RealSense based system for tracking Hand Pose

Damindarov Ruslan, Fam Cham An

25th February 2021

Introduction

The interest to the touch screens is growing up every year. The reason is that it is not obligatory to wear auxiliary tools, such as glasses or trackers which may cause discomfort for the users. The designing of so called touchless touch screens would definitely attract people, especially in this pandemic period. The task of these devices may be different and it depends on the goal of designing them. For example, they can detect the hand, fingers or even objects on the surface.

The purpose of our project is to create a mechanism of tracking the hand pose using the RealSense depth camera. This method may be useful in developing of interactive screens.

Related work

The earlier work considered the using Kinect V2 as a depth camera for projectable interactive surfaces for children's education purposes [1, 2]. The authors pointed out the portability of camera-projector systems, which means that they can work at any surface. It can be a floor, table or some screen. And the cost of such systems is not high, so any family can afford to buy them.

Idea of implementation

The idea is the following: we have the projector that projects some screen on the surface and the camera mounted above the projector. When we run the program the camera turns on and the program tries to detect a hand and track its position (fig.1). In our case we took the center of the hand as the point to track. In the first method we also track the furthest point of the hand (it can be forefinger or other fingers).

For better understanding of how the system works we have launched the C# project using the Kinect V2 camera. Then we focused on the implementation of the RealSense D435, wrote the code on Python and tested the performance.

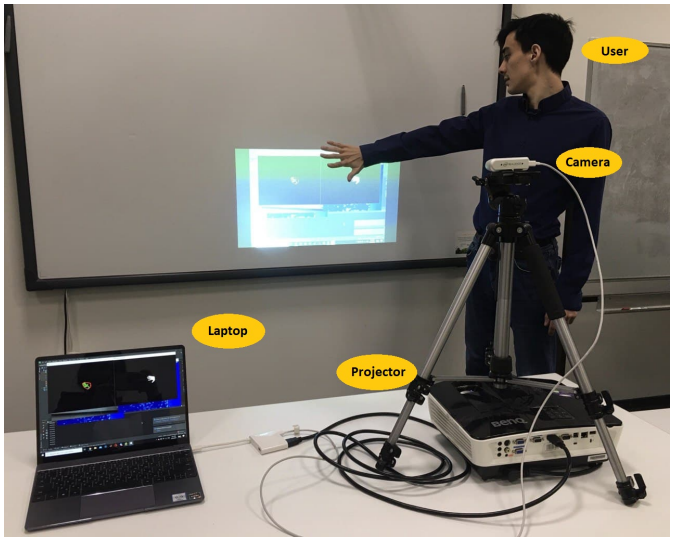

Figure 1: Setup configuration

First method

The hand was detected by using skin detection. The touch is detected by the change of the distance in the moment of interaction and the tracking of the hand position is possible due to tracking the center point of the hand (fig.2, 4). Depth map is used for measuring the distance from the camera to the object.

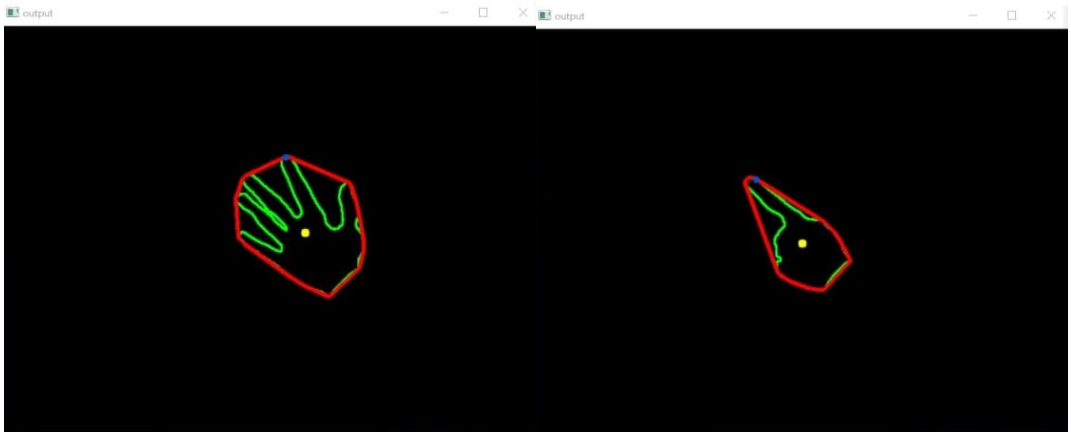

Figure 2: Detection of the hand (the contour of the hand is labeled as red, the yellow point is the center of the hand, the blue point is the furthest point)

Second method

In the first method other parts of the body may be detected (for example, the face, the whole arm). This prevents an accurate tracking. For this reason we tried implement another way for detecting the hand. Here only depth map is used. The depth is limited from 0.5 to 0.7m and in this range the hand if detected by transforming depth map to the color map (fig.3, 4).

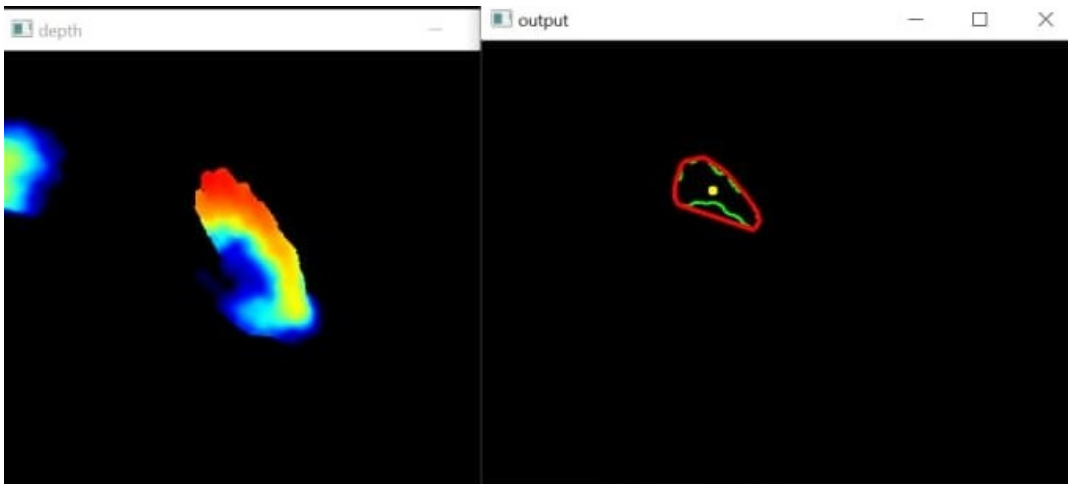

Figure 3: Detection of the hand (on the left - depth map; on the right - the contour of the hand is labeled as red, the yellow point is the center of the hand)

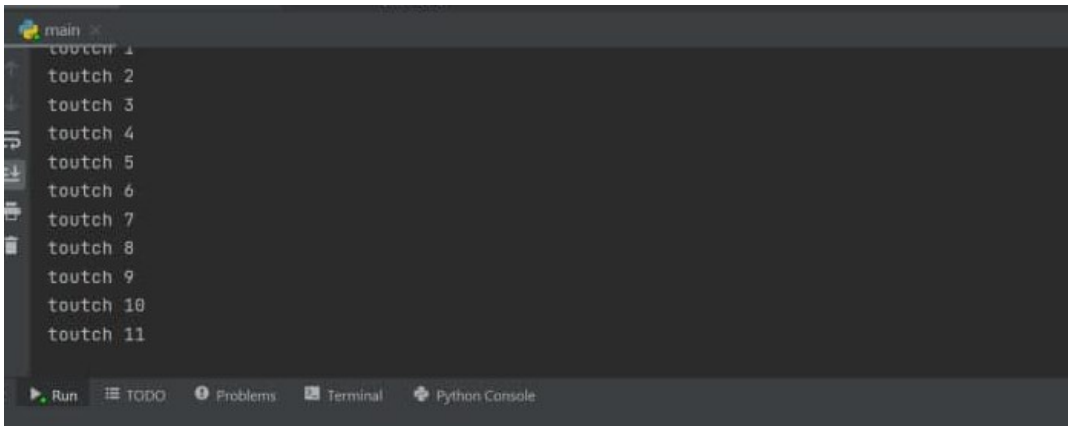

Figure 4: Registering the touch

Future work

We are planning to do the calibration for better positioning and we also will track the hand with another configuration setup which will ease the task of tracking.

Link to GitHub

https://github.com/Damindarov/HandPoseTracking

References

[1] P Sharma, RP Joshi, Riby Abraham Boby, SK Saha, and Takafumi Matsumaru. Projectable interactive surface using microsoft kinect v2: Recovering information from coarse data to detect touch. In *2015 IEEE/SICE International Symposium on System Integration (SII)*, pages 795–800. IEEE, 2015.

[2] Riby Abraham Boby, R Prakash, SK Saha, T Matsumaru, P Sharma, and S Jaitly. Calibration and statistical techniques for building an interactive screen for learning of alphabets by children. In *International Journal of Advanced Robotic Systems*, volume 14(3), 2017.