixel level displacement to real world displacement
6. Shift ROI region by pixels' average displacement
7. Do Step 2 again until end

D.Experiment & Result

a. Precision Machinary

In the first experiment, we installed the camera on the precision machine. The precision machine moves 0.5cm in each frame which is the ground truth. We use structure from motion to calculate the camera's trajectory. Obtaining the camera pose in each frame. We obtained the 3D point cloud by triangulation as Figure 4 shown. To recover our cylinder, it is known that the height of the cylinder is 21cm. We fit the cylinder's 3D point clouds as shown in Figure 4. Besides, we use RAFT to calculate the displacement in x and y direction to compare the result from structure from motion.

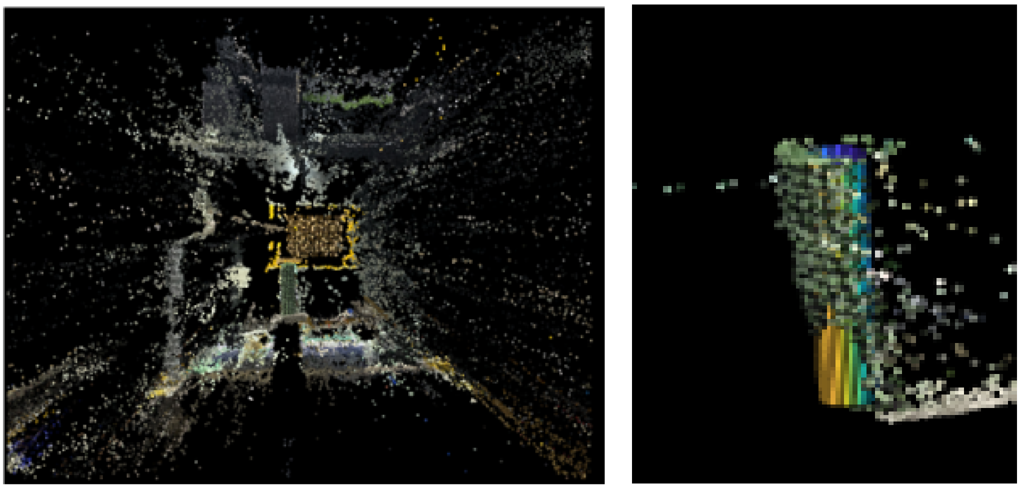

Figure 4. 3D point cloud in Exp1.and 3d point cloud fit cylinder.

By tracking the background features, the 6DOF motion of the camera was estimated as shown in Figure 5. As is demonstrated in this figure, the rotation of the camera is correct due to no rotation during this test. The translation in x and y direction were compared to RAFT, which is shown that both displacement values tend to be equal unless the sign.

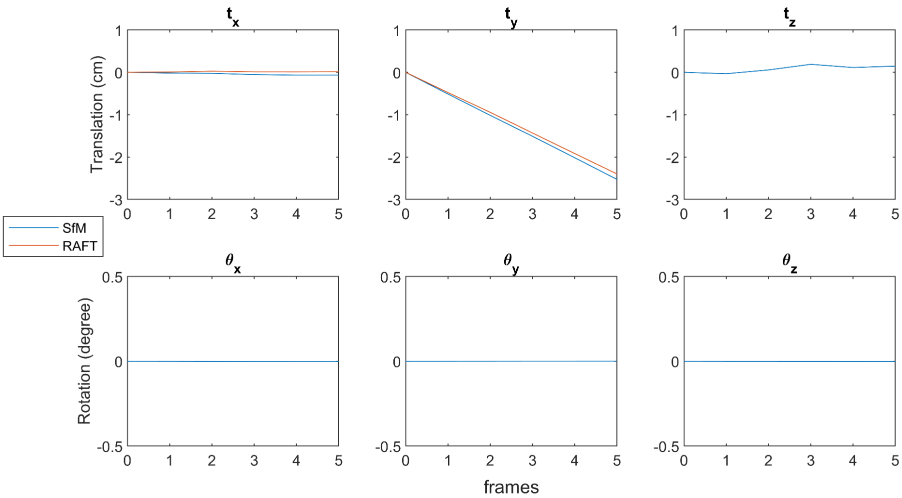

Figure 5. 6DOF motion of the camera and optical flow.

b. Camera motion with hand drift

The second experiment simulates the UAV motion by hand. Because of taking it by hand, there is no ground truth in this experiment. In this experiment, we purposely move the camera in x direction and no rotation in our video. After getting the video. We compute the camera pose from video by SfM and triangulate the 3D point clouds as shown in Figure 6.

Figure 6. 3D point cloud in Exp2.

By tracking the background features, the 6DOF motion of the camera was estimated as shown in Figure 7. As is expected, the x and y direction had the strongest similarity with the result from RAFT.

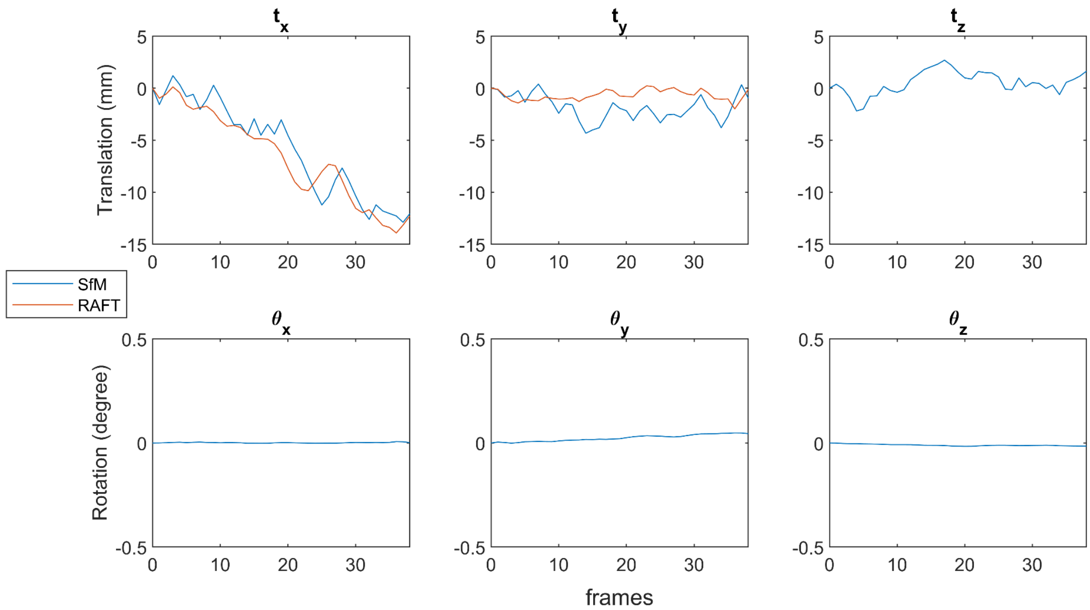

Figure 7. 6DOF motion of the camera and optical flow

E. Conclusion

In experiment 1, we obtained ground truth of camera pose by precision machinery, and used SfM to estimate camera pose and using optical flow to estimate target's (checkboard) displacement, the result shows that our method is correct. (the translation of camera pose is negative of optical flow.) In experiment 2, we do camera movement by hand, so we do not have ground truth this time, and the result also shows that our method can correctly estimate object (checkboard) displacement and that the translation of camera pose is similar to negative optical flow.

F. How to Run Code

We wrote this in our README.md .

G.Participation

a. 蔡承恩
- Structure foem Motion implementation
- Experiment
- Presentation
- Report

b. 吳泓毅
- Optical flow implementation
- Presentation
- Report

c. 江昱翰
- Video clip

H.Reference

1. Ribeiro, D., et al. "Non-contact structural displacement measurement using unmanned aerial vehicles and video-based systems." Mechanical Systems and Signal Processing 160 (2021): 107869.
2. Teed, Zachary, and Jia Deng. "Raft: Recurrent all-pairs field transforms for optical flow." European conference on computer vision. Springer, Cham, 2020.