Dumb Charades Project

Assignment 2

**Data Mining**

CSE 572 | Spring 2018

Submitted on:

03/23/2018

Submitted to:

Professor Ayan Banerjee

Ira A. Fulton School of Engineering

Arizona State University

Submitted by

GROUP 28

Tariq M Nasim (tnasim@asu.edu)

Akash Rathi (aarathi@asu.edu)

Shrinivas Bhat (sbhat10@asu.edu)

Prateek Jain(pjain43@asu.edu)

Madhulika Dolke(mdolke@asu.edu)

Table of Contents

[1 Introduction 3](#_Toc509615027)

[1.1 Phase 1: Data Collection 3](#_Toc509615028)

[1.2 Phase 2: Feature Extraction 3](#_Toc509615029)

[2 Task 1: Data Preparation 3](#_Toc509615030)

[2.1 Task 2: Feature Extraction 4](#_Toc509615031)

[2.1.1 Gesture 1: ABOUT 5](#_Toc509615032)

[2.1.2 Gesture 2: AND 6](#_Toc509615033)

[2.1.3 Gesture 3: CAN 7](#_Toc509615034)

[2.1.4 Gesture 4: COP 9](#_Toc509615035)

[2.1.5 Gesture 5: DEAF 10](#_Toc509615036)

[2.1.6 Gesture 6: DECIDE 11](#_Toc509615037)

[2.1.7 Gesture 7: FATHER 12](#_Toc509615038)

[2.1.8 Gesture 8: FIND 13](#_Toc509615039)

[2.1.9 Gesture 9: GO OUT 14](#_Toc509615040)

[2.1.10 Gesture 10: HEARING 15](#_Toc509615041)

[2.1.11 Summary of feature extraction method: 16](#_Toc509615043)

[3 Task 3: Feature Selection (PCA) 18](#_Toc509615048)

[3.1 Subtask 1: Arranging the feature matrix 18](#_Toc509615049)

[3.2 Subtask 2: Execution of PCA 19](#_Toc509615050)

[3.2.1 Gesture: ABOUT 19](#_Toc509615051)

[3.2.2 GESTURE: AND 20](#_Toc509615052)

[3.2.3 GESTURE: HEARING 21](#_Toc509615053)

[3.2.4 GESTURE: FIND 22](#_Toc509615054)

[3.2.5 GESTURE: DEAF 23](#_Toc509615055)

[3.2.6 GESTURE: DECIDE 24](#_Toc509615056)

[3.2.7 GESTURE: COP 25](#_Toc509615057)

[3.2.8 GESTURE: CAN 26](#_Toc509615058)

[3.2.9 GESTURE: GO OUT 27](#_Toc509615059)

[3.2.10 GESTURE: FATHER 28](#_Toc509615060)

[3.3 Conclusion 29](#_Toc509615061)

# Introduction

This is a project report for Assignment 2 of CSE 572 Data Mining course. The ultimate goal of this project is to develop a procedure to identify human gestures which includes:

1. identifying known gestures
2. segment sequence of gestures
3. identify unknown gestures.

There are three phases of this project.

## Phase 1: Data Collection

Gesture data was collected at the iMAPCT Lab at Arizona State University from 37 people doing 10 gestures 20 times each. The gestures performed are:

ABOUT, AND, CAN, COP, DEAF, DECIDE, FATHER, FIND, GO OUT, HEARING

A **Myo armband** was used to record the data. The device is normally used to control remote technologies wirelessly using various hand motions. This device uses following sensors to recognize gestures:

* Electromyographic (EMG) sensors that can sense electrical activity in the forearm muscles (EMG0L/EMG0R, EMG1L/EMG1R etc.)
* Gyroscope (GLX/GRX, GLY/GRY etc.)
* Accelerometer (ALX/ARX etc.)

## Phase 2: Feature Extraction

In this phase we implemented the feature extraction and feature selection methods on the raw data that was collected in Phase 1. The tasks of Phase 2 are described in the next section.

# Task 1: Data Preparation

For applying feature selection and feature extraction method, the data needed to be organized in a particular manner. To achieve this a MATLAB code was written to do the following:

* For each of the gestures 10 csv files (classes) were created.
* Each of them consisting of all the 20 actions of each person’s sensor data aligned row-wise.
* And time series data for each sensor column-wise. The class sample (ABOUT) csv contains the following data.

ABOUT Action 1 Acc X 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

ABOUT Action 1 Acc Y 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

.

.

.

ABOUT Action 2 Acc X 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

ABOUT Action 2 Acc Y 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

.

.

.

ABOUT Action 20 Acc X 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

ABOUT Action 20 Acc Y 2 3 4 5 1 5 1 6 2 7 8 3 2 1 3 ----------------

.

.

**ABOUT gesture csv file of 1 group.**

Thus each .csv file **contains - Number of rows - 20\*34** (number of actions \* number of sensors) and **Number of Columns** – **45 – 55 columns** for each person’s recording. So, 37 (number of group) were created. Arranging the data in this manner was helpful in task 2 to identify valid group’s recordings.

## Task 2: Feature Extraction

In order to use the above data for feature extraction method, we needed good sample time-series data (meaning the samples which have 45 column entries and which show significant variation). This was achieved using MATLAB code which writes the actions that have 45 time-series data into a new file and ignore the group that doesn’t meet these criteria.

The above criteria weren’t sufficient since not all groups, that had 45 time-series data, were having good nature of raw data. For example some of the sensor data were supposed to vary but did not. These groups were filtered out by the following steps:

* In MATLAB each person’s sensor data is represented by different color. Since MATLAB has only 7 colors, we plotted graphs in groups of 7.
* The plot to group mapping was done using legend feature of MATLAB.

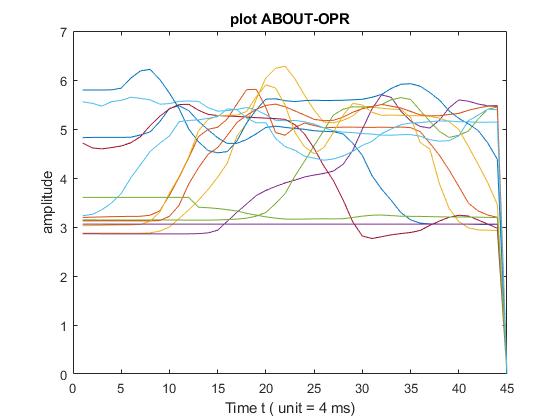
**Final valid groups:** Total **14 Groups** (01, 07, 09, 13, 16, 19, 20, 22, 23, 26, 27, 31, 32, 34).

These 14 groups have 20 instances each. For the ease of visualization we selected one of the 20 instances for each of the groups by looking at the plots of each of 20\*14 instances.

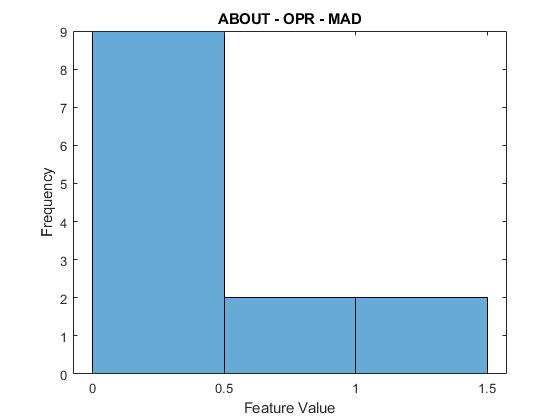
For each gesture, we have first selected a sensor that we expected to help us clearly identify that gesture. Then we plotted the raw data for that along with feature extracted out of the assembled data. Following list of gesture contains feature extraction to distinguish the gesture from others, along with the intuition to choose the feature extraction method and the plots.

### Gesture 1: ABOUT

1. We have selected the sensor ‘OPR’ and ‘EMG-R’ for ABOUT.
2. Reason: OPR sensor gives us the orientation pitch data of the right hand. For ‘ABOUT’, since right hand comes in-front of the chest and makes a roll movement. We thought due to changes in the orientation of the right hand has unique pattern among other gestures.
3. Following is the figure that contains all the 14 Groups plot for ABOUT.



**Fig 1. ABOUT-OPR raw data.**



**Fig 2. ABOUT-OPR MAD data.**

1. Our intuition was right: It is evident from the fig 1 that right hand will be stationary initially and then due to rolling action OPR reading goes high in the middle and then gradually come back to zero meaning resting position. To extract this amplitude change, we plotted median absolute deviation as in Fig. 2. Where, we can clearly see high amplitude feature value occurs for low number of time, indicating rolling movement of right hand for short duration.

### Gesture 2: AND

1. We have selected the sensor ‘ARX’ for AND.
2. Reason: Here our intuition says that the movement of right hand along only X-axis can be distinct feature as compared to other sensor, since the action involves moving of right hand from left to right side of the chest.
3. Following is the figure that contains all the 14 groups plot for AND.

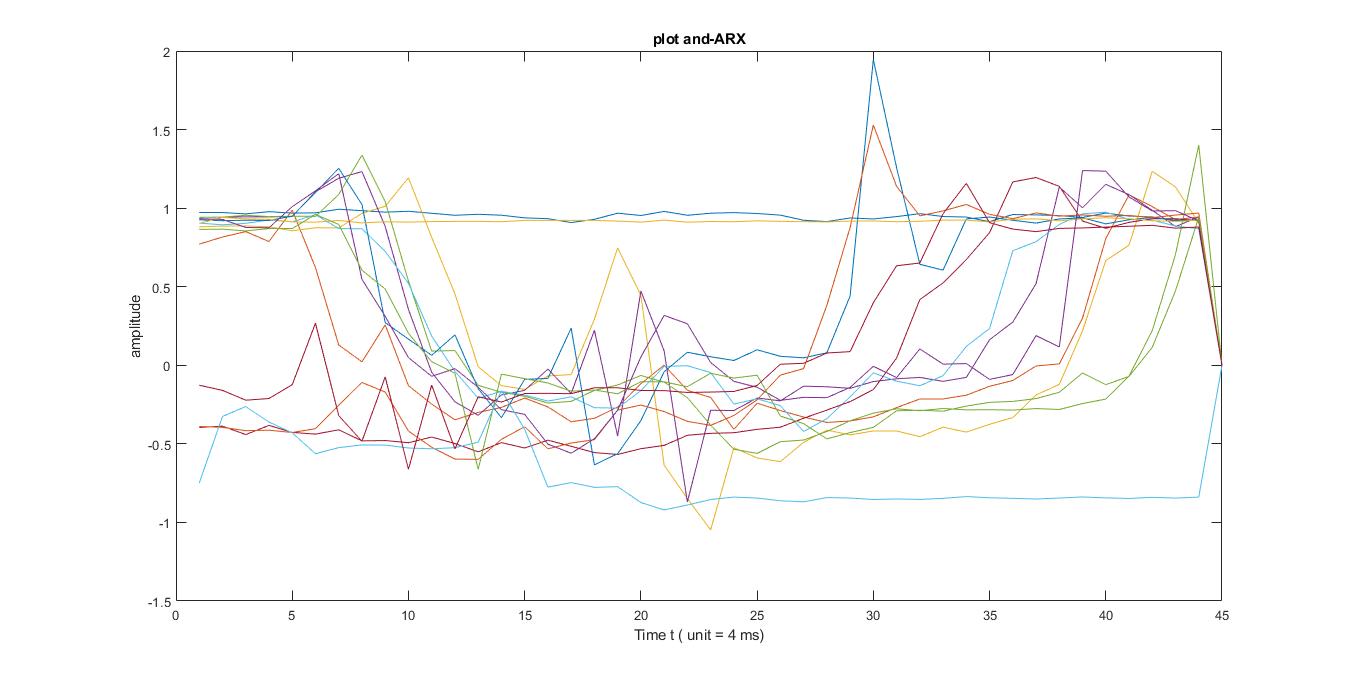


Fig 3. AND-ARX raw data.

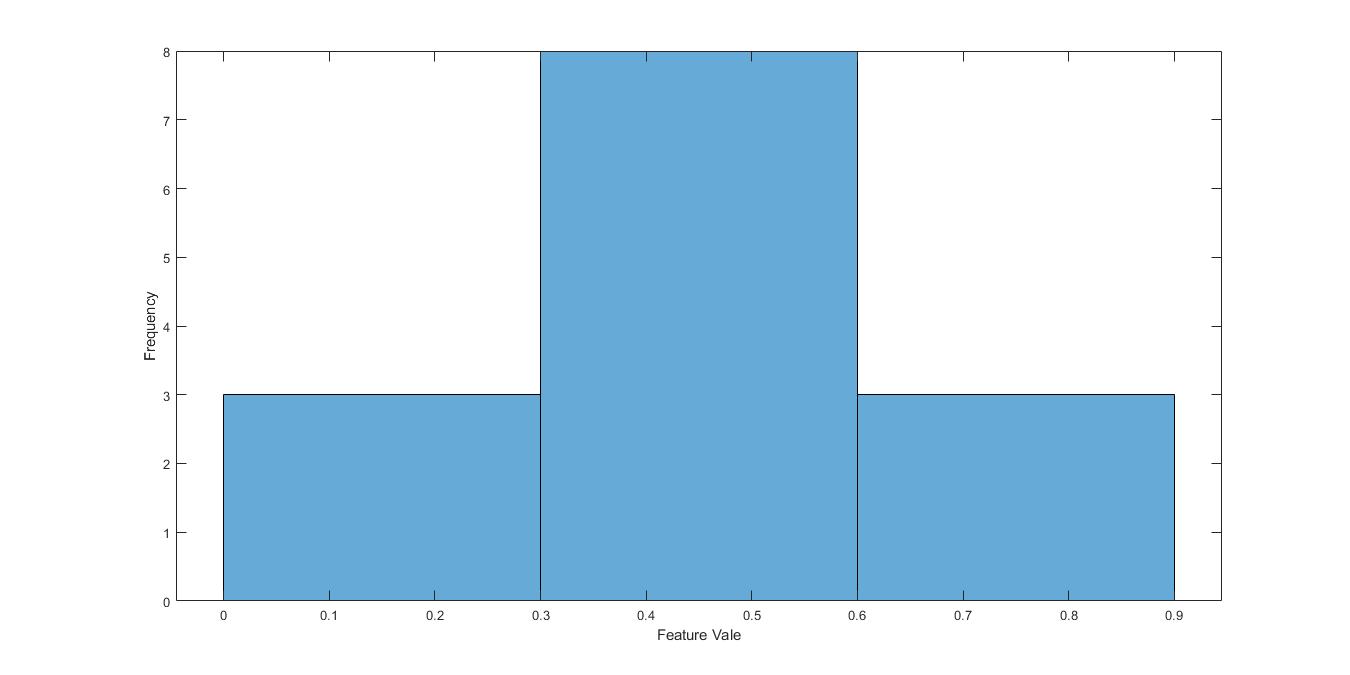
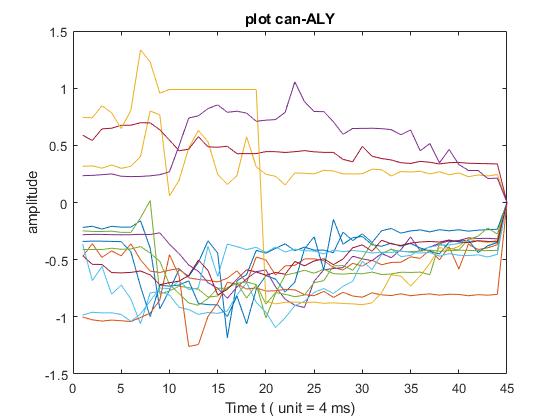


Fig 4. AND-ARX variance plot

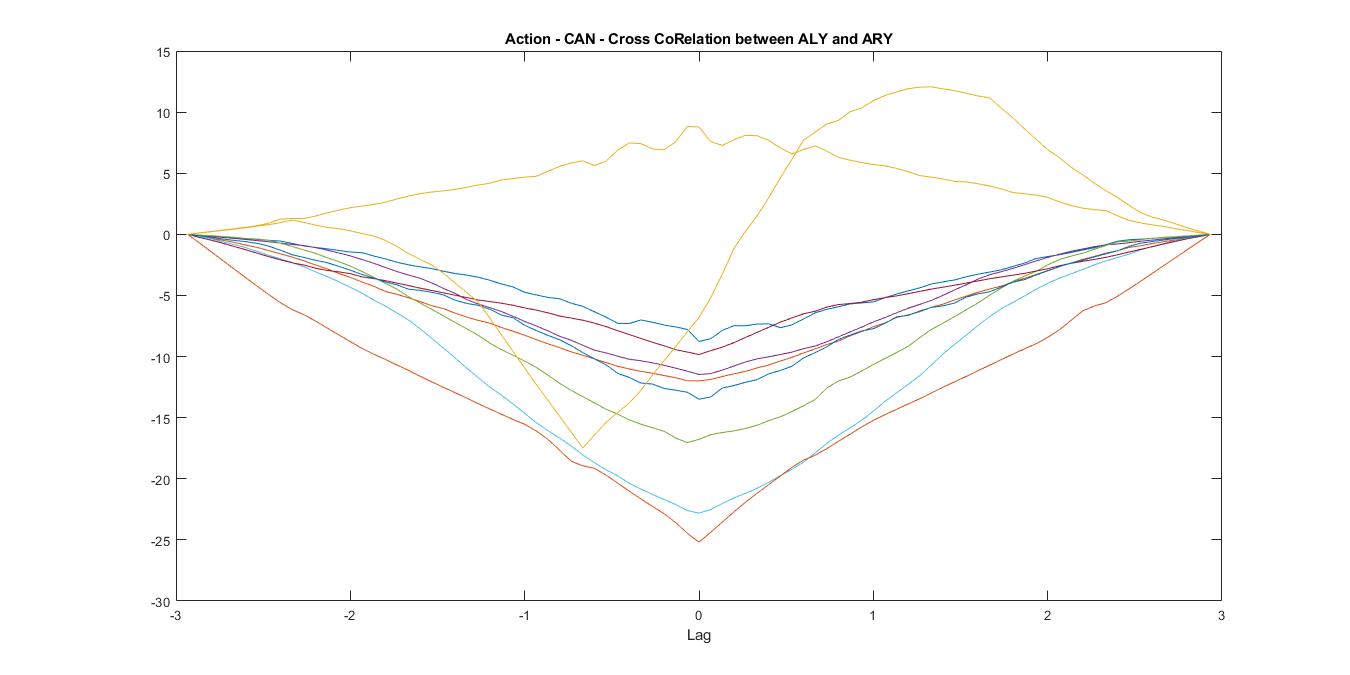
1. Our intuition was right: We can see that from fig 3, the right hand along x axis remains stationary and moves with increasing acceleration and gradually decelerate before coming to stand still. Plotting variance histogram graph it’s evident that for constant-moderate acceleration frequency is more meaning distance covered is more. For immediate acceleration (from stand still to motion)and deceleration (from motion to stand still) number of frequency is less indicating brief movement.

### Gesture 3: CAN

1. We have selected the sensor ‘ALY’ and ‘ARY’ for CAN.
2. Reason: Our intuition was that for left hand, acceleration along Y axis should give a Gaussian curve structure. Because movement of left hand is very rare among the other gestures, even if there is any movement it cannot be like Gaussian curve.
3. Following is the figure that contains all the 14 groups plot for CAN.



**Fig 5. CAN-ALY raw data.**

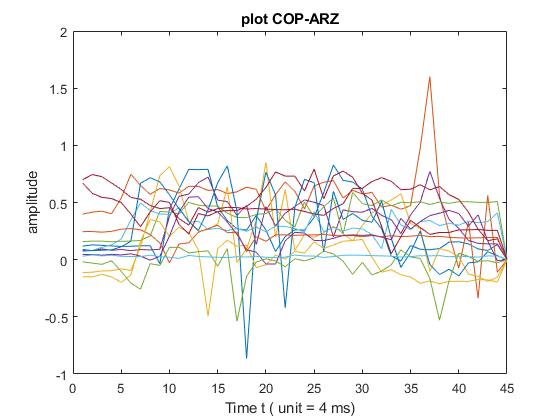


**Fig 6. Can - Cross correlation between ALY-ARY.**

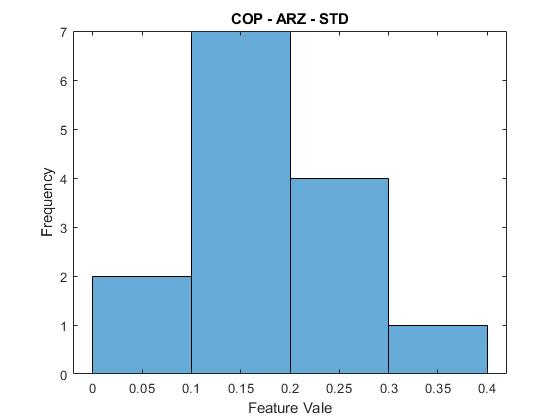
1. Our intuition was wrong: Since most of the groups data collected neither resembled the right hand accelerometer data (ARY), nor followed the Gaussian curve (As shown in Fig 5). But plotting cross correlation between ALY and ARY as shown in Fig 6 gave us the lag between ALY and ARY which are least correlated when delay was 0 and are greatly correlated otherwise.

### Gesture 4: COP

1. We have selected the sensor ‘ARZ’ for COP.
2. Reason: We figured out that the sign has periodic movement of right hand along z axis for short duration.
3. Following is the figure that contains the entire 14 groups plot for COP.



**Fig 7. COP-ARZ raw data.**

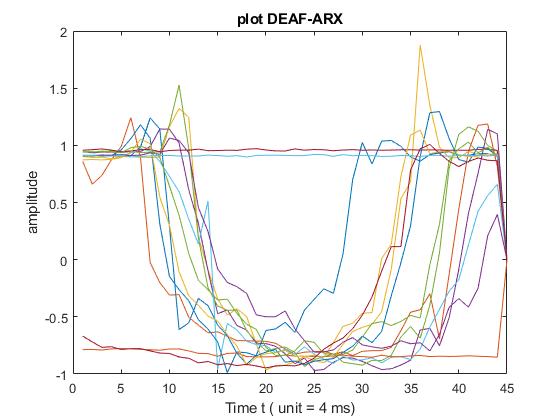


**Fig 8. COP-ARZ standard deviation.**

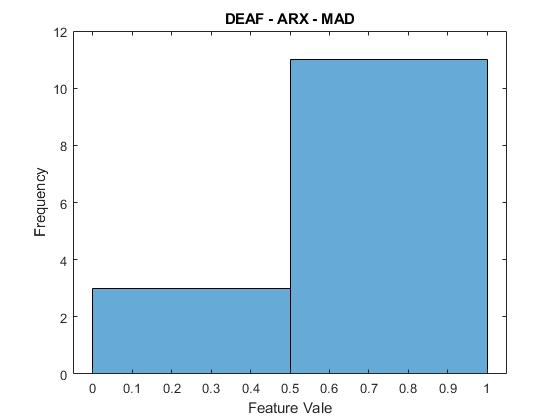
1. Our intuition was right: It is evident from standard deviation plot (as shown in Fig 8) of ARZ that 78.5% of people were found to have ARZ values between 0.1-0.3 (meaning periodic hand movement was of short range). Thus ARZ feature can be extracted using Standard deviation extraction method for COP gesture.

### Gesture 5: DEAF

1. We have selected the sensor ‘ARX’ for DEAF.
2. Reason: Since the right index finger moves in x direction (from mouth to cheek) we can see the data variation in ARX sensor as shown in Fig 9.
3. Following is the figure that contains the entire 14 groups plot for DEAF ARX.



**Fig 9. DEAF – ARX raw data**

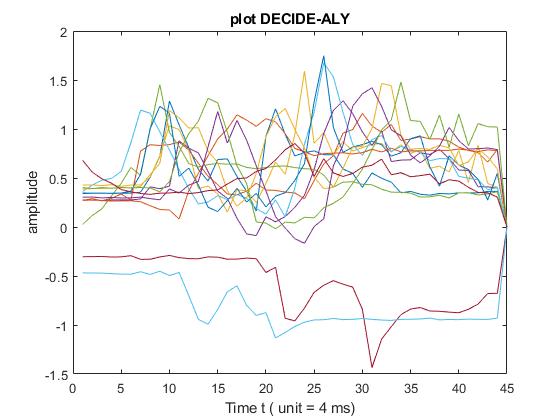


**Fig 10. DEAF – ARX MAD**

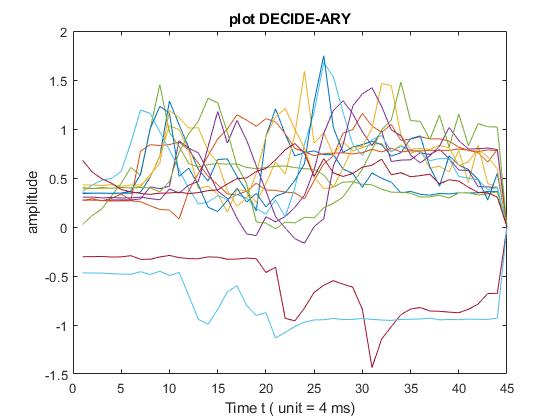
1. Our intuition was right: As shown in Fig 10 we applied MAD (mean absolute deviation) on the ARX sensor data and found that 78.5% of people were found to have ARX amplitude between 0.5 – 1. Thus applying MAD on ARX was useful to identify ‘DEAF’ gesture.

### Gesture 6: DECIDE

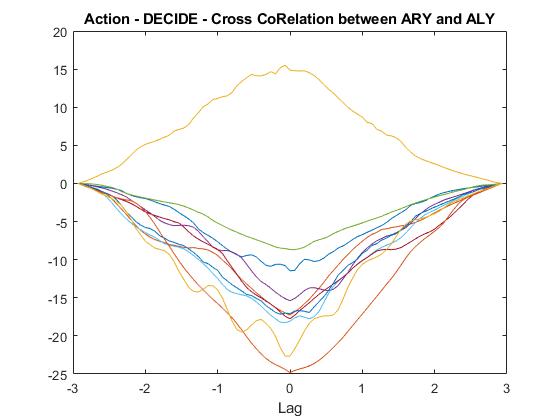
1. We have selected the sensor ‘ARY’ and ‘ALY’ for DECIDE.
2. Reason: Since this action requires movement of right hand negative Y axis and movement of left hand along positive Y axis.
3. Following is the figure that contains the entire 14 groups plot for DECIDE.



**Fig 11. DECIDE-ALY raw data.**



**Fig 12. DECIDE-ARY raw data.**

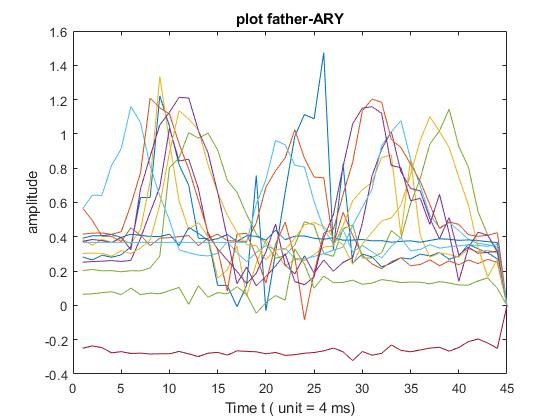


**Fig 13. DECIDE ARL-ARY cross correlation.**

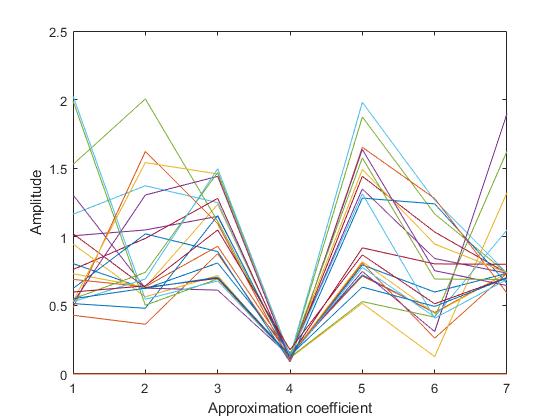
1. Our intuition was right: The cross correlation between ARY and ALY sensors should be in negative since they move in opposite direction. This is evident in Fig 13 – cross correlation between ARY and ALY.

### Gesture 7: FATHER

1. We have selected the sensor ‘ARY’ for FATHER.
2. Reason: The ASL sign has right hand movement along both Y and Z axis. Hence we decided to observe the ARY data as shown in Fig 14.
3. Following is the figure that contains the entire 14 groups plot for FATHER.



**Fig 14. FATHER-ARY raw data.**

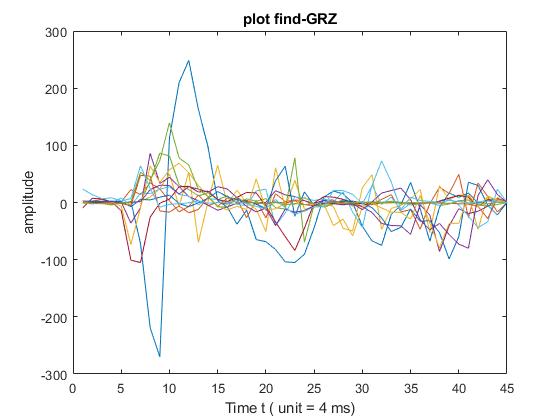


**Fig 15. FATHER-ARY DWT plot.**

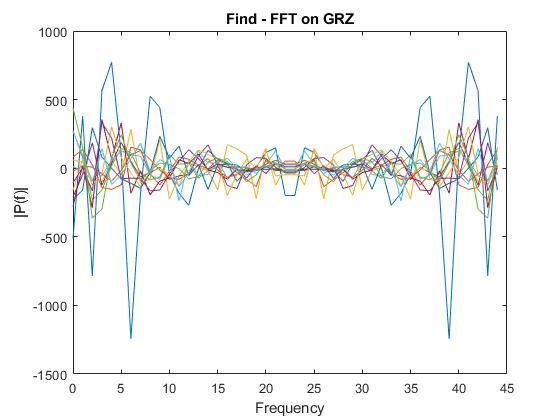
1. Our intuition was right: The movement captured in Fig 14 contains two peaks that indicate movement along Y axis is rapid initially and the gradually slows down to become a periodic wave and again increases. By applying DWT (Discrete wavelet transformation) with ‘haar’ wave filter (Haar wavelet), we could identify the two major peaks as shown in Fig 15.

### Gesture 8: FIND

1. We have selected the sensor ‘GRY’ for FIND.
2. Reason: As per ASL sign the right hand has spinning motion. So we concluded that gyroscopic sensor would be useful. Since the movement is along Z axis we chose GRZ sensor for this gesture.
3. Following is the figure that contains the entire 14 groups plot for FIND.



**Fig 16. FIND-GRZ raw data.**

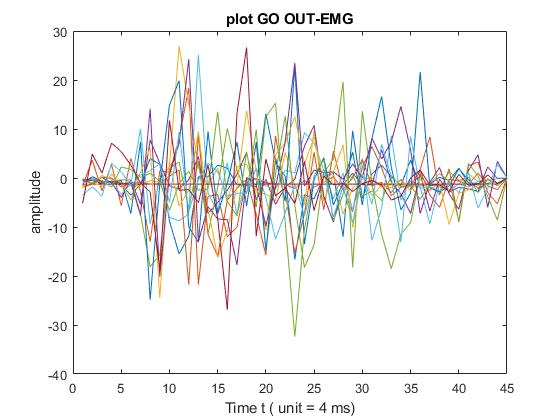


**Fig 17. FIND-GRZ FFT data.**

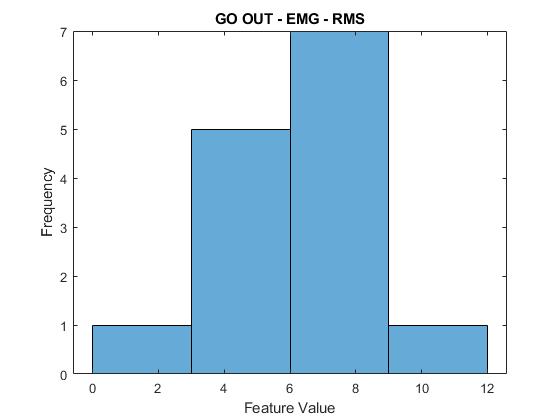
1. Our intuition was partially right: The periodic movement of GRZ data had frequency variation, hence, by applying FFT for the raw GRZ data we could observe that 2 peaks (where the first peak represents frequency of the signal). Thus the hand movement was found to be periodic depicting ‘FIND’ action.

### Gesture 9: GO OUT

1. We have selected the sensor ‘EMG-R’ for GO OUT.
2. Reason: The gesture involves a lot of muscle contraction. So, using a EMG sensor for right hand can provide us with lot of information about the gesture. So, EMG-R was opted.
3. Following is the figure that contains the entire 14 groups plot for ‘GO OUT’.



**Fig 18. GO OUT – EMGR raw data**

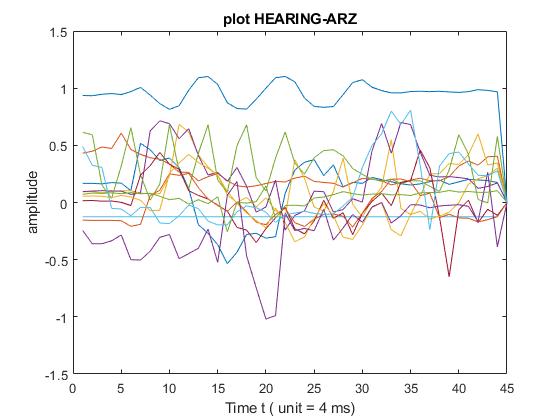


**Fig 19. GO OUT – EMGR RMS (Root mean square) data**

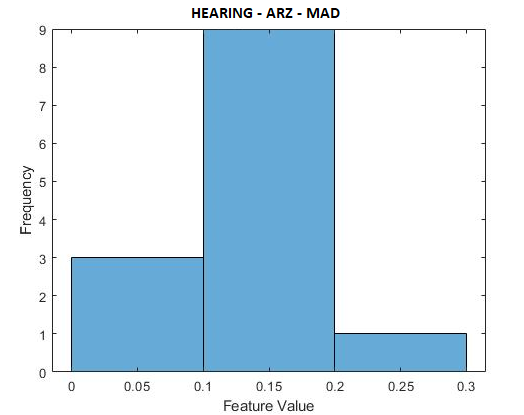
1. Our intuition was right: From fig 19 we could observe that 85.7% of the people were found to have RMS value of EMG between 3-9 hence this feature can be used to identify GO-OUT gesture.

### Gesture 10: HEARING

1. We have selected the sensor ‘ARZ’ for HEARING.
2. Reason: As per ASL sign the movement of right index finger is periodic and varies along Z axis. The raw data plot for ARZ is found to be periodic as shown in Fig 20.
3. Following is the figure that contains the entire 14 groups plot for ‘HEARING’.



**Fig 20 HEARING-ARZ raw data.**



**Fig 21 HEARING-ARZ MAD (Median absolute data) data.**

1. Our intuition was right: Since the median absolute value for the raw data of ARZ sensor had 64% of people having frequency in the range (0.1-0.2) which indicated that ARZ varies in the range of 0.1-0.2 for most of the people, hence this feature depicts the HEARING gesture.

### Summary of feature extraction method:

The feature extraction methods used are as follows:

#### Standard Deviation

This measure the variation in data values, a low STD can be useful since most of the data is accumulated near the mean, which is the expected value.

Since we are applying standard deviation on one sensor row (vector), it gives a scalar data as output.

MATLAB syntax: S = std(A,w,dim);

Where, A is vector on which STD is applied,

w is weighting scheme when w=0 ,S is normalized by N-1,

dim to select column(1) or row(2).



#### Mean Absolute Deviation

This gives the mean of the absolute deviation from the middle value.

Here m(X) is mean value of the vector X.

MATLAB syntax: Y = mad(X,0);

Where, Y is the mean absolute deviation,

X is the vector on which MAD has to be applied.

#### Median Absolute Deviation

This gives the mean of the absolute deviation from the middle value.

Here m(X) is the **median** value of the vector X.

MATLAB syntax: Y = mad(X,1);

Where, Y is the mean absolute deviation,

X is the vector on which MAD has to be applied

#### Root Mean Square

Root mean square the square root of mean of the squares of vector X.

MATLAB Syntax: y = rms(x,dim) calculates the RMS of vector X along the dimension.

#### Fast Fourier Transform

FFT samples a periodic signal over time and divides it into its frequency components. The FFT algorithm computes the Discrete Fourier Transform of the sequence. It can be used to identify the frequency of the periodic signal.

MATLAB Syntax: Y = fft(X,n,dim)

Where X is a vector, n is n point DFT, dim is the dimension along which transformation is applied wave.

#### Discrete Wavelet Transform

Wavelets are wave-like oscillation with an amplitude that begins at 0. Wavelet transform are discretely sampled wavelets. The key advantage of DWT over FFT is that it has temporal resolution.

MATLAB Syntax: [cA,cH,cV,cD] = dwt2(X,wname)

Where X is a vector and wname is a analyzing wavelet.

# Task 3: Feature Selection (PCA)

## Subtask 1: Arranging the feature matrix

The Dimensionality reduction technique – Principle component Analysis takes only one matrix. So, we have to arrange our feature matrix in such a way that the new feature matrix can be obtained by simply multiplying the eigen vector with old feature matrix.

For each gesture, there is a separate feature matrix. The way the feature matrix is created is shown in table 1 below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gestures | Feature1 | Feature2 | Feature3 | Feature4 |
| ABOUT | MAD on OPR 1:1 | RMS on OPR 2:2 | MAD on EMG0R 3:3 |  |
| DEAF | MAD on ARX 1:1 | VAR on EMG0l 2:2 | MAD on GRX 3:3 | MAD on ORR 4:4 |
| AND | VAR on ARX 1:1 | VAR on ARY 2:2 | MAD on EMG0R 3:3 | MAD on EMG2R 4:4 |
| HEARING | VAR on OPR 1:1 | RMS on ARZ 2:2 | STD on GRX 3:3 | MAD on ORR 4:4 |
| CAN | VAR on EMG0R 1:1 | VAR on EMG0L 2:2 | MAD on ALY 3:3 | MAD on ARY 4:4 |
| COP | STD on ARZ 1:1 | VAR on OPR 2:2 | MAD on OPR 3:3 |  |
| DECIDE | VAR on ALY 1:1 | VAR on ARY 2:2 | MAD on EMG3L 3:3 | MAD on ORR 4:4 |
| GO OUT | RMS on EMG0R 1:1 | VAR on ARY 2:2 | MAD on GRX 3:3 | MAD on ORR 4:4 |
| FIND | FFT on GRY 1:55 | VAR on ARX 56:56 | MAD on EMG0R 57:57 | MAD on EMG2R 58:58 |
| FATHER | DWT on ARY 1:28 | RMS on OPR 29:29 | MAD on EMG0R 30:30 |  |

Table 1: Arrangement of Feature Matrix

## Subtask 2: Execution of PCA

The eigen vectors of the top two principal components are shown in the following plots generated using MATLAB’s biplot function.

**Scree Plot** – It is used to plot the variance among all the principle components and useful for finding which component is most useful.

**Bi-Plot –** It is used to visualize the magnitude of variables contribution to the principle components.

### Gesture: ABOUT

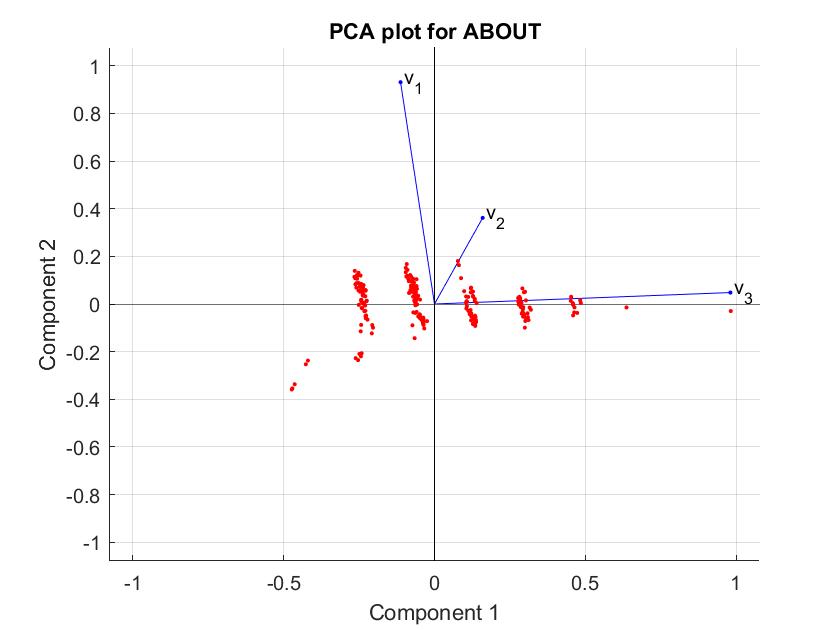


Fig 22: PCA plot for ABOUT

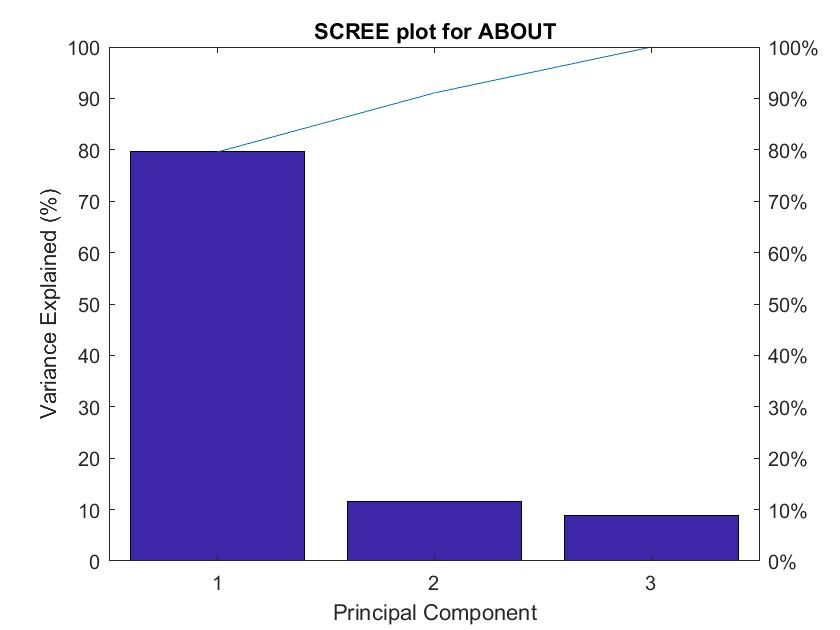


Fig 23: Scree plot of ABOUT

#### Making Sense of PCA:

From the biplot we can find most variance along the vector v3, which represents the feature “MAD on EMG0R”. Also, from the scree plot we can also observe that the first component represents 80% of the variance. Using this component, we can reduce the dimensionality.

#### Result of PCA

The results show that out of all the extracted features, MAD of EMG0R is the principal component.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using MAD on OPR. MAD on EMG0R turns out to be a better feature for the gesture ‘ABOUT’.

### GESTURE: AND

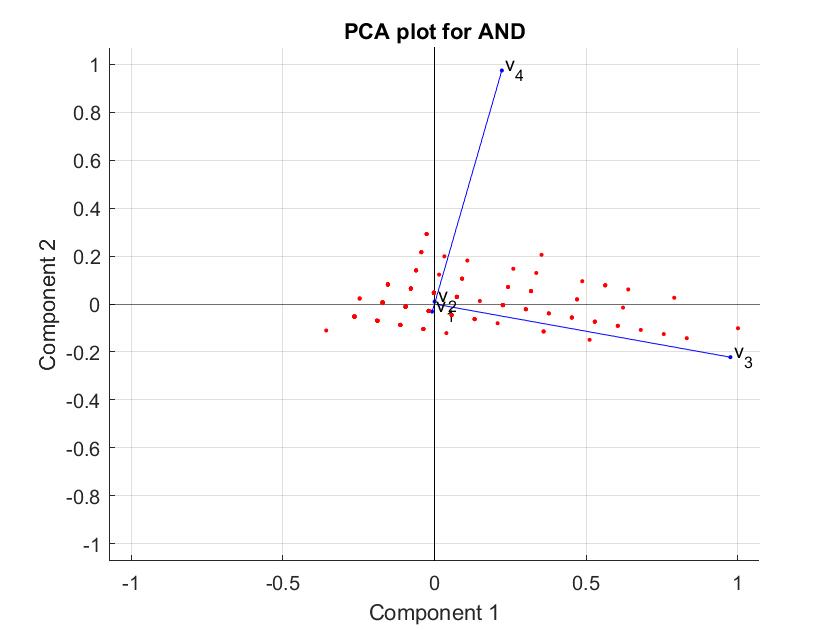


Fig 24: PCA plot of AND

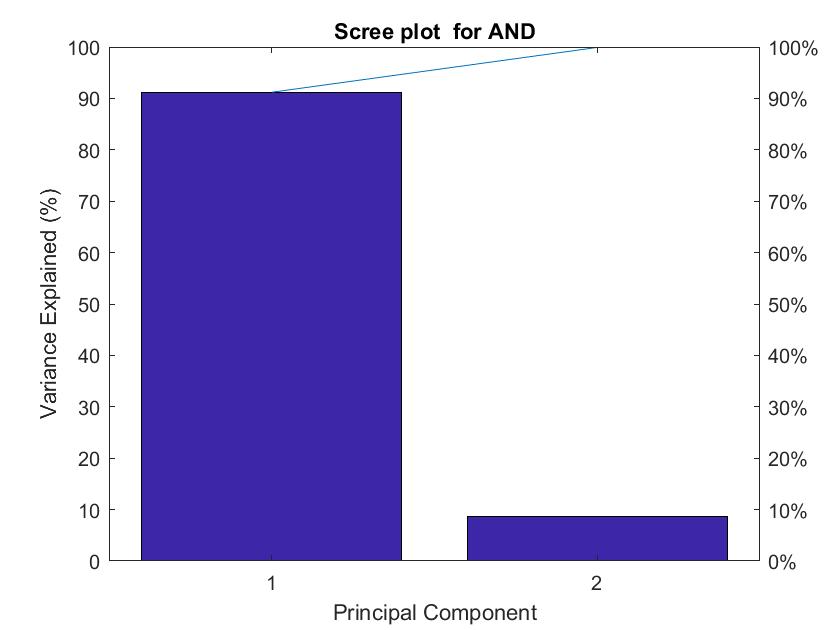


Fig 25: Scree plot of AND

#### Making Sense of PCA:

From the biplot we can see most variance along the vector v3, which represents the feature “MAD on EMG0R”. Also, from the scree plot we can also observe that the first component represents 90% of the variance. Using this component, we can reduce the dimensionality.

#### Result of PCA

The results show that out of all the extracted features, MAD of EMG0R is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using VAR on ARX. But “MAD on EMG0R” turns out to be a better feature for the gesture ‘AND’.

### GESTURE: HEARING

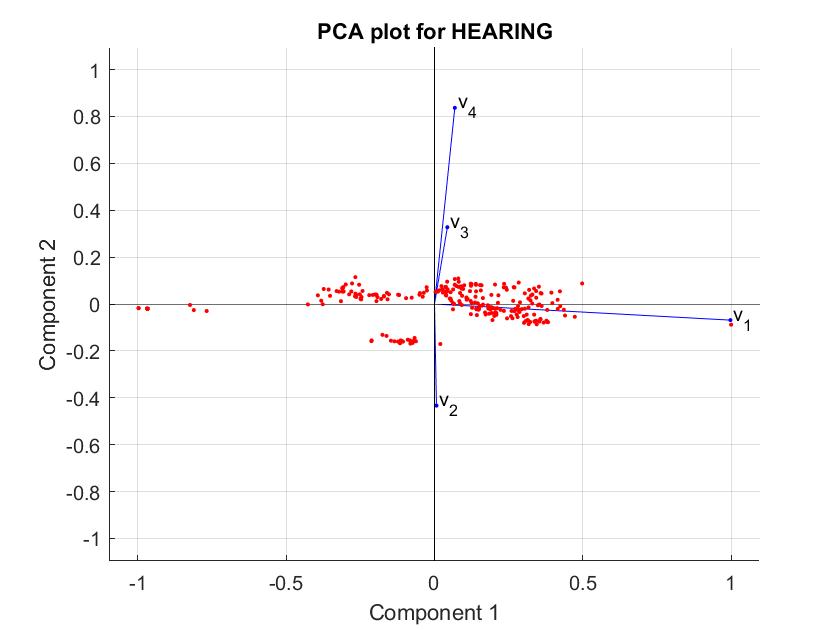


Fig 26: PCA plot of HEARING

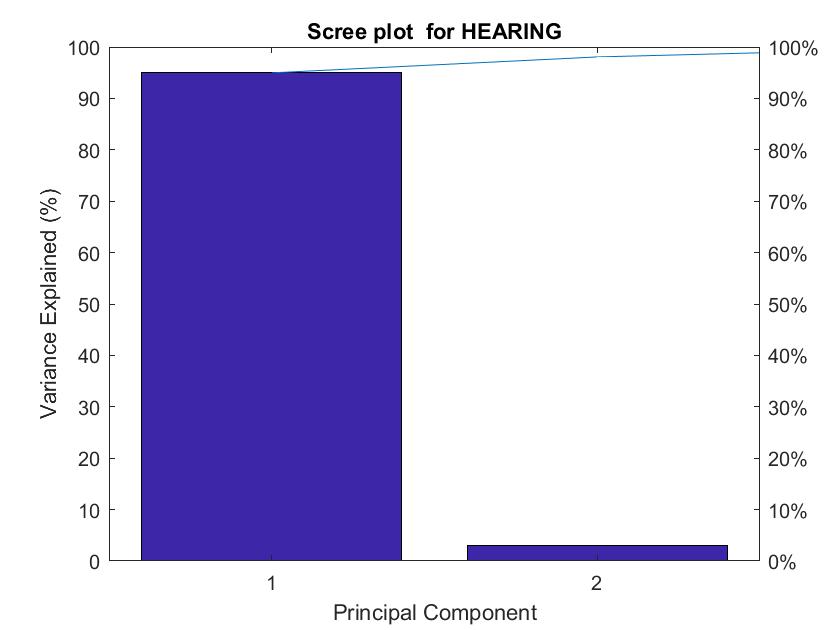


Fig 27: Scree plot of HEARING

#### Making Sense of PCA:

From the biplot we can see most variance along the vector v1, which represents the feature “VAR on OPR”. Also, from the scree plot we can observe that the first component represents 95% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, VAR on OPR is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using RMS on ARZ. But “VAR on OPR” turns out to be a better feature for the gesture ‘HEARING’.

### GESTURE: FIND

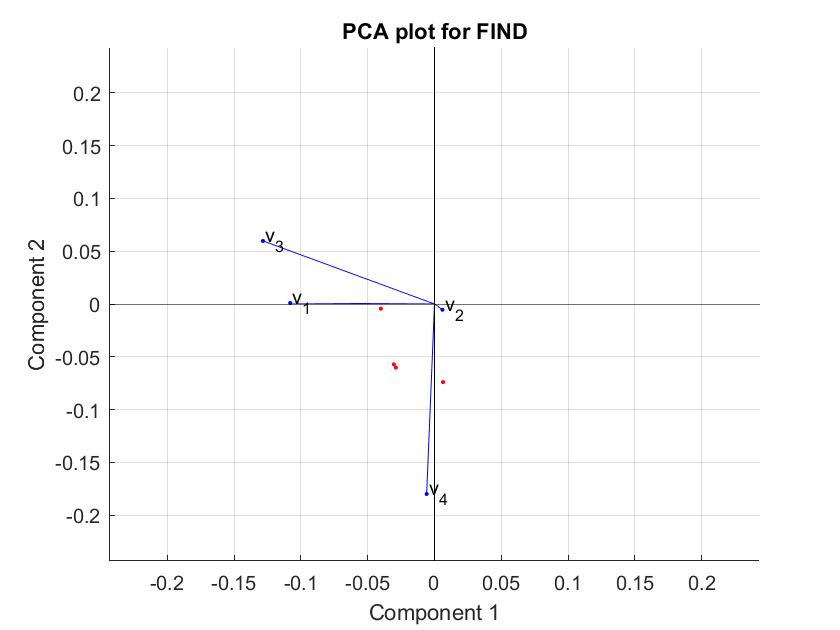


Fig 28: PCA plot of FIND

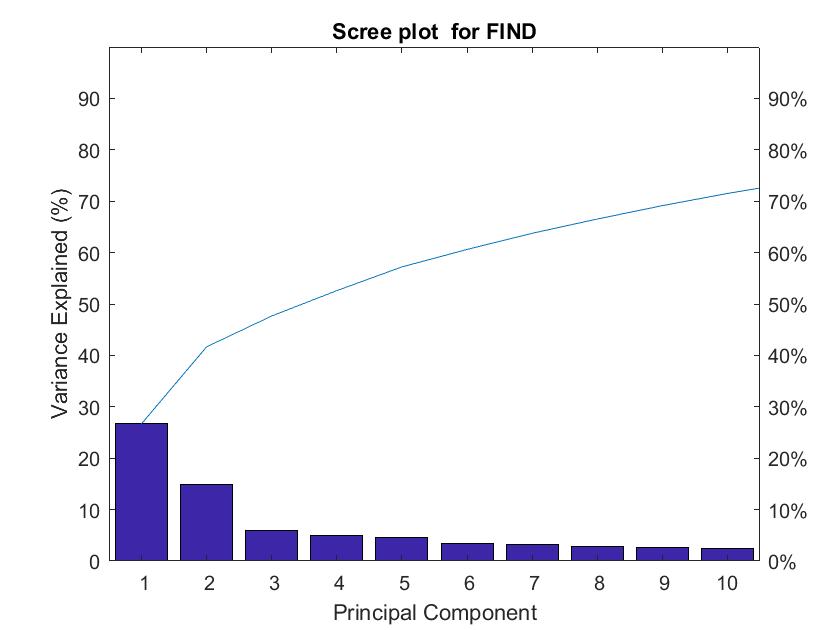


Fig 29: Scree plot of FIND

#### Making Sense of PCA:

From the biplot that most variance can be seen along the vector v4, which represents the feature “MAD on EMG2R”. Also, from the scree plot we can also observe that the first two components represents 40% of the variance. Using these components, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, MAD on EMG2R is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using FFT on GRY. But “MAD on EMG2R” turns out to be a better feature for this gesture.

### GESTURE: DEAF

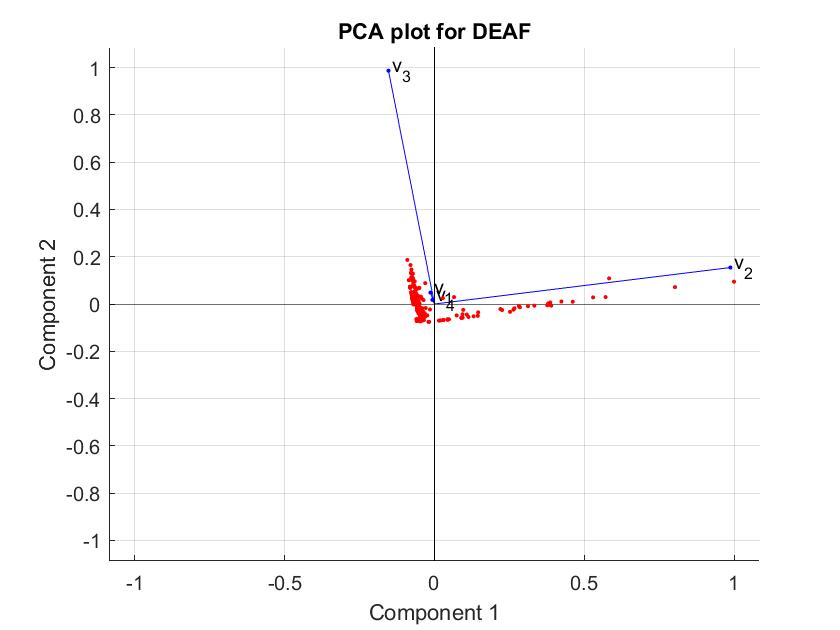


Fig 30: PCA plot of DEAF

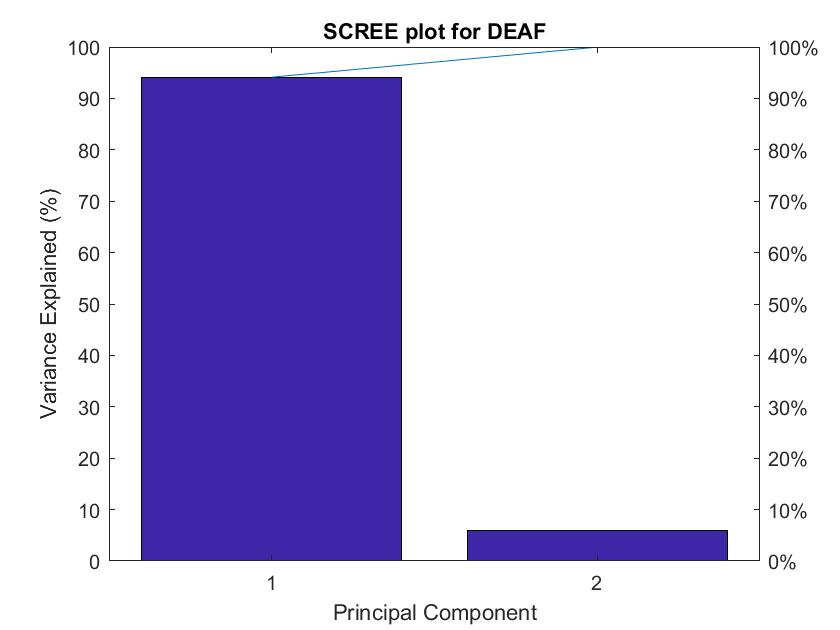


Fig 31: Scree plot of DEAF

#### Making Sense of PCA:

From the biplot we can see most variance along the vector v2, which represents the feature “VAR on EMG0L”. Also, from the scree plot we can observe that the first component represents nearly 95% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, VAR on EMG0L is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using MAD on ARX. But “VAR on EMG0L” turns out to be a better feature for the gesture ‘DEAF.

### GESTURE: DECIDE

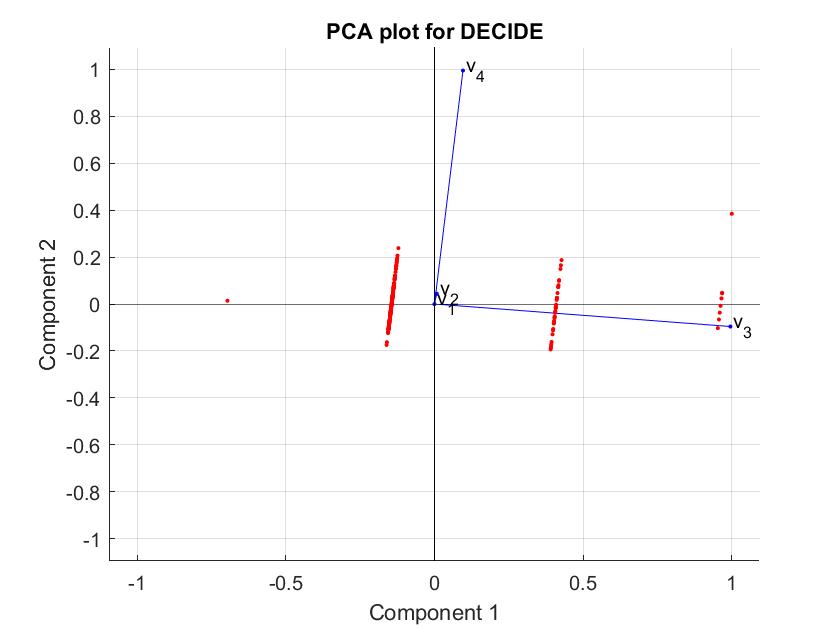


Fig 32: PCA plot of DECIDE

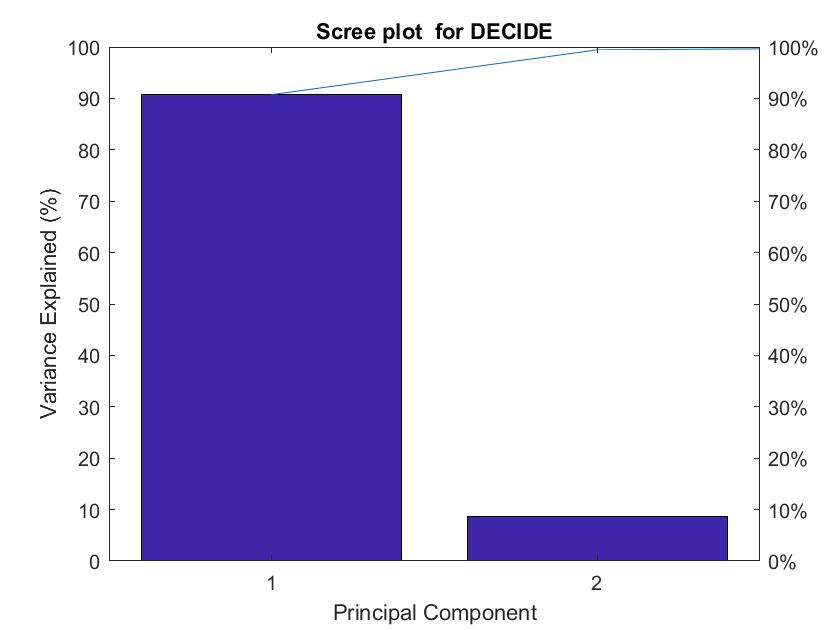


Fig 33: Scree plot of DECIDE

#### Making Sense of PCA:

From the biplot we can see most variance along the vector v3, which represents the feature “MAD on EMG3L”. Also, from the scree plot we can observe that the first component represents nearly 90% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, MAD on EMG3L is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using the sensors ALY and ARY. But “MAD on EMG3L” turns out to be a better feature for the gesture ‘DECIDE’.

### GESTURE: COP

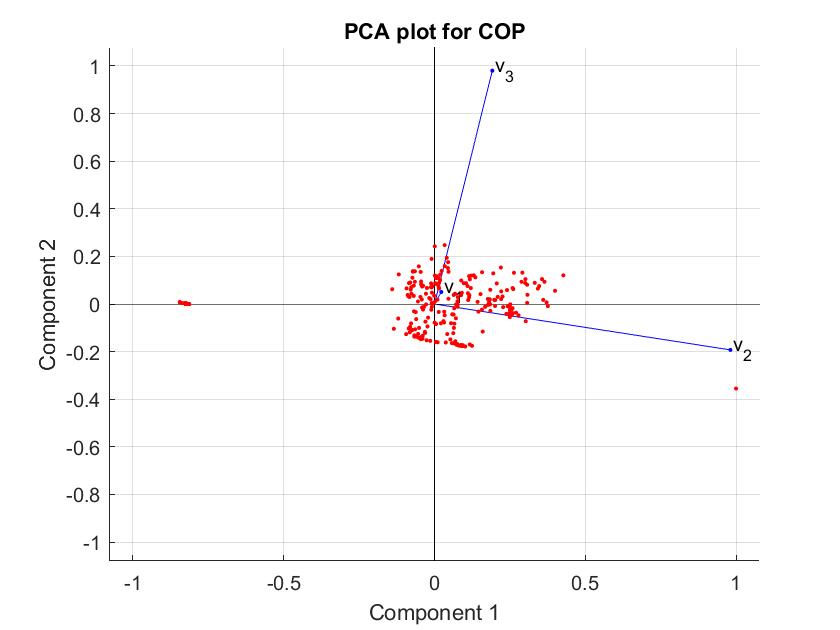


Fig 35: PCA plot of COP

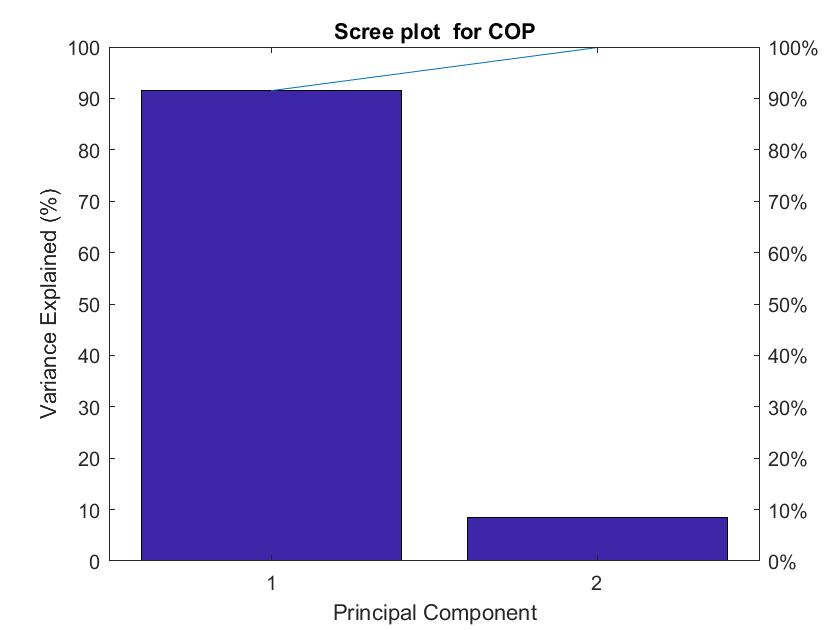


Fig 36: Scree plot of COP

#### Making Sense of PCA:

From the biplot we can see most variance along the vector v2, which represents the feature “VAR on OPR”. Also, from the scree plot we can observe that the first component represents nearly 90% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, VAR on OPR is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using STD on ARZ. But “VAR on OPR” turns out to be a better feature for the gesture ‘COP’.

### GESTURE: CAN

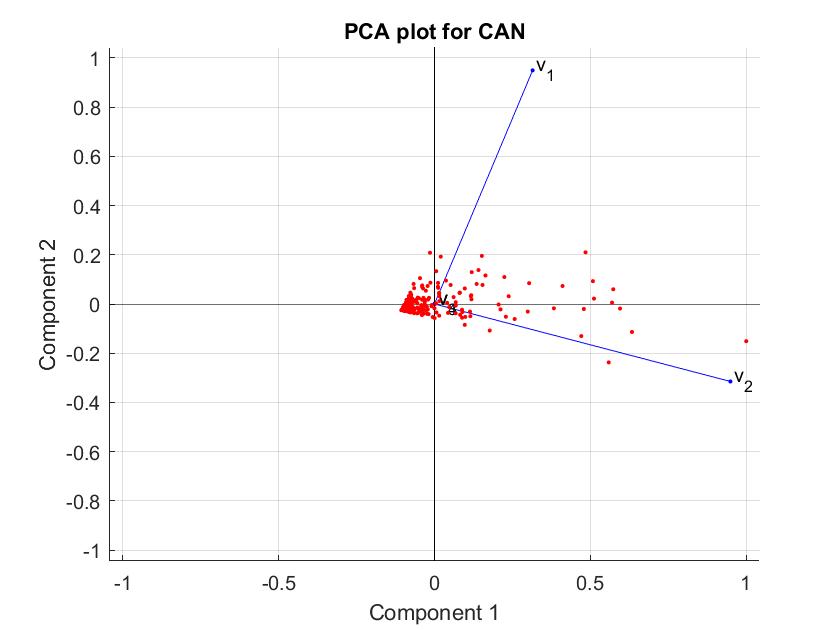


Fig 37: PCA plot of CAN

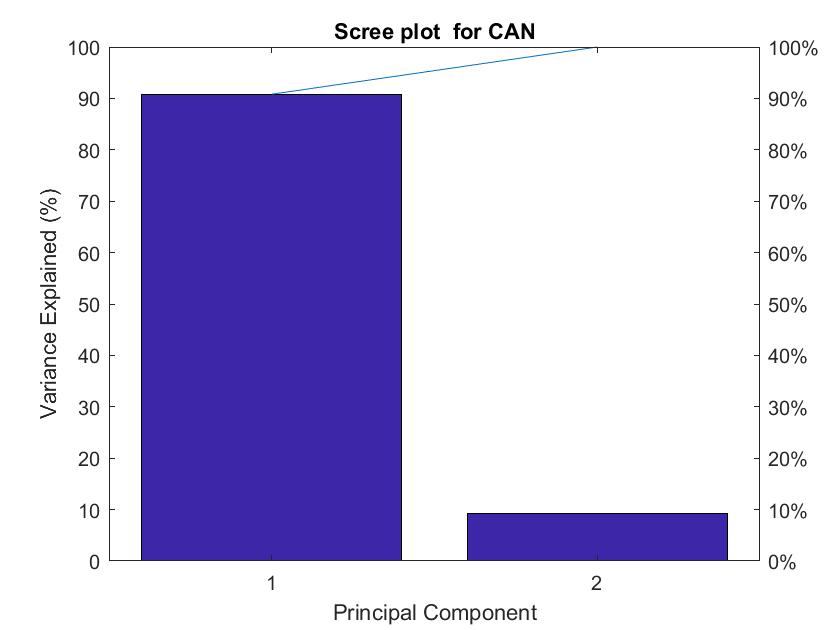


Fig 38: Scree plot of CAN

#### Making Sense of PCA:

From the biplot we can see most variance along the vectors v1 and v2, which represents the feature “VAR on EMG0L” and “VAR on EMG0R”. Also, from the scree plot we can observe that the first component represents nearly 90% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, VAR on EMG0L is the principal component for this gesture.

#### Argument about the result

Doing PCA was helpful because it corrected our initial intuition of using the sensors ALY and ARY. But “VAR on EMG0L” turns out to be a better feature for the gesture ‘CAN’. This shows that, both the features have the potential to identify this gesture.

### GESTURE: GO OUT

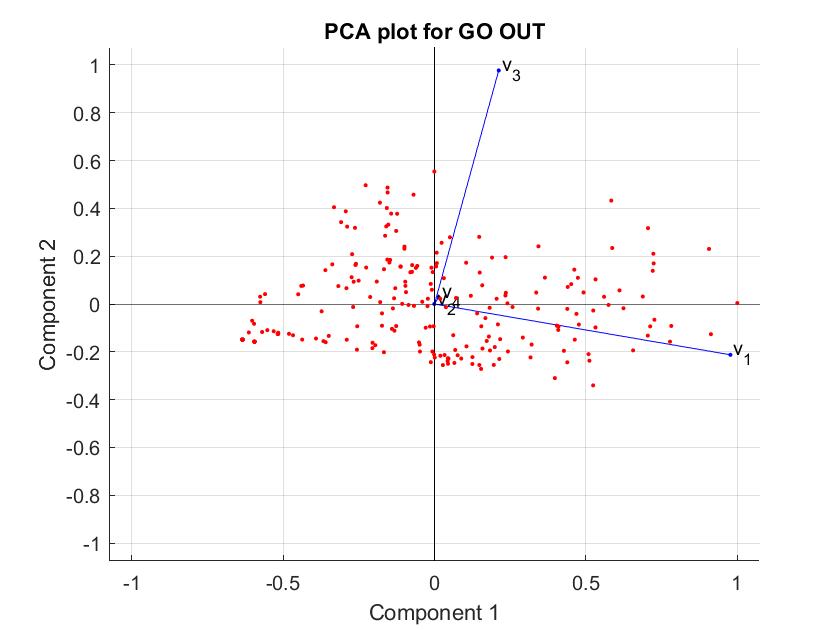
****

Fig 39: PCA plot of GO OUT

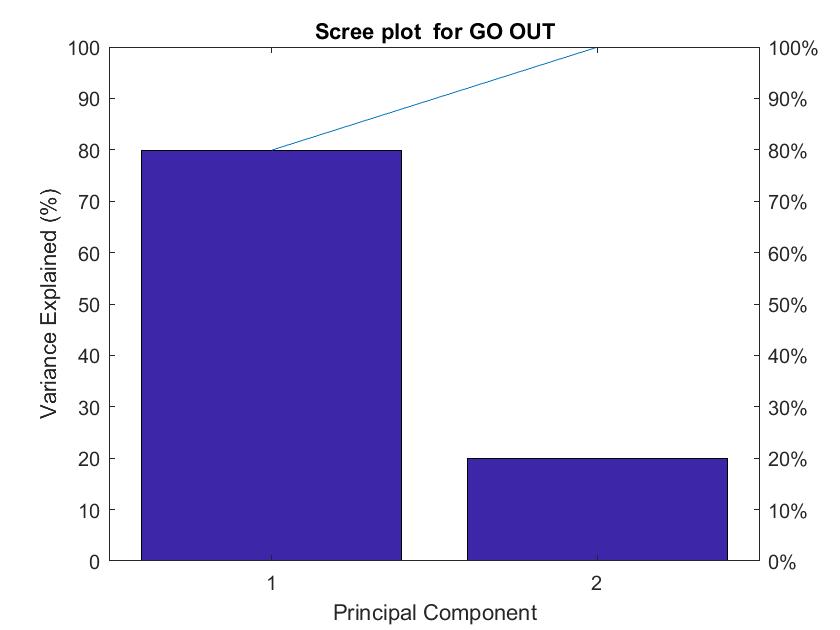


Fig 40: Scree plot of GO OUT

#### Making Sense of PCA:

From the biplot we can see most variance along the vectors v1, which represents the feature “RMS on EMG0R”. Also, from the scree plot we can observe that the first component represents nearly 80% of the variance. Using this component, we can reduce the dimensionality significantly.

#### Result of PCA

The results show that out of all the extracted features, RMS on EMG0R is the principal component for this gesture.

#### Argument about the result

The results of PCA shows that our initial intuition of using the “RMS on EMG0R” was correct.

### GESTURE: FATHER

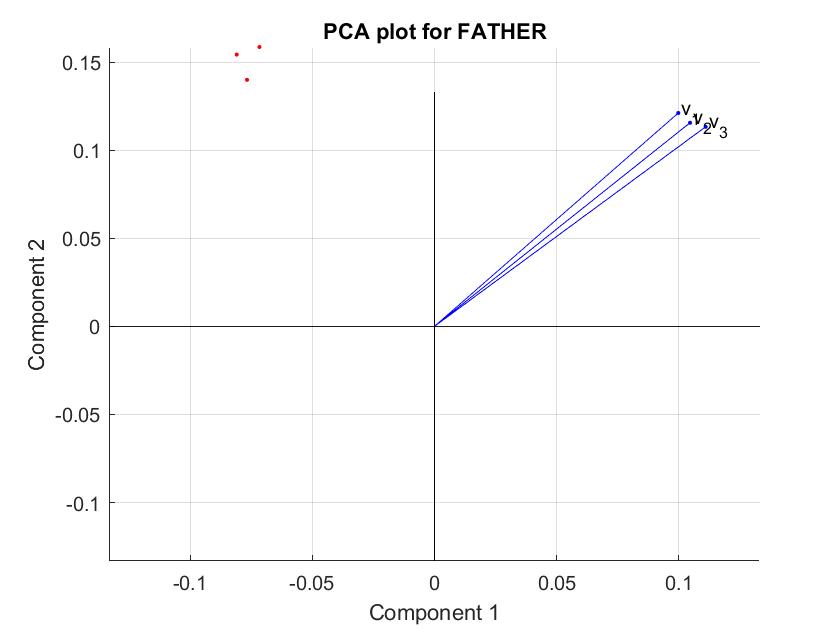


Fig 41: PCA plot of FATHER

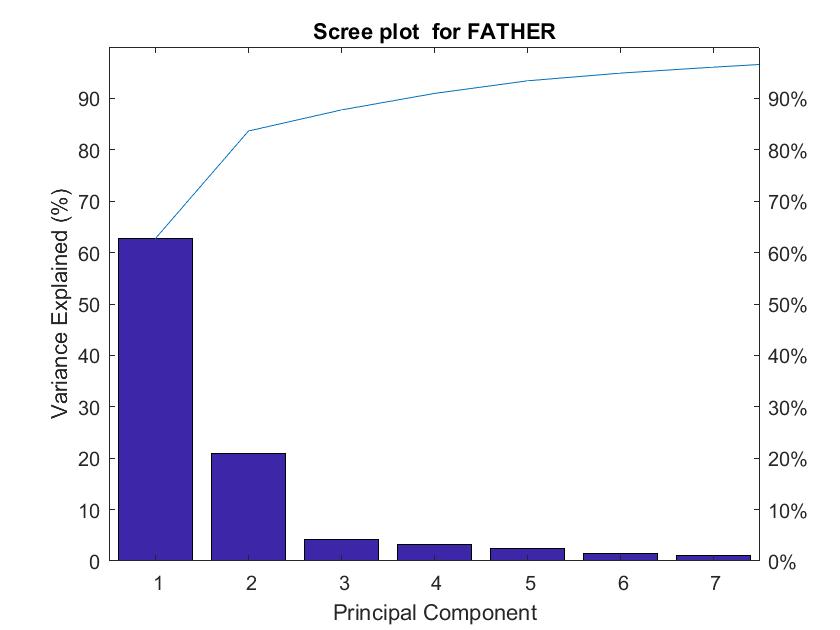


Fig 42: Scree plot of FATHER

#### Making Sense of PCA:

From the biplot we can see most variance along the vectors v1, v2 and v3 which represents the feature “DWT on ARY”. Also, from the scree plot we can observe that the first two components represents nearly 80% of the variance.

#### Result of PCA

The results show that all three components has same variance.

#### Argument about the result

Doing PCA was not very helpful because we cannot make a decision on the first three vectors (since they represents only the first three columns of the first feature, i.e. DWT on ARY). Probably for this, we should have plotted all the 30 columns of the feature matrix instead of the first three.

## Conclusion

Out of the 10 gestures, our intuition was correct for only one gesture (GO OUT).

From this phase of the project we have learnt several things.

Arranging data properly helps us analyze the data in later stages.

Initial human guess on the sensors might not come out to be true after applying the feature extraction and dimensionality reduction.