

Hand Gesture Recognition

Abstract :

Use of hand gestures is a common way of communication and we do a lot of hand gesture in our daily life. Imagine a presentation being done without the use of any remote or mouse control but using gestures. There are devices which are used to measure these hand gestures. The project is focused on the conversion of the extracted voltage fluctuation output from the device into a meaningful gesture.

Device and Data :

The device that is used for generating the data points is home made. We[1] have created a glove which uses flex sensor, accelerometer and gyroscope to measure finger and hand movements. The device uses Arduino2560 programmer for generation of the signals and writing it to any serial port.

Using the above configuration, around 10 training and test data (gestures) is generated.

ML Proposal:

The data from accelerometer and gyroscope are time series data points. Use of Dynamic Time Warping (DTW) algorithm seems to be more relevant in this case. DTW is used for measuring similarity between two temporal sequences. Let us assume you have made a gesture of the alphabet 'O'. Now to do that all you have to do a circle with your hand. This dimensions and speed of the gesture might vary from person to person but they are all similar. DTW can measure this similarity in a efficient manner. We will discuss this further along the path.

Finger movement are detected by flex sensors which has a range of value and more or less same for all person. Using all the three data types as one dimension does not make much sense. I am proposing a ensemble learning and gain more info along the progress.

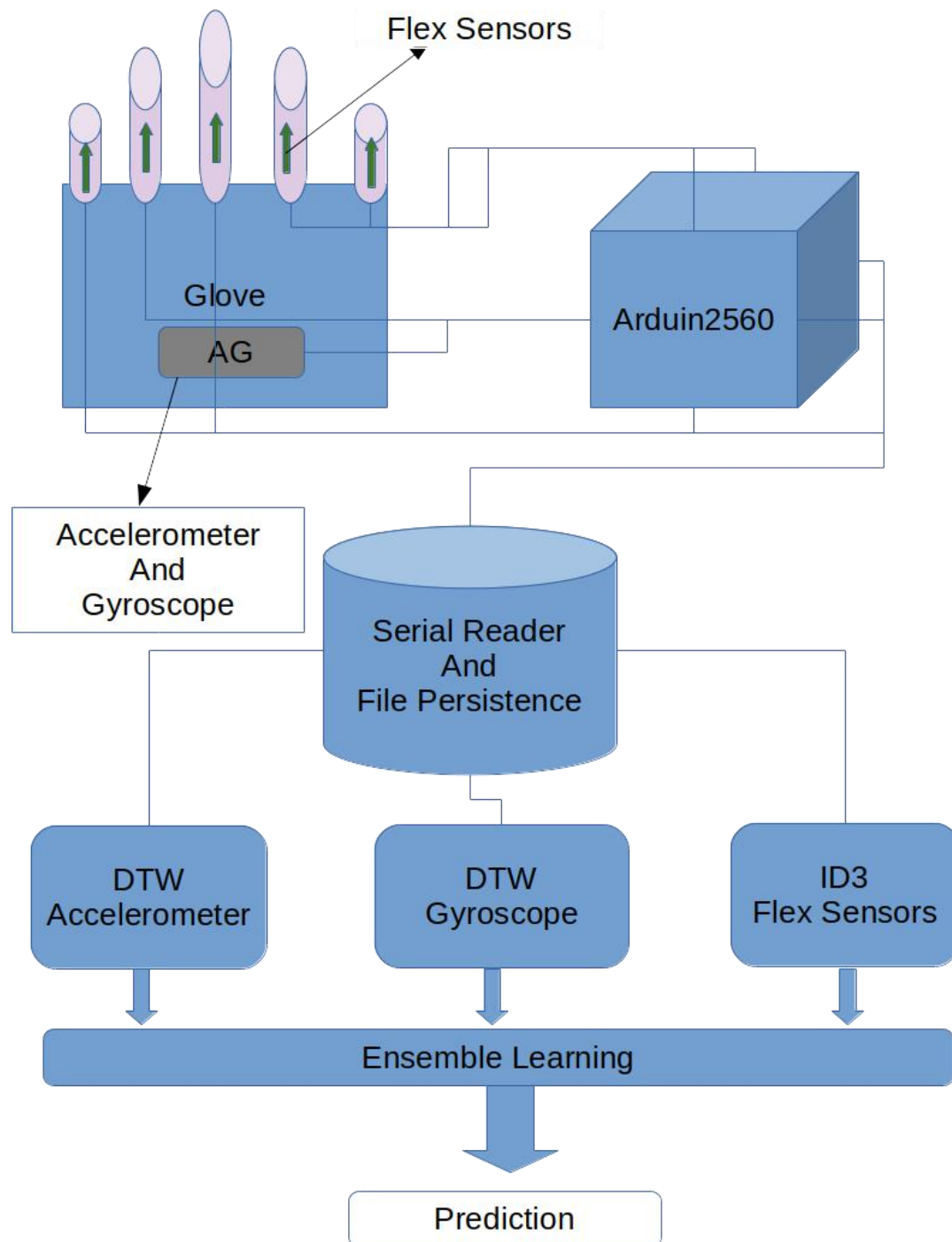
Data Metrics

$X = [3.61, 15.16, -4.81, -0.07, 1.07, 1.52, 99.00, 60.00, 54.00, 99.00, 2.00]$

First three values are x,y and z from accelerometer.

Next three values are angular velocity of x,y and z co-ordinates from gyroscope.

Last five values are finger position right from thumb to little finger (pinky) from flex sensors (5).



ML Description

As said earlier , the data will be saved as 11 dimension array and in a set of 10-20 rows. These are gesture data captured over a period of time. Below is a sample data set generated for the gesture hi ('hand of waving')

```
16.52,-1.22,4.37,-0.23,1.11,-0.35,20.00,22.00,14.00,27.00,3.00
16.22,-1.28,4.70,-0.15,0.43,0.09,21.00,23.00,15.00,27.00,3.00
16.20,-1.59,4.33,0.24,0.14,0.17,18.00,23.00,17.00,28.00,3.00
16.39,-1.97,3.75,0.00,0.17,-0.10,20.00,23.00,14.00,28.00,3.00
16.31,-1.46,4.14,-0.64,1.27,-0.95,20.00,24.00,16.00,27.00,3.00
16.20,-0.94,4.71,0.27,0.75,-0.87,20.00,23.00,15.00,26.00,3.00
15.96,-1.35,4.58,-0.39,0.37,-1.07,21.00,21.00,15.00,25.00,3.00
15.50,-1.26,8.57,5.29,-0.06,-0.64,30.00,27.00,15.00,31.00,3.00
15.97,0.12,4.80,4.49,-4.37,0.28,59.00,56.00,32.00,64.00,3.00
16.20,-2.01,2.18,2.31,-4.06,0.70,70.00,63.00,39.00,75.00,3.00
17.12,-0.82,1.10,-3.14,2.17,-0.64,66.00,56.00,34.00,64.00,3.00
16.84,-1.49,4.14,-4.03,7.53,-0.60,34.00,34.00,22.00,41.00,3.00
17.17,-1.75,4.82,0.46,3.69,-1.41,21.00,29.00,19.00,32.00,3.00
15.85,-0.82,5.02,-2.70,0.73,-2.21,27.00,30.00,18.00,32.00,3.00
16.74,-0.74,4.62,5.97,-3.19,-0.63,81.00,47.00,29.00,55.00,3.00
13.47,-0.19,2.75,3.46,-5.78,0.96,104.00,60.00,39.00,71.00,3.00
16.53,-0.88,2.63,-2.98,-0.87,1.61,101.00,57.00,39.00,68.00,3.00
16.89,-1.61,4.72,-4.14,2.85,0.18,60.00,40.00,24.00,47.00,3.00
18.23,-0.83,4.05,0.21,3.77,-0.92,27.00,33.00,20.00,35.00,3.00
16.23,-0.60,3.81,-0.60,3.85,-3.30,22.00,28.00,17.00,31.00,3.00
16.92,-1.17,6.31,2.68,3.86,-1.55,50.00,39.00,30.00,51.00,3.00
14.28,-0.74,1.10,3.75,5.88,-1.58,76.00,63.00,41.00,71.00,3.00
15.75,-0.56,4.79,-3.48,-1.63,1.40,69.00,54.00,32.00,61.00,3.00
17.77,-1.70,9.41,-5.03,0.02,-0.67,35.00,38.00,24.00,42.00,3.00
16.74,-0.82,5.74,1.45,-2.10,0.69,24.00,31.00,20.00,36.00,3.00
16.51,-1.19,4.69,0.21,-1.41,0.39,25.00,29.00,17.00,33.00,3.00
```

The data is split into accelerometer , gyroscope and flex output. Accelerometer and Gyroscope data trained with DTW and flex with a decision tree algorithm. The training of flex sensor output requires more study and might change in future.

The output from three different algorithms will be used to predict the type either through simple voting or by some other mechanism yet to decide.

References:

[1] The device was a result of a Hackathon competition within my organization. The creation of device includes many other participants including me. We could not able interpret the entire data set and finally settled with flex sensor. My effort is to continue where we left and use ML to train and predict the gesture using the same device.