CS512 Final project

Zhiquan Li A20381063

Fangzhou Xiong A20376382

Illinois Institute of Technology

I. Problem Statement

The paper[2] propose a more efficient tracking algorithm, which is a hybrid version of the traditional optical flow approach and efficient second order minimization algorithm(ESM). Optical flow tracking the object by finding similar point in the neighbor of a set of feature points, however, the feature points are limited, thus the tracking result is not robust.
ESM is a region-based algorithm, it is tracking the whole object of a template, the algorithm could have large iteration and thus low efficiency.
The hybrid algorithm proposed by the paper trying to solve the defects of the both algorithm.

II. Problem Statement

II.1 Oriented FAST and Rotated BRIEF(ORB)

ORB is an algorithm which combines FAST keypoints detection and BRIEF descriptor. First it uses FAST to detect keypoints in image and then use Harris corner to find the top N corners among them. After detecting feature points in template and image scene, ORB will use brute force to compare each pair of BRIEF descriptors corresponding to keyponits in template and image scene to find the minimum hamming distance for each pair. Then ORB finds the relationship of corner between two images. At last, ORB can only detect the feature points of marker in the scene.

II.2 Optical Flow(OF) algorithm

OF is the pattern of apparent object motion in a consecutive frames in movement of object or camera. There are two assumptions to make sure OF works: 1. the pixel intensities of an object do not change between consecutive frames. 2. Neighboring pixels have similar motion.
For a pixel in image, OF will compute the distance in next frame taken after dt time, so we can say,

$$\mathrm{I}(\mathrm{x}, \mathrm{y}, \mathrm{t}) = I(x + dx, y + dy, t + dt)$$

Then, OF could transform the equation to get optical flow equation by using taylor series approximation of right-hand-side with simple modification,

$$f_x u + \ f_y v \ f_t = 0, \text{ where } f_x = \frac{\partial f}{\partial x}, f_y = \frac{\partial f}{\partial y}, \mathrm{u} = \frac{\partial x}{\partial t}, \mathrm{v} = \frac{\partial y}{\partial t}$$

II.3 Efficient Second Order Minimization

Efficient Second Order Minimization is a region-based algorithm, it does not rely on the feature points, instead, the algorithm try to match the whole object by minimizing the difference of the current warped image intensity and the template. For tracking planar, the homograph matrix of the current image planar to the template should be the object. In order to optimizing the final result, we could using iterative method to minimizing the errors.
Let the function $w(x)$ be a warping function transform the template point p to current image point p', the x is the parameter variables in the homograph matrix.
Set $x = e$ as the original parameters, that $p = w(x)(p)$, and x_c as parameters for last image, Δx be the increase of the parameters from x_c to current image. So we could represent the current image point as $w(x)(w(x_c)(p))$. for the Let $I(p)$ be the intensity of the point p.
So we could have the error as:

$$E(\Delta x) = \sum_{\{i=0\}}^{\{n\}} \left(I(w(x_c)(p)) - I1\left(w(\Delta x)\big(w(x_c)(p)\big)\right)\right)^2$$

Equation 1

To minimizing the error, the necessary condition is:

$$\frac{dE(\Delta x)}{d\Delta x} = 0$$

Than we can approximate the Δx with second order approximation according to paper [5] as:

$$\Delta x \approx -2\big(J(e) + J(X_c)\big)^{+} \Delta I$$

Equation 2

For each point P, we could have approximation as:

$$\Delta I_k \approx -1/2(\frac{\partial I(P)}{P} + \frac{\partial I(w(x_c)(P))}{P})\frac{\partial w(x)(P_k)}{x}\Delta \mathrm{x}$$

Equation 3

For give x = e, and P $=P_k$. So, we could append $-1/2(\frac{\partial I(P)}{P} + \frac{\partial I\big(w(x_c)(P)\big)}{P})\frac{\partial w(x)(P_k)}{x}$ as a row of $J(e) + J(X_c)$.
Since ΔI_k is know, and $(J)^{+} = (J^{TJ})^{-1}J^{T}$, we could have the estimation of Δx for one iteration. And calculate the new x_c for next iteration.
For each iteration, we can calculate the L2 norm of the Δx, if the norm is small enough, we consider the approximation has converged and output the final homograph matrix.

II.4 Advanced Image Tracking Algorithm

[Advanced Image Tracking Algorithm]

Computing the template features points P_t

1

for each new frame:

get new features point P_2 using OF algorithm.

find homograph $H_\Delta$base on previous frame feature points P_1 and P_2.

find homograph $H_2$to template base on previous homograph to template H_1

find estimate homograph $\widehat{H_2}$ using ESM algorithm with H_2

estimate $\widehat{P_2}$ base one $\widehat{H_2}$ and P_t

P_1 = $\widehat{P_2}$

H_1 = $\widehat{H_2}$

We known $p_1 = H_1 p_t$, and $p2 = H_\Delta p_1$, so we could have $p_2 = H_\Delta H_1 p_t$, we could have $H_2 = H_\Delta H_1$.

III. Implementation detail

III.1 Oriented FAST and Rotated BRIEF(ORB)

In marker detection, we implement ORB feature extraction by using functions provided by openCV. Firstly, we create an ORB object, then use detectAndCompute function to find feature keypoints and brief descriptors from template and image scene. Secondly, we use BFmatcher function to find the keypoint index with minimum hamming distance of brief descriptors. Thirdly, we only select keypoints which are matched with the keypoints in template from the image scene.

III.2 Optical Flow

In image tracking, we implement optical flow algorithm by using calcOpticalFlowPyrLK function provided by openCV. For this function, it needs three key parameters to find the feature points in consecutive frame, which are previous frame, current frame and feature points needs to track in previous frame. Then the function will return the position of tracking feature points. After getting the feature points in current points, we continue to assign the current frame to the “previous frame”, current feature points to feature points in “previous frame” and assign next frame to “current frame”, then we can continue tracking an object movement in a consecutive frame.

III.3 Efficient Second Order Minimization

III.3.1 Homograph parameterization

Following the homograph parameterization discussion in the paper[5]. Assuming 8 degree free of the homograph matrix H(x). To prevent big displacement of the camera to have $h_{33(x)}$, which is the last row and last column value of H(x) to be 0, we should choose H(x) with determine to 1. So let:

$$H(x) = e^{A(x)}$$

$$A(x) = \sum_{i=1}^{8} x_i G_i$$

where trace of all the G_i is 0, and independent to each other hence determine of H(x) is 1.

So, assuming $x = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8,]$, One of the solution for H(x) could be:

$$H(x) = \begin{bmatrix} x_7 & x_1 & x_2 \\ x_3 & -x_7 - x_8 & x_4 \\ x_5 & x_6 & x_8 \end{bmatrix}$$

III.3.2 Computing $\mathrm{J(e) + J(x_c)}$

Using the equation 3, we could compute each row of $\mathrm{J(e) + J(x_c)}$ matrix. as $\frac{\partial I(P)}{P} + \frac{\partial I(w(x_c)(P))}{P}) \frac{\partial w(x)(P_k)}{x}, where$:

$$\frac{\partial I(P)}{\partial P} = [I_{tx}, I_{ty}]$$

the 1 by 2 matrix is the gradient of the template at point P_k this could be calculated when the templated is selected.

$$\frac{\partial I(w(x_c)(P))}{\partial P} = [I_{wx}, I_{wy}]$$

the 1 by 2 matrix is the gradient of the warped image at point P_k, this matrix has to calculated for each iteration.

$$\frac{\partial w(x)(P_k)}{\partial x} = \frac{\partial f(e^{A(x)}))(P_k)}{\partial x} = \frac{\partial f(\hat{x})}{\partial \hat{x}} \frac{\partial A(x) P_k}{\partial x}$$

where f(x) is a function the input is a matrix $x = [x_1, x_2, x_3]$, the output is:

$$f(x) = [\frac{x_1}{x_3}, \frac{x_2}{x_3}]$$

The derivative is the value in $x = e$. Give $P_k = [u, v, 1]$

$$\frac{\partial f(\hat{x})}{\partial \hat{x}} = \begin{bmatrix} 1 & 0 & -u \\ 0 & 1 & -v \end{bmatrix}$$

$$\frac{\partial A(x)P_k}{\partial x} = \begin{bmatrix} v & 1 & 0 & 0 & -u^2 & -uv & u & -u \\ 0 & 0 & u & 1 & -uv & -v^2 & -v & -2v \end{bmatrix}$$

Assuming $\left(\frac{\partial I(P)}{\partial P} + \frac{\partial I(w(x_c)(P))}{\partial P}\right) = [I_x, I_y]$, we could have a row vector of point P_k in $\mathrm{J(e)} +$ $\mathrm{J(x_c)}$ matrix as:

$$\left(\mathrm{J(e)} + \mathrm{J(x_c)}\right)_k = [vI_x, I_x, uI_y, I_y, -u^2 I_x \; - uvI_y, -uvI_x \; - v_y^{2I}, uI_x - vI_y, uI_x - 2vI_y]$$

III.3.3 Warping Image

To warp the image, we assuming template is rectangle, and the coordinate of template start from (0,0) to (x,y), where x is the length of the template, y is the width of the template, both in pixel.

Than converting the coordinate to 2D homogeneous coordinate and warp with homograph matrix to have the coordinate in the compared image.

III.3.4 Stabilization and efficiency

The L2 norm of Δx has important impact to the efficiency and stabilization of the algorithm, if the norm is too small, the iteration number could be large, and increasing the risk of calculation errors lead the matrix to be singular and time complexity, the small number of norm could lead to the failure of tracking. We use 0.003 as threshold.

The efficiency of ESM is a key part for it's tracking performance, since the algorithm rely of continues image of object, it time for one image processing is long, ESM could find previous and current frame image being very different in real time processing. To lower the time complexity of ESM, we use "cython" to implementing the algorithm.

III.4 Hybrid Image Tracking

The hybrid algorithm rely on OF and ESM algorithm, we can use algorithm provided in II.2 with function provide by ESM and OF section to implement the algorithm.

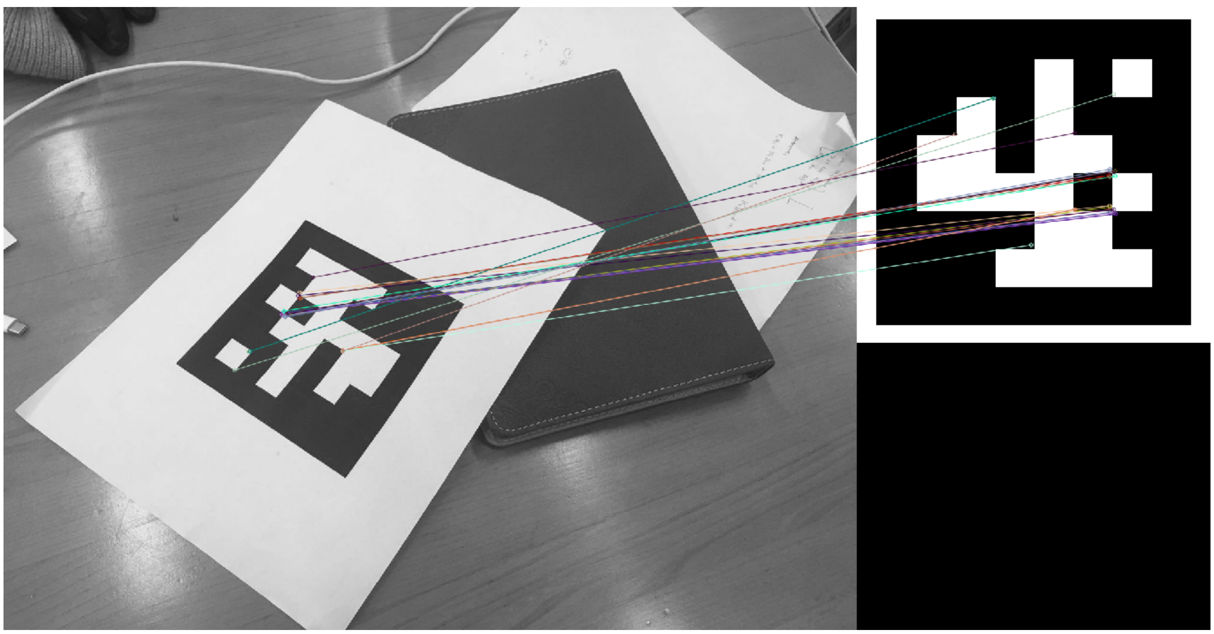
III.4.1 Detecting ORB feature points for template

We using brute force matching function provided by opencv[4] to find best matched feature points from the image, To avoid noise, we select top 20 best matched points as feature point.

IV. Result and Discuss

IV.1 Result of ORB

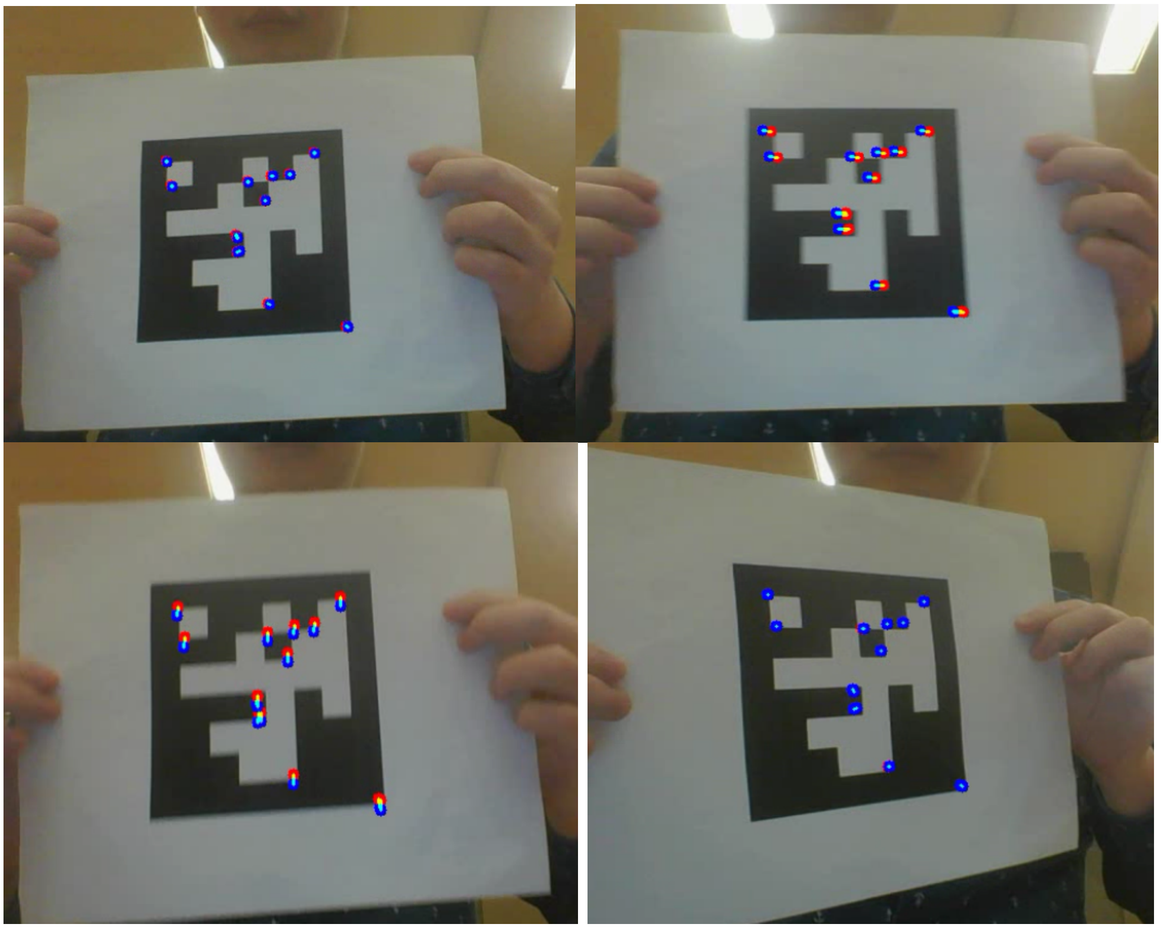
After implemented ORB feature extraction and feature mapping, we can have result below:



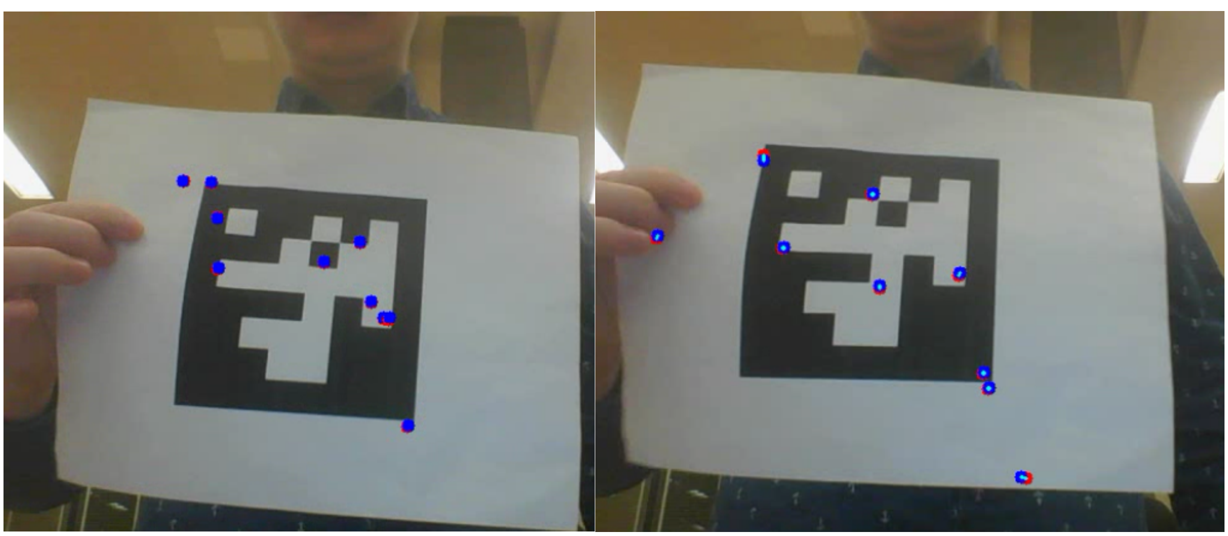
We can see the algorithm has well performance to find the feature points in marker. Usually, the number of feature points in image scene will greater than template,

IV.2 Result of OF

After implemented OF algorithm, the result is showed below:



We can see optical flow has well performance in image tracking, however, there is drawback of optical flow, which is it often fails tracking when the movement of image or camera too fast, the keypoint will move to other position in the image. We can see the drawback below:



By shaking my hand quickly, we can see several points shift to other position, the reason is fast movement will conflict the second assumption of optical flow, which is neighboring pixels have similar motion

IV.3 Result of ESM

IV.3.1 Tracking performance

The tracking performance for rotation, translation, moving in depth direction, are showing below:

Figure 1 Moving in depth direction

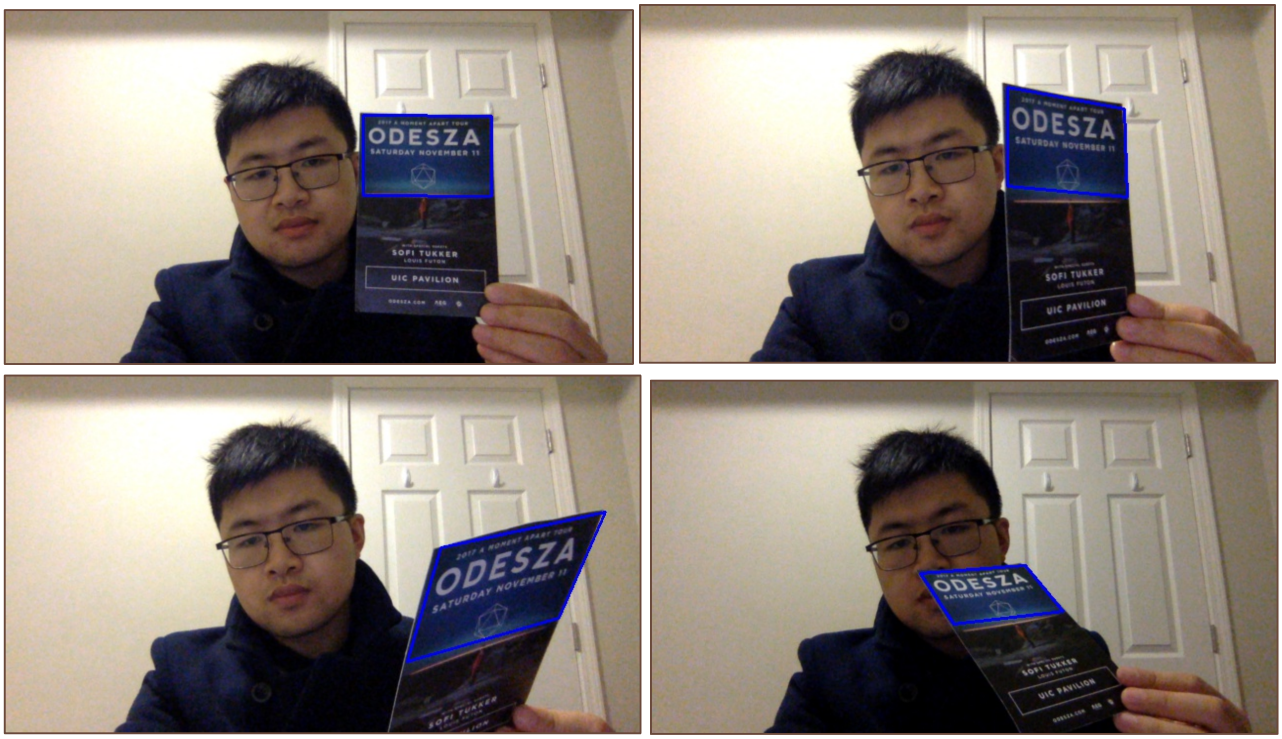


Figure 2 Roation Performance

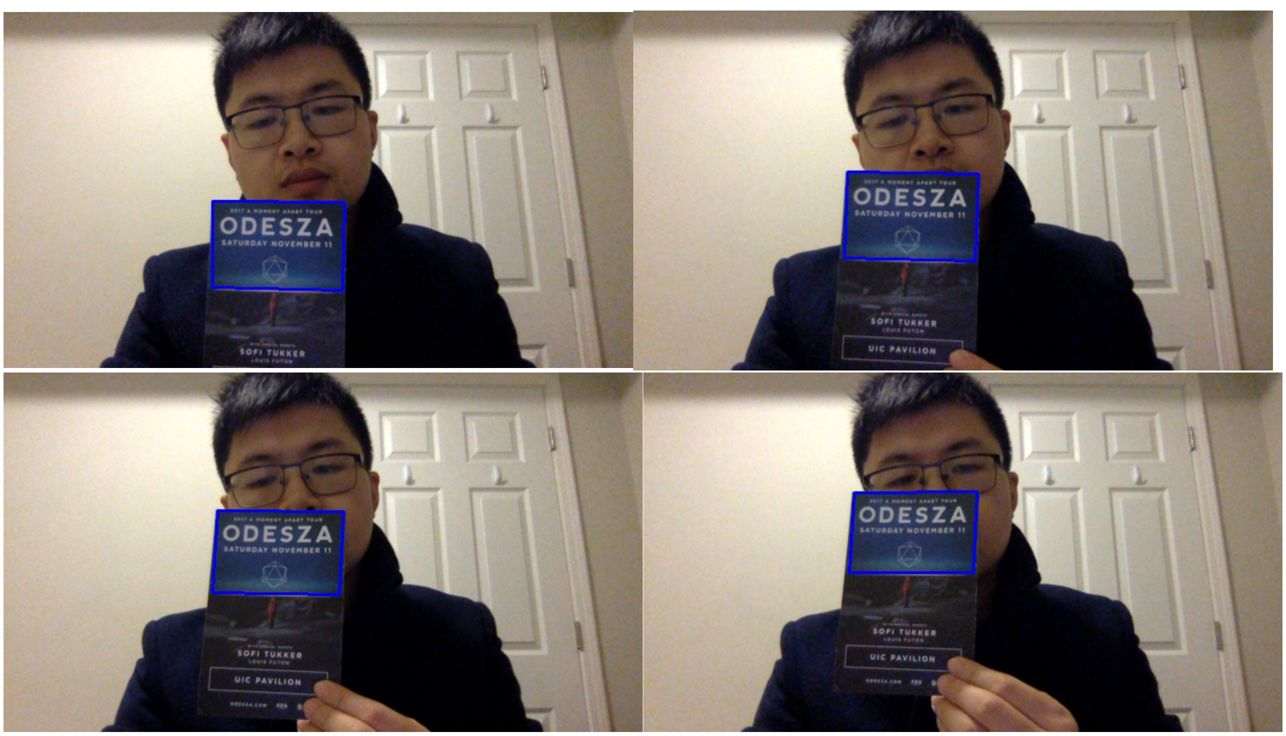
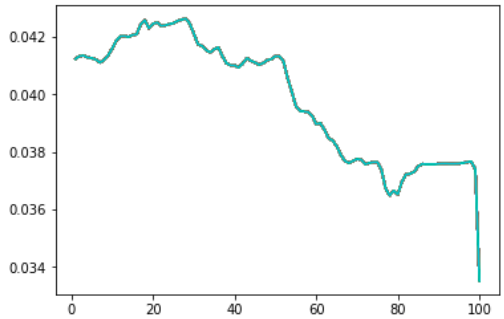


Figure 3 Translation Performance

IV.3.2 Converge rate

We collected the L2 norm to the ∆x in the experiments of testing Translation performance above, several sample are selected, and change of ∆x to the time are showing below:



IV.3.3 Iteration number

We test iteration number in three scenario, static, dynamic and moving(translation) situation, the dynamic situation is when user hold the marker and shaking in a small degree. Max iteration number is 100

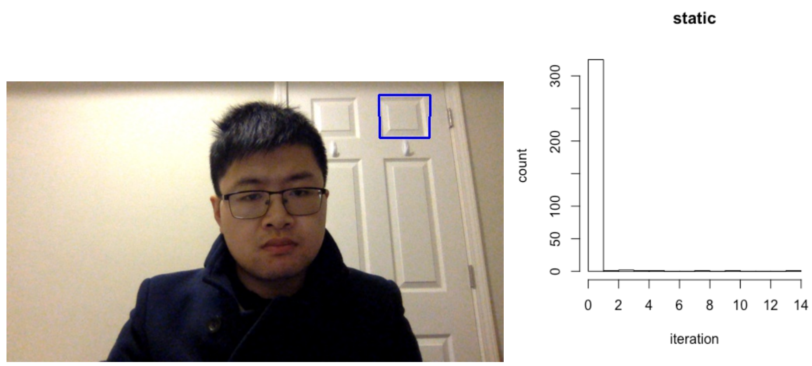


Figure 4 Iteration number distribution when tracking static object

Figure 5 Iteration number distribution when tracking dynamic object

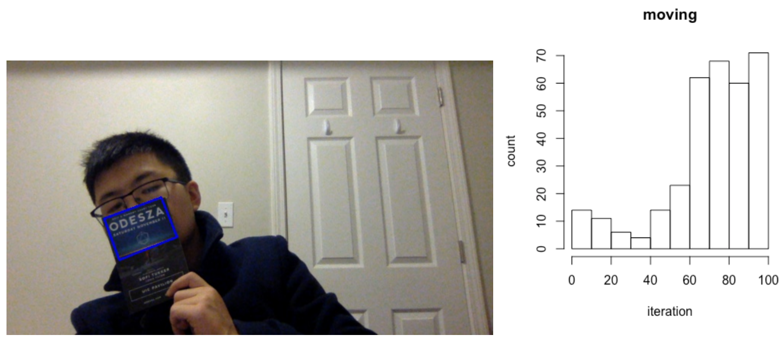
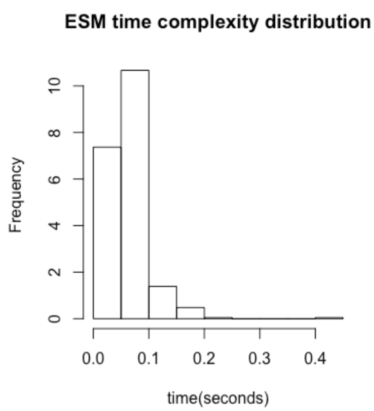


Figure 6 Iteration number distribution when tracking moving object

IV.3.4 Time complexity

The time data is sampled from the above operation, the matrix is no more than 200*200, no less than 100*100. The average time is 0.05966747 seconds, 0.001035929 seconds for min, 0.4021769 for max, the distribution graph is listed below:



References

[1] rabalho de Gradua ̧ca ̃o , Claudio Jung Orientador , 3D Tracking for an Augmented Reality System

[2] Gorovyi, M. & Sharapov, S. (2017).Advanced Image Tracking Approach for Augmented. Retrieved from http://ieeexplore.ieee.org/document/8053687/

[3] Tutorial_py_lucas_kanade. Retrieved from https://docs.opencv.org/trunk/d7/d8b/tutorial_py_lucas_kanade.html

[4] ORB (Oriented FAST and Rotated BRIEF) . Retrieved from https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_orb/py_orb.html?highlight=orb

[5] Selim BENHIMANE and Ezio MALIS (2014) , Real-time image-based tracking of planes using Efficient Second-order Minimization

[6] Ethan Rublee Vincent Rabaud Kurt Konolige Gary Bradski , ORB: an efficient alternative to SIFT or SURF

[7]Efficient Second Order Minimization (ESM), Retrieved from https://www.edwardrosten.com/cvd/cvd/html/group__gEsm.html

,