Human Gait Analysis

Report submitted in fulfillment of the requirements for the Exploratory Project of
Second Year, IDD.

By:
Utsav Krishnan

Supervisor:
Dr. Hari Prabhat Gupta

April 18, 2017

Department of Computer Science and Engineering

Indian Institute of Technology (BHU) Varanasi

<u>Certificate</u>

*This is to certify that the work contained in this report entitled "**Human Gait Analysis**" being submitted by **Utsav Krishnan (Roll No. 15074014)**, carried out in the Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, is a bona fide work of our supervision.*

Place: IIT (BHU) Varanasi
Date:April 18, 2017

Dr. Hari Prabhat Gupta
Department of Computer Science and Engineering,
Indian Institute of Technology (BHU) Varanasi,
Varanasi, INDIA 221005.

Abstract

In this project, we compare the biometric gait of the human body using accelerometer sensor present in the smartphone.

The accelerometer data are acquired while the phone is in the trousers pocket in fixed orientation.

Data is then analyzed in time domain where average gait cycle is extracted for both legs, and their similarity is evaluated using Dynamic Time Wrapping.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

Biometrics refers to metrics related to human characteristics. Biometric authentication authentication is used in computer science as a part of identification and access control[1].

There being mainly two sub discipline namely behavioral and physiological. Behavioral characteristics deals with gait, voice, rhythm and the latter deals with body size and shape like fingerprint, iris, DNA, etc[1].

In this project work we focus on gait analysis. Human gait can be analyzed different way, the common being image processing and accelerometer based, this latter being used for this project..

1.2 Motivation of the Research Work

Biometric analysis can be of great use in medical field. There has been news of doctors using Fitbit to monitor patients heart beat in hospitals. It could be of great value to people who needs to keep track of their heart beat.

Although there are many gadget similar to that, but they are costly and common people doesnt considers investing in it. But almost every people have smartphone and even the basic ones these days have built in accelerometer. Here we use android app which logs accelerometer readings and process it to understand gait cycle and compare it with other leg.

Variation arise in gait pattern of both legs when a person is ill or tired/not feeling well. This app can inform the subject if he/she needs to consult a doctor and more importantly being cost effective.

1.3 Organisation of the Report

- Data acquisition and storage
- Noise removal and smoothing
- Feature template extraction and arriving at conclusion

Chapter 2

Data Acquision and Storage

2.1 Data acquisition

Data is acquired by accelerometer present in the smartphone. Data logging is triggered by the subject by pressing the start button and the smartphone is kept in the trousers pocket, in a fixed orientation. Such that the +Y axis, +Z axis and X axis of the smart phone points in top,forward,sideways direction respectively.

Then the subject walks considerable distance with the phone kept in the pocket with each leg.

2.2 Storage

The data is saved in txt format, which can be easily read on the computer. Two options are provided for storing accelerometer data, either in internal storage or in the external storage.

Each time data is recorded the data log file is overwritten. We use sensor delay fastest for logging data. All three axis reading are stored in the log file.

1016	-2.2146366	11.829751	0.9911995
1032	-2.1308396	11.346122	1.0486604
1036	-1.7262194	11.020511	0.93134445
1055	-0.78769237	10.630256	0.59376204
1076	-1.4628572	10.311827	0.34955344
1077	-2.255338	9.792285	0.6488287
1078	-1.5490485	10.156203	0.92895025
1086	-2.753332	8.719683	0.8523358
1097	-2.918532	8.966286	1.2497733

Figure 2.1 The first, second, third and fourth columns of the data log files are time, X-axis, Y-axis and Z-axis respectively.

The app can also be used for standalone data logger.

Only Z-axis is used here for further analysis. In this orientation we could have used both Z and Y axis effectively, but using X-axis is avoided because the time series graph obtained doesnt provide meaningful gait cycle feature template.

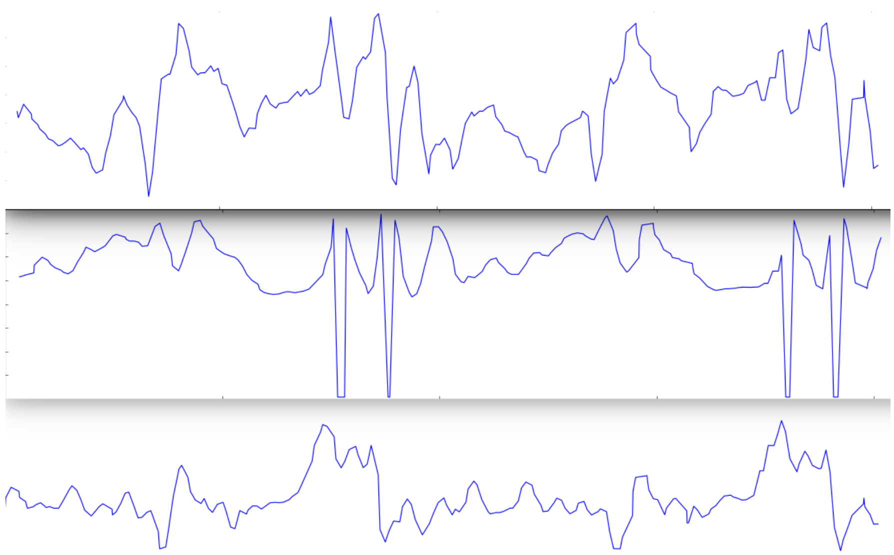

Figure 2.2 Raw X, Y, Z axis readings, from top to bottom.

Chapter 3

Noise reduction and smoothing

3.1 Noise removal

Following low-pass filter is used for removing high frequency noise from the data.

The extent of filtering is set by the value of alpha as shown in the equation, the value of alpha represents the weight given to the previous data point called the inertia from the previous output.

```
for i from 1 to m
    y[i] := y[i-1] * (1-alpha) + x[i] * alpha
```

[2]

3.2 Smoothing

There being various method of smoothing, moving average was selected for our analysis, as it smooths the graph while maintaining the characteristic of the graph.

After many iterations the window size of 5 was chosen, if a large value of window size had been used then the variations could have not been so obvious. After experimenting it was also found out that running moving average 3-5 times with small

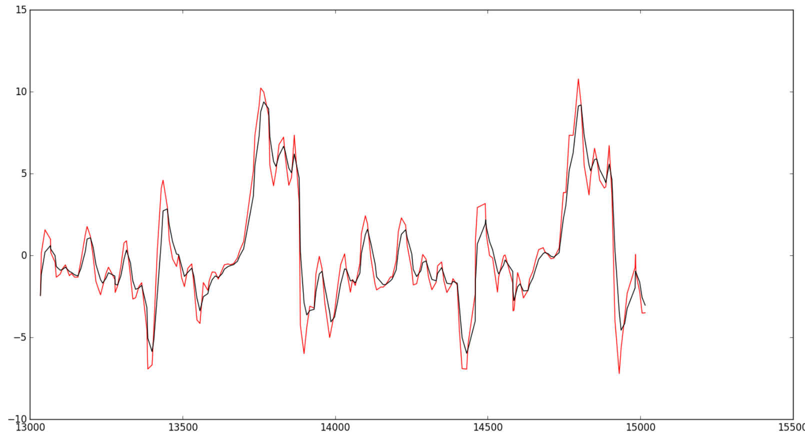

Figure 3.1 Raw z-axis(red) readings vs low-pass filtered(black) data.

window size proved to be better than running 1-2 window average with large window size as the gait cycle period points are short.

The average of window size of data points are taken as average and the value is set to the median element. The weight given to every points in the window size over summation is the same.

$$X_{\frac{m+n}{2}} = \frac{X_m + X_{m+1} + \cdots + X_n}{n - m + 1} = \frac{1}{n - m + 1}\sum_{i=m}^{n} X_i \quad [3]$$

3.3 Normalization

Data is then normalized to reduce the distance between the maximum and minimum of gait cycle points while preserving the characteristics of the graph.

$$y_i = \frac{x_i - min(x)}{max(x) - min(x)} \quad [4]$$

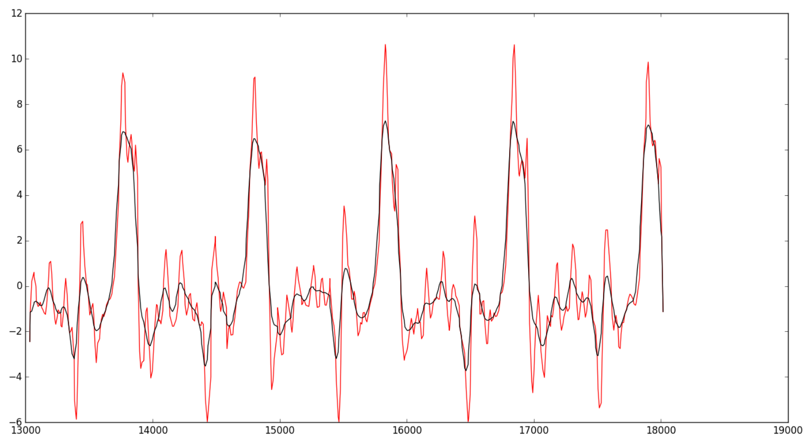

Figure 3.2 Moving Average applied on low-pass filtered data.

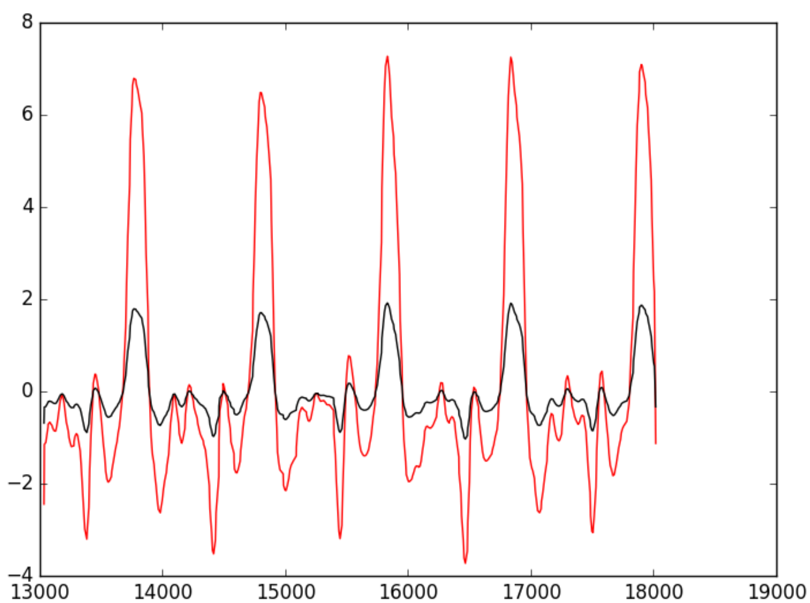

Figure 3.3 Normalization done on moving averaged data.

Chapter 4

Feature Extraction

4.1 Maximal Points

First all the points with local maxima are extracted, the points which are greater than its previous and next value. Then the points with value less than a threshold value is discarded. The threshold value is obtained from the equation[.], where u is mean, sd is standard deviation and k is the whole number value set by the user. The value of k is found out by trial and error method, and depends from user to user.

4.2 Gait Cycle Points

The remaining points forms the gait cycle. cycle is sometimes called the walking cycle. It exists between any two consecutive heel strike of the same leg[5].

4.2.1 Average Gait Cycle

Average Gait Cycle(AGC)[5] is defined as shown in the following equation.

$$AGC = \left\{ gaitcycle_k \mid \quad d_k = argmin\left(\frac{1}{N}\sum_{j=1}^{N} dtw(gaitcycle_k, gaitcycle_j)\right) \right\} \text{ [5]}$$

The term average could be misleading here as we are not taking average of all gait

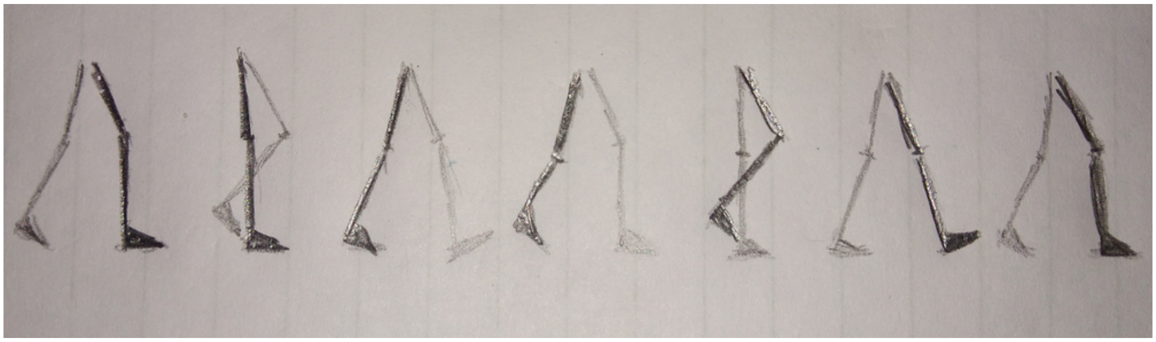
[5]

Figure 4.1 Single Gait Cycle.

cycles,instead we compare each gait cycle with every other gait cycle and the one which has maximum resemblance to other gait cycles is considered as the average gait cycle.

It has an advantage, every extracted gait cycle may not be a gait cycle in reality instead it can be noise. Hence those noisy gait cycle will least resemble to other gait cycle and could not be an average gait cycle.

Had we been taken the average of all gait cycle the noisy gait cycle would have been considered distorting average gait cycle.

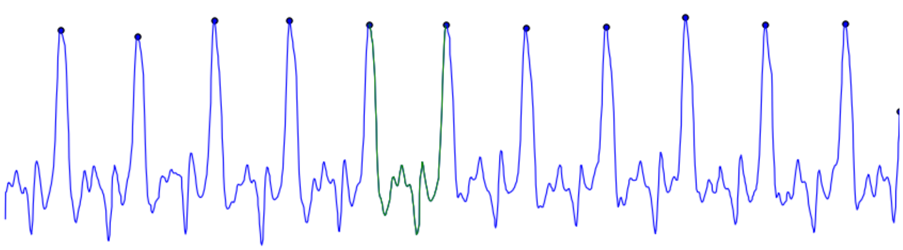

Figure 4.2 AGC of left leg(in black).

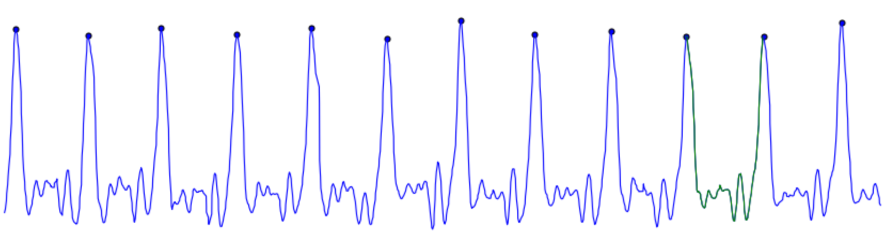

Figure 4.3 AGC of right leg(in black).

4.3 Dynamic Time Wrapping

The time domain technique that we used to compare any two gait cycle is Dynamic Time Wrapping(DTW)[6]. DTW is a pattern matching technique. We dont have to normalize the length/duration of each gait cycle taking advantage of the property of DTW.

```
dwt(array x,array y){
    cost[0][0] := d(x[0], y[0]);
    for i from 1 to nRows-1
        cost[i][0] := cost[i-1][0] + d(x[i], y[0]);
    for i from 1 to nCols-1
        cost[0][i] := cost[0][i-1] + d(x[0], y[i]);
    for i from 1 to nRows-1
        for j from 1 to nCols-1{
            mm := min( cost[i][j-1], cost[i-1][j-1] ) ;
            cost[i][j] := min( cost[i-1][j], mm ) + d(x[i], y[j]);
        }
    return cost[nRows-1][nCols-1];
}
WHERE d(x, y) = | x - y |
```

[6]

Chapter 5

Conclusion & Discussion

This project illustrates that applying moving average with small window size multiple times can be a good way to smooth a time series accelerometer data instead of applying moving average with large window size once, while preserving the characteristics of the gait cycle.

The similarity score between two average gait cycle template is reported as a whole number, a smaller value means greater resemblance.

Further Scope

- The gyroscope reading can be used to get the orientation of the smartphone and thus the rotation matrix. Hence the subject no longer has to keep the smartphone in fixed orientation in the trousers pocket.

- Post processing time can be reduced by at least half, by using multi-threading technique. As until extracting average gait cycle both legs data are processed independently.

- Although for low-pass filter we have used the value of alpha as 0.5 it can be calculated from the data set dynamically.

Bibliography

[1] https://en.wikipedia.org/wiki/Biometrics

[2] https://en.wikipedia.org/wiki/Low-pass_filter

[3] http://www.statisticshowto.com/moving-average

[4] http://machinelearningmastery.com/normalize-standardize-time-series-data-python

[5] Hoang Minh Thang, Vo Quang Viet1, Nguyen Dinh Thuc, Deokjai,year 2012 Choi http://web.engr.oregonstate.edu/ hoangmin/pub/c/2012_ICCAIS.pdf

[6] Muller, M., Information Retrieval for Music and Motion, Ch. 4 (available online at http://www.springer.com/cda/content/document/cda_downloaddocument/9783540740476-1.pdf?SGWID=0-0-45-452103-p173751818), Springer, 2007,ISBN 978-3-540-74047-6