



TOSHKENT SHAHRIDAGI INHA UNIVERSITETI
INHA UNIVERSITY IN TASHKENT

School of Computers and Information Engineering
Spring semester 2020

Multimedia Computing (SOC4020)
Term Project Report

Hand and Finger Gesture Recognition in Computer
Vision using Python and OpenCV

Name:

Sardor Allabergenov

ID:

U1610202

Group:

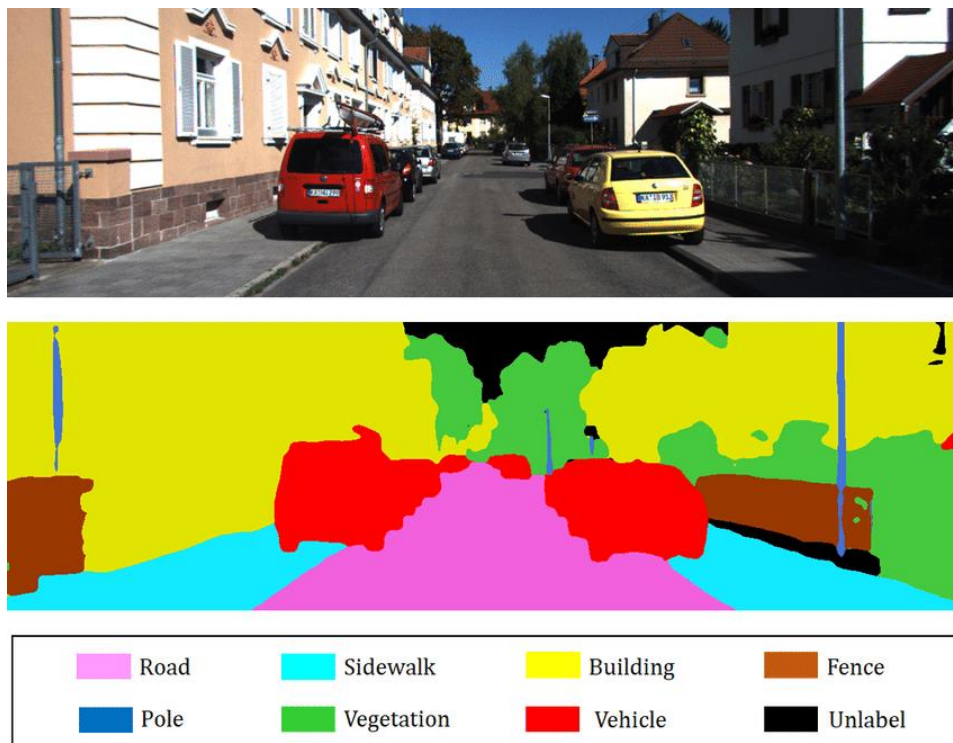
CSE 16-2

Abstract

In this project, I investigate the hand gesture recognition and finger count to do some operation by Computer Vision. I abstractly choice the control of audio by counting the fingers and do some operations like song playing, sound up and so on. I significantly determined that the controlling without some remote-control systems faster than the using remote controller. There is no used the Artificial Intelligence or Machine Learning it is only finding gestures with contours of hand and fingers. Hand gesture recognition is a cool task to begin for a Computer Vision devotee as it includes a natural bit by bit technique which could be handily seen, so you could manufacture progressively complex stuff on these ideas.

Introduction

The gesture recognition problematic issue for Computer Vision at all in case of object detection in real time. In the same time to looking to one image by eyes of human and computer the different thing. Because the object in picture human can detect and can segmentation of foreground object from a background without any difficulties, whereas computer should do segmentation by its contours of object in picture. But it is challengeable in real time motion detection to semantic segmentation. You can see the image below to get some imagination about that:



<https://www.researchgate.net/publication/326875064/figure/fig3/AS:659518916681730@1534252971987/Example-of-2D-semantic-segmentation-Top-input-image-Bottom-prediction.png>

The semantic segmentation is a way to focus on our segment or region to get foreground object.

This picture portrays above the semantic division issue where the goal is to discover various locales in a picture and label its relating names. For this situation, “road”, “sidewalk”, “building”, “fence”, “pole”, “vegetation”, “vehicle”, and “unlabeled”. As this is an exceptionally troublesome issue to explain, we will limit our concentration to pleasantly section one frontal area object from a live video grouping.

To do the semantic segmentation for this I used the opencv and python libraries and methods of them. Like edge detection, find contours and so on.

My approach

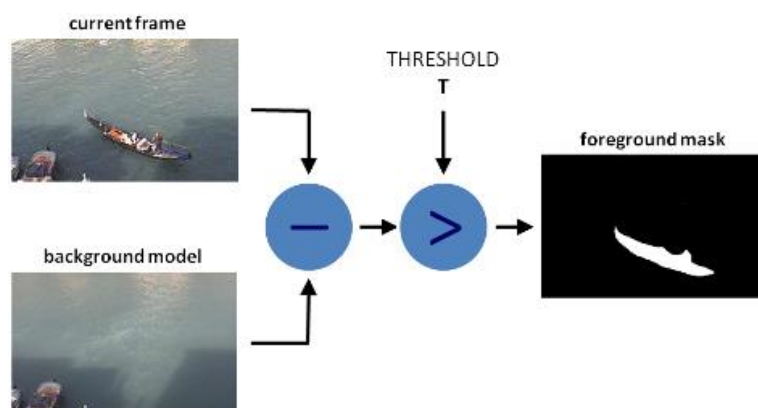
To begin with, we are going to recognize it by video sequencing. We grab the video into frames to process with them.

I divide into the sections like:

- Segment the hand region
- Subtraction it from background
- Motion detection
- Thresholding the difference image
- Finding Contours
- Counting fingers

First of all, I draw a segment to get rid of another parts of our frame which is not needed by using ROI in opencv library.

To subtraction from background we using the concept of running averages. When we grabbing into frames, we compute the running average over the current and the previous frames.



https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtraction.html

After determine the background model using running averages, we calculate the absolute difference between the background model and the current frame to obtain a difference image that holds the newly added foreground object which is our hand.

To distinguish the hand district from this difference picture, we have to threshold the distinction picture, with the goal that solitary our hand area becomes noticeable and the various undesirable areas are painted as dark with the method threshold() by opencv function.



<https://docs.opencv.org/2.4/doc/tutorials/imgproc/threshold/threshold.html>

After that, we find the contours in the resulting image. To do that we assume that the largest area is will be our hand.

To find the and count up our fingers, we use the convex hull with extreme points. Then we find the center of palm using these extreme points which we found before. After that to construct circle with the maximum Euclidean distance by using palm center. The bitwise and operation between hand image and circular ROI reveals the finger slices. At the end, we can calculate the number of fingers by them.



Original



Result

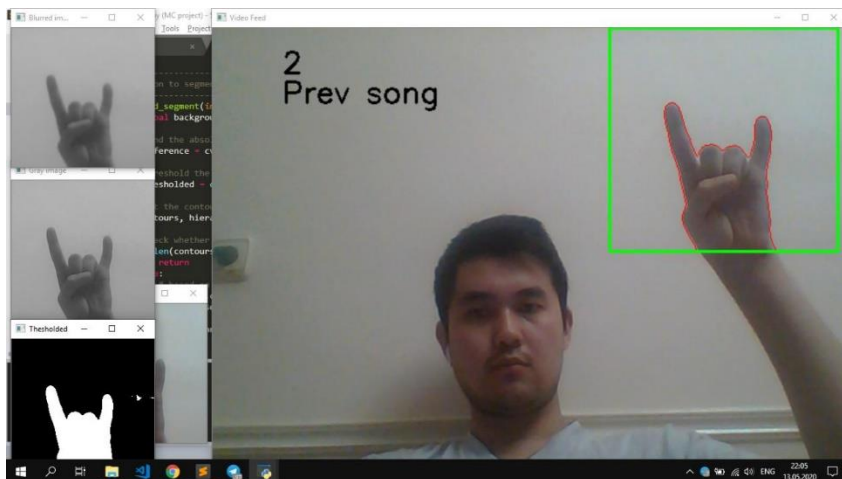
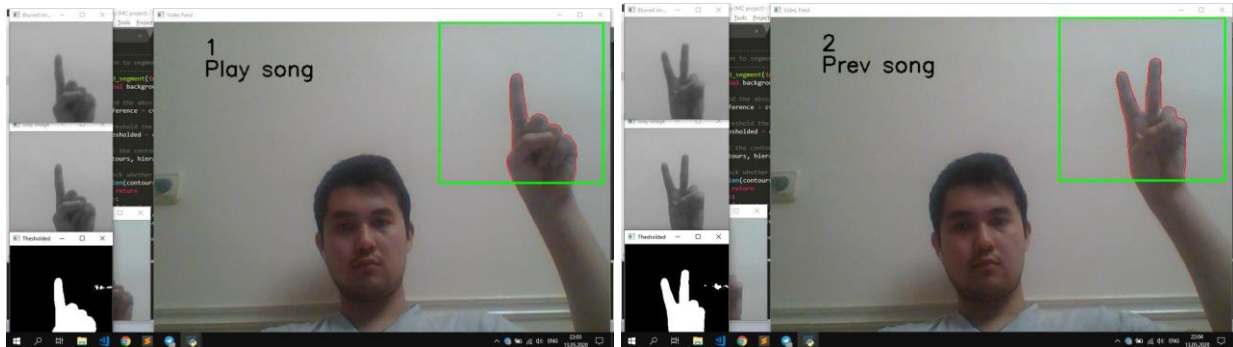
https://docs.opencv.org/3.4/d7/d1d/tutorial_hull.html

Experiment result

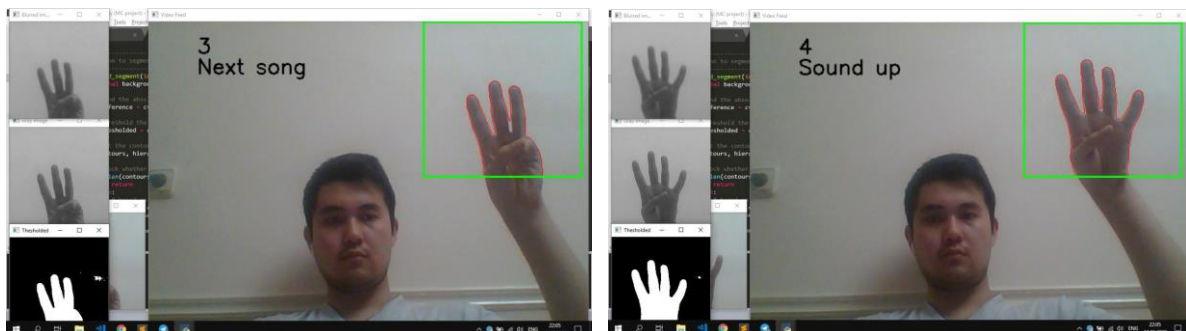
The project implemented by using Python and OpenCV libraries. You can download the entire code to perform Hand and Finger Gesture Recognition from [GitHub](#).

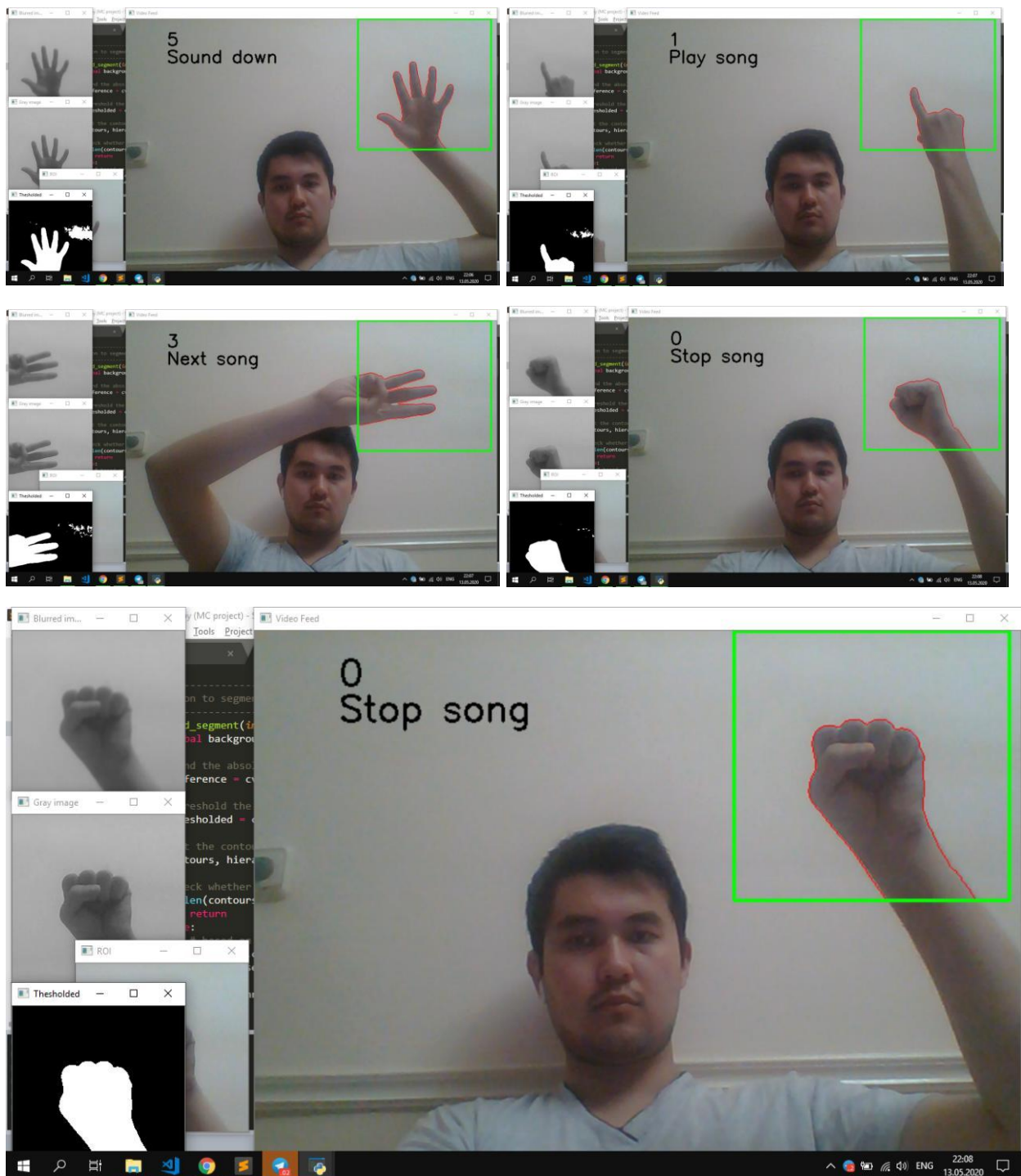
I assume that this project is focused on controlling many devices like Smart TVs, projectors and etc.

Here below you can see some resulting images:



Here you can see the different cases of detecting fingers.





Here you can see the hand gesture which is closed and no finger is detected. I assume the concept of audio controlling in smart house. And here is a main function to control system the audio player.

References:

https://docs.opencv.org/3.4/d7/da8/tutorial_table_of_content_imgproc.html

https://docs.opencv.org/3.4/d7/d1d/tutorial_hull.html

https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtraction.html

https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html