COMPUTER VISION
ENEL808

ASSIGNMENT 1 – GESTURE RECOGNITION

MIN CHOI 1132188

Introduction

The purpose of this assignment was to try and code in MATLAB, a program that can detect hand gestures from a video. The requirements were to detect how many fingers are held up and which direct the hand swiped across then output a string that shows all the gestures detected. Before I begin the report, I must confess that I was not able to successfully complete all the tasks.

Methods

First step was to try and separate the hand from the background. The final method I have used was the “vision.ForegroundDetector”. this made things very simple to extract the foreground from the frame. Before discovering this method, I have spent couple of days trying to use simple image processing we have learned in the laboratory. I also tried to use the frame where no hand is present and using them to try and establish a baseline then subtracting it from the frames where the hand is present. This wasted bulk of my time before I finally discovered the method, I ended up using. The problem was that the colour of the hand was too similar to the background for any simple image processing methods to work properly.

After separating the hand from the background, I have cleaned up the noise that was also detected as foreground. The noises were from the shadow and change in the intensity of the light. I final method I used were “imopen” and “Imclose”. “imopen” to try remove all the small noise pixels then “imclose” to try and connect the near by pixels. I have also tried “bwareaopen”, “imdilate” and “imerode” but ended up using just the afore mentioned two methods because I was spending too much time trying to dial in the processing just right for every video. Because I have only used two image processing to clean up the frames, I ended up getting perfect result for some videos and for some there was huge noise from a shadow, but I had to skip to next step because of the time constraint. The hand also has disjointed fingers because some parts of the background is too similar to the finger and I did not have enough time to try and find a setting that will work best over all the videos.

Next step after separating hand from background and removing noise was to detect the swipes and fingers. For the fingers I have used “imfindcircles” method to try and

detect the round edges of the finger. This worked great for some videos and very bad for others. For swipe detection I created a bounding box around the whole hand that was detected then used the coordinate of the bounding box. I have first set up a trigger for the swipe, so it doesn't report slight movement sideways as a swipe. The trigger was set to detect swipe if the hand moved more than just over a half of the screen width. I have implemented this by taking the maximum value of the x coordinate then subtracting the minimum of the x coordinate to try and see if the hand has moved more than the set number of pixels. Then to detect whether the hand moved towards the left or right I found the column number where the maximum and minimum is present in, then subtracted coordinate of the maximum from the coordinate of the minimum. If the maximum appeared earlier than the minimum in the array, that means that the hand started from right side then moved to the left.

This is as far as I have gotten for this assignment. I was not able to implement the feature of outputting the string of detected gestures.

Result

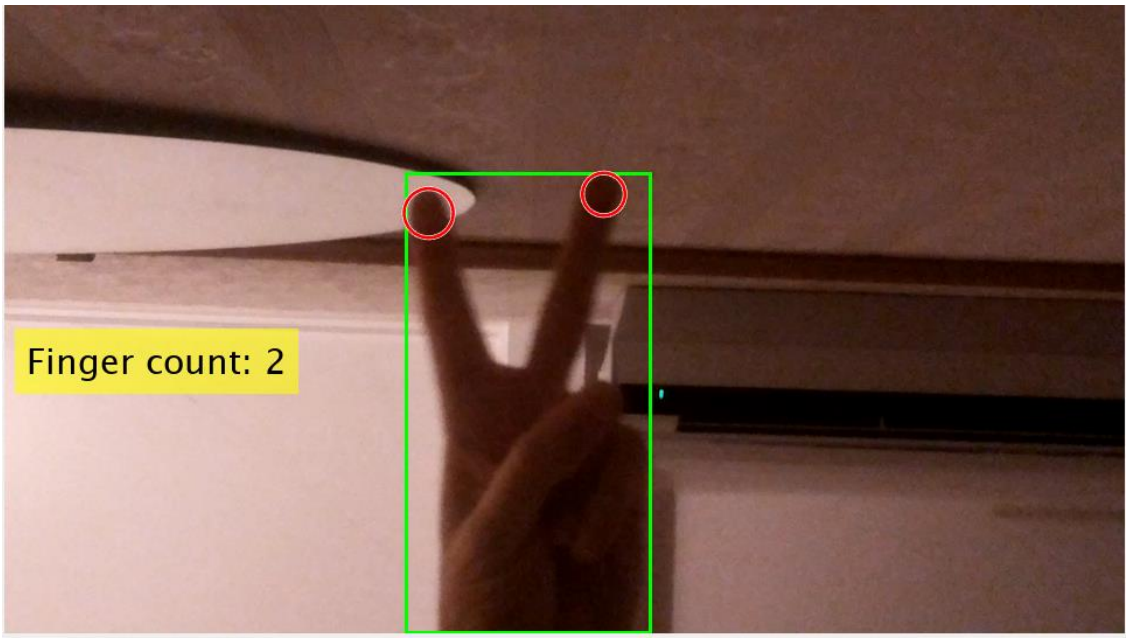

Figure 1 - Sample of the Result

(Figure 1) above shows one of the results for finger detection.

Speed

My current program is fairly fast but as it is not fully working yet it is not really viable indicator. The part that I have noticed took longest was the image processing such as “imopen” etc. “imfindcircles” was another part that slows down the process especially as I increased the pre-set to find larger circles. One thing I have done to speed up the process was to resize the frame to half. I have tried going smaller but that just made it more inaccurate.

Future work

As I was not able to successfully complete the assignment there are a lot of work that still needs to be done. Firstly, I have to make MATLAB output a string of detected gestures. For me to be able to do this I must first be able to save the number of detected fingers. The method I was going to try use was to detect when the hand comes up then goes down trigger a function to save the current count of fingers while hand is in between going up and going down when it is steady. I was successful in detecting the movement of the hand in the y axis but ran out of time to try and program the string output feature. After I finish coding the function to save the finger count, I still have a long way to go to improve the accuracy of the finger count and swipe detection. Especially for the video where I get lots of noise from the background shadow. So, the image pre-processing needs to be vastly improved.

Reflections

I found this paper very interesting and fun, but assignment was very hard as this is my first-time using MATLAB to this extent and also my first time learning anything related to computer vision. No wonder only 6 students signed up.

The most important thing I remember is to use mean not average. I think computer vision is very interesting field because it is advancing very fast with the introduction of deep learning and AI technology and improving computer power. I would like to learn more about computer vision and also try finish my assignment code if I get a chance.