Intelligent Systems

Project

Developing of a neural system, a Mamdany-type fuzzy inference system and a Sugeno-type fuzzy inference system for the identification of a volunteer position/activity

Francesco Paolo Culcasi

30/05/2016

This page intentionally left blank.

SUMMARY

1 - PROJECT SPECIFICATIONS

The project requires the analysis of a set of medical data. The application context is dermatology and, in particular, the compression therapy by means of bandages in the treatment of venous ulcers of the leg.

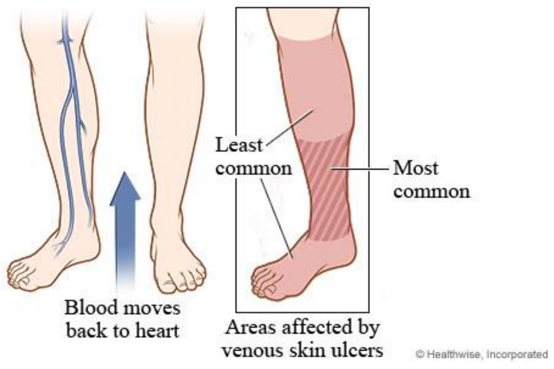
Least common
Most common
Blood moves back to heart
Areas affected by venous skin ulcers
© Healthwise, Incorporated

FIGURE 1 - COMPRESSION THERAPY

In dermatology, venous ulcers are very frequent lesion of the skin of the legs, due to a compromised functioning of the blood circulation. To repair these ulcers, the blood circulation must be enhanced by exploiting the calf muscle pump which allows the blood to be moved back to the heart. The basic way to re-establish the blood circulation is the compression therapy.

In the compression therapy, the doctor applies a compression bandage on the leg of the patient providing a given pressure of the ulcer area. The pressure applied by the bandage allows to repair the ulcer, typically, in a few months.

The pressure applied by the bandage, and thus the efficiency of the therapy, depends on several factors:

- The type of bandage (depending, e.g., on the elasticity);
- The correct application of the bandage: an appropriate pressure should be applied in several parts of the leg according to the position of the ulcer, and the pressure should remain constant as much as possible between bandage applications;
- The activities performed by the patient during the therapy (e.g., walking, standing, etc.).

1.1 - DATA DESCRIPTION

The data refer to 10 volunteers. A compression bandage was applied to their calf. Three sensors were applied in three different position to measure sub-bandage pressure. Each volunteer wears the bandage for 12 minutes. During this time, the volunteer perform different activities or maintains their positions. Each activity/position is performed/maintained for about 3 minutes. The sensors measure the pressure with the sampling time of about 82 ms. The positions/activities taken into account are the following:

A. Supine position
B. Dorsiflexion standing
C. Walking
D. Stair climbing

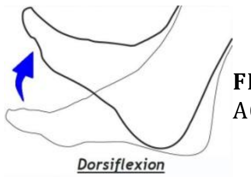

FIGURE 2 - DORSIFLEXION ACTIVITY

The data are organized as follows. One folder for each volunteer is provided. In each folder there are 4 files, one for each position/activity performed. Each file contains the measurements of the three sensors (in the first three columns), and the corresponding sampling time (in the last column). Please, note that the sensor measurements are electrical resistance and are measured in Ohms.

1.2 - OBJECTIVE

The objective of the project is twofold: on the one hand, we want to identify the position/activity of the volunteer by analyzing the pressure of the bandage, i.e., to distinguish among supine position, dorsiflexion standing, walking and stair climbing; on the other hand, we want to find out the least temporal interval that is necessary and sufficient to recognize the position/activity.

2 - DATA ANALYSIS

Given the data organization decrypted in the previous chapter, each signal coming from a sensor is a set of approximately 2000 samples. Rather than give all these data as input of our system, we prefer to appropriately represent the signal in terms of a reduced number of features that can summarize the information deriving from the original.

For this reason on the following paragraphs we are going to evaluate the signal after a preliminar feature extraction phase.

2.1 – DATA INITIALIZATION

At the starting point we have to deal with raw signals that are quite difficult to analyze. This implies a pre-processing phase with the purpose to get some structures in order to ease the computations on data (*Appendix A.1*). In the pressure measurement of the activity we observe that the three different sensors may show different shapes and shifts due to their position on the calf sub-bandage. For this reason we also introduced a fourth "virtual" sensor, obtained as the sum of the former three, to summarize the overall measurement.

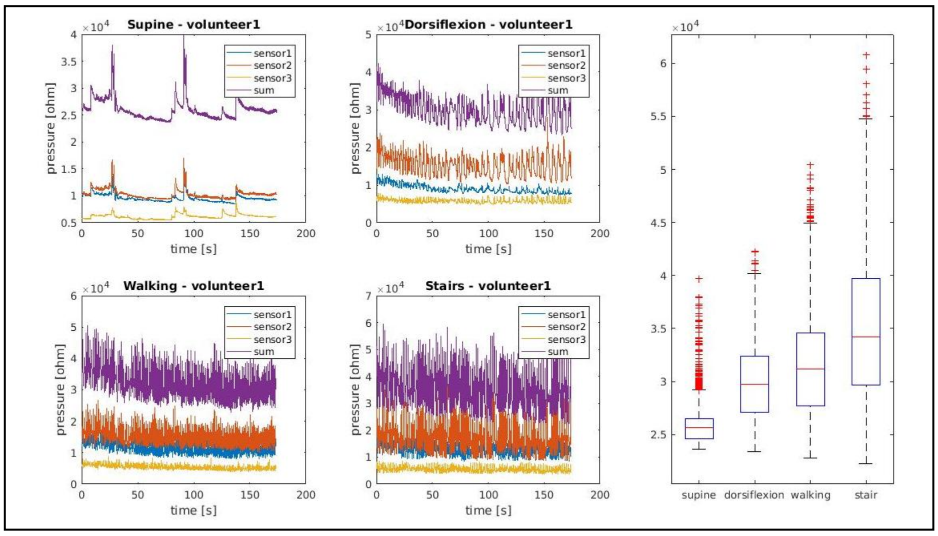

FIGURE 1 *SIGNALS AND BOXPLOTS OF POSITION/ACTIVITIES MEASURED ON VOLUNTEER 1*

Moreover our data should be sliced to obtain several pieces of the signal with smaller duration than the 3 minute signal for each activity. This has twofold profit: we split every signal to find out the least temporal interval that is necessary and sufficient to recognize the position/activity and even more we have much more signals that we can use to train/test/valuate our neural network (*Appendix A.2*).

We notice from the signals that it has been sampled with a sampling period of ~82ms (or with a 12.2Hz sampling frequency). Therefore we can compute the number N of samples needed to form a M-seconds-long piece of the signal with the following formula:

$$N = 12.2Hz \times M\,sec$$

A vector with several of these values has been defined to faster switch between different temporal intervals in our project. The values are computed for the following intervals: 3sec, 5sec, 10sec, 12sec, 15sec, 20sec, 30sec, 1min, 2min.

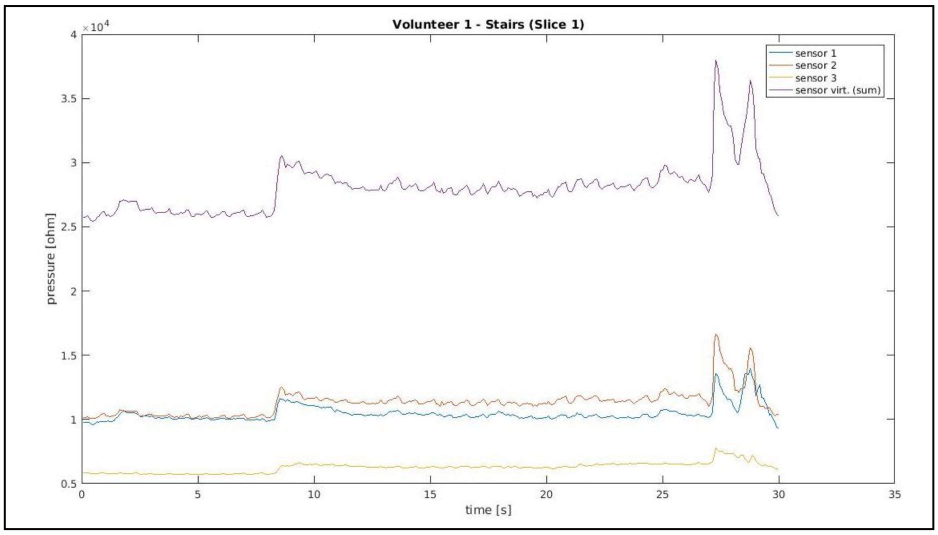

FIGURE 2 *FIRST 30-SECONDS-LONG SLICE OF THE SUPINE SIGNALS FROM VOLUNTEER 1*

2.2 – DATA CLEANING

The signals we are going to processing are obtained from pressure sensors. As like as every real world signal they are affected from noise that deteriorates the signal.

For this reason it is a good rule to apply some cleaning techniques before any feature extraction.

2.2.1 – SIGNAL SMOOTHING

One of the most common smoothing technique is the use of Savitzky-Golay filter. Savitzky–Golay filter is a digital filter that can be applied to a set of digital data points for the purpose of smoothing the data, that is, to increase the signal-to-noise ratio without greatly distorting the signal. This is achieved, in a process known as convolution, by fitting successive sub-sets of adjacent data points with a low-degree polynomial by the method of linear least squares.

For our project we used a third-order polynomial and a frame length of 7 (*Appendix A.3*).

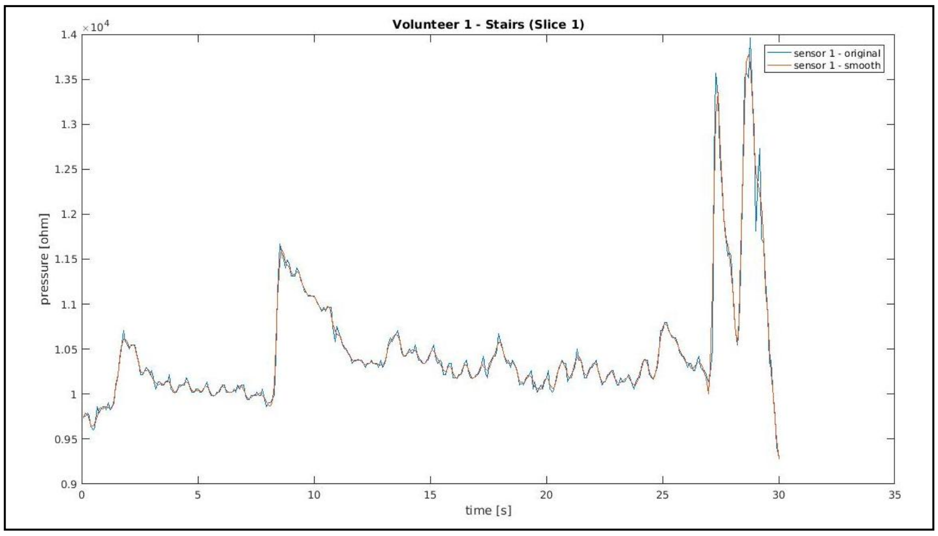

FIGURE 3 *SAVITZKY-GOLAY FILTER APPLIED TO A SIGNAL SLICE*

2.2.2 – NORMALIZATION

Another important point to deal with is the following: signals may have different scale and shift (see ***Figure 2Figure 2** First 30-seconds-long Slice of the supine signals from volunteer 1*). This is a huge problem since we would like to compare signals. To solve the problem we apply a normalization phase.

Since our dataset may contain some outliers we use Z-score normalization of the signal X, that is not affected from outliers. It can be described with the formula belove:

$$Z = \frac{X - \mu}{\sigma}$$

where we define with μ and σ respectively the signal mean and standard deviation.

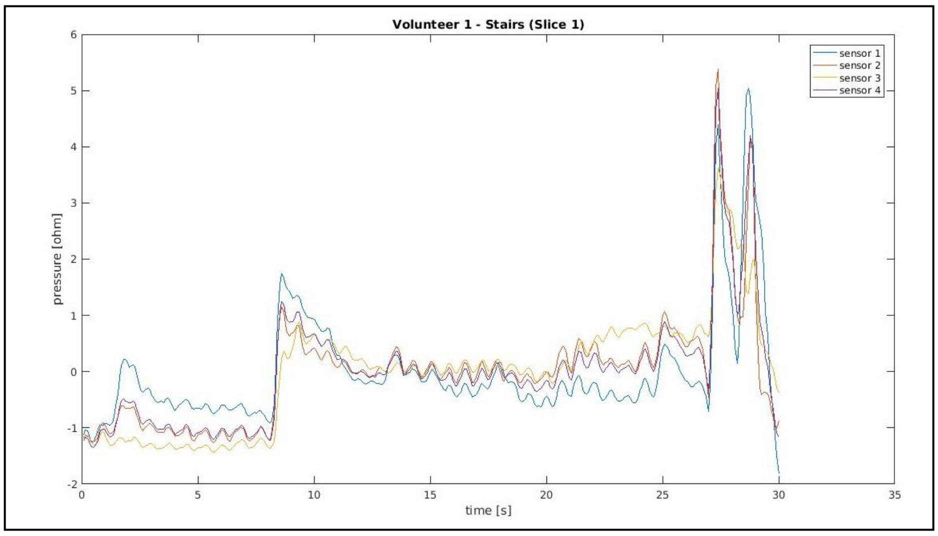

***FIGURE 4** NORMALIZED SIGNALS*

2.3 - FEATURE EXTRACTION

The goal is to reduce the number of inputs sufficient to distinguish among positions/activities. In order to do that we start analyzing signals shape.

From ***Figure 1***, representing the 4 activities of the volunteer 1, it become immediately evident that signals belonging on different activities can be discriminated by the following temporal characteristics:

- **Max/Min**: the maximum and minimum values of each signals may be significatively important;
- **Standardized moments**:
  - **1st std. moment (mean)**: after the normalization phase the mean value has been nullified;
  - **2nd std. moment (variance)**: the same as for mean value.
  - **3rd std. moment (skewness)**: is a measure of the asymmetry, computed as
    $$\widetilde{\mu_3} = \frac{E[(X-\mu)^3]}{\sqrt{(E[(X-\mu)^2])^3}};$$
  - **4th std. moment (kurtosis)**: is a measure of the "tailedness", measured as:

$$\widetilde{\mu_4} = \frac{E[(X-\mu)^4]}{\sqrt{(E[(X-\mu)^2])^4}};$$

- **Autocorrelation**: is the correlation of a signal with a delayed copy of itself as a function of delay, and can be obtained with
$$R_{xx}(\Delta) = \frac{E[(X_t-\mu)(X_{t+\Delta}-\mu)]}{\sigma^2}$$
For this project we take in consideration the first 20 lags (Δ=1,2,…,20).

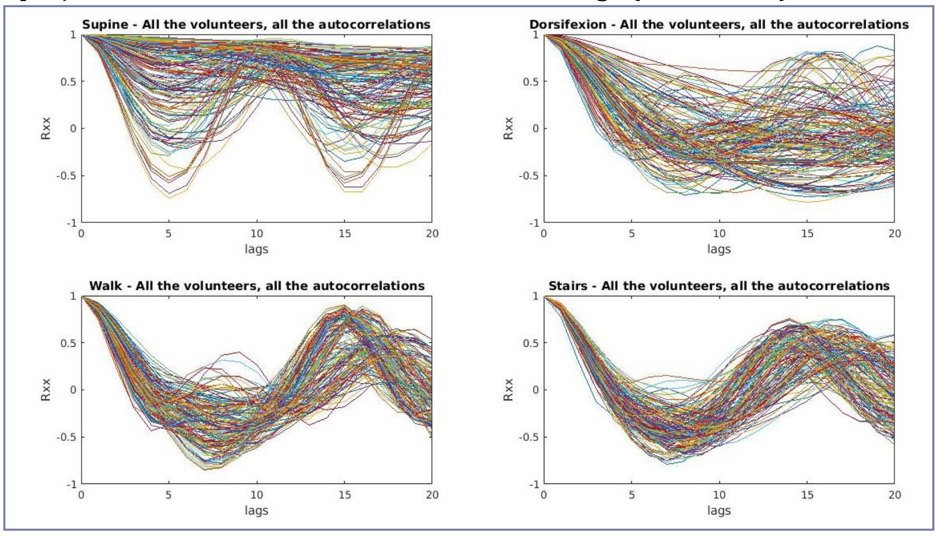

FIGURE 5 *AUTOCORRELATIONS FOR ALL THE ACTIVITIES OF EVERY VOLUNTEER*

In addition, performing some spectral analysis, we can infeer common frequential features for spectra for the same positions/actvities (see ***Figure 6***). We can introduce some frequency features that may look appropiated:

- **Fundamental frequency** (from now on f_0): achieved by detecting the maximum peak abscissa of the signals' spectrum, obtained by Fast Fourier Transform (FFT) formula,
$$X_n = \sum_{k=0}^{N-1} x_k e^{-i\frac{2\pi}{N}kq} \ (\mathrm{q} = 0,1,\dots,\mathrm{N}-1)$$
To faster compute f_0 we can approximate it with the maximum value of the FFT;
- **Amplitude of the** f_0 **peak** (amp);
- **Power Spectral Density** (PSD): obtained from the spectrum through the formula,
$$P_X = \sum_{n=-\infty}^{+\infty} |X_n|^2.$$

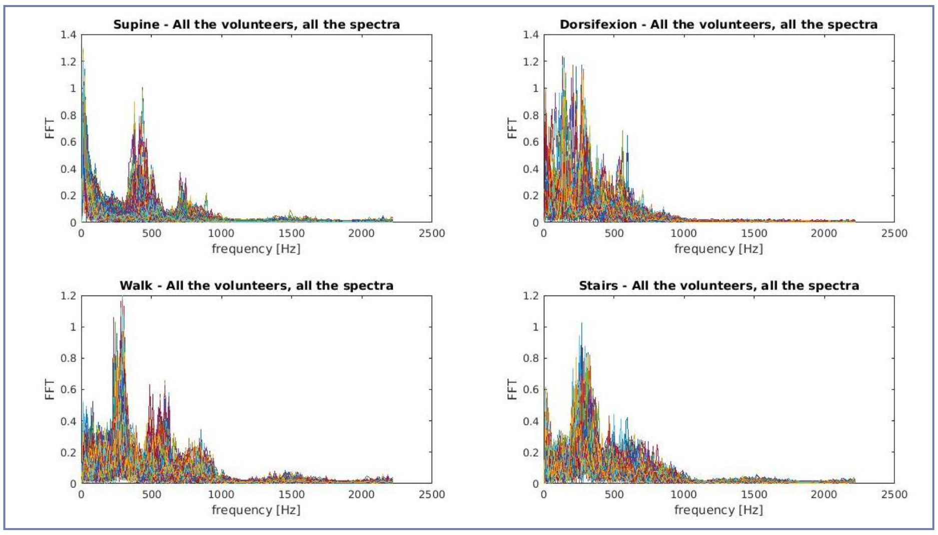

FIGURE 6 *FAST FOURIER TRANSFORMS FOR EVERY VOLUNTEER SEPARATED ON ACTIVITY/POSITION*

The feature extraction phase is showed in the *Appendix A.4*.

2.4 – FEATURE SELECTION

So far we have obtained 25 features: max, min, skewness, kurtosis, autocorrelation (20), f0, amp and PSD. But we have to compute these for each of the 4 sensors (3 + 1 virtual). Then we have 108 features.

They are still a lot of inputs for the system, thus we have to select only the more significant for the activity classification.

2.4.1 – SEQUENTIAL FEATURE SELECTION

To obtain the most important features to discriminate among positions/activities we can take advantage of the tools MATLAB makes available.

One of this is “`sequentialfs`”, that selects a subset of features from the input data matrix basing on the the result of the criterion function (*Appendix A.5*). Every detail can be found in the MATLAB documentation.

3 - NEURAL NETWORK

APPENDIX A - MATLAB SCRIPTS

A.1 INIT.M

```
%% 01 Data initialization - init.m
% - Data are grouped by activity/position and collected in a structure.
addpath('gitProjects/intelligent-system/');
load('ProjectWS.mat');

% This variable disable plots, so to fast computations.
showPlots = false;

% Neglect first row, it's just a row of zeros for every activity.
supine = {V01A(2:end,:), V02A(2:end,:), V03A(2:end,:), V04A(2:end,:),...
    V05A(2:end,:), V06A(2:end,:), V07A(2:end,:), V08A(2:end,:), ...
    V09A(2:end,:), V10A(2:end,:)};

dorsiflexion = {V01B(2:end,:), V02B(2:end,:), V03B(2:end,:), ...
    V04B(2:end,:), V05B(2:end,:), V06B(2:end,:), V07B(2:end,:), ...
    V08B(2:end,:), V09B(2:end,:), V10B(2:end,:)};

walking = {V01C(2:end,:), V02C(2:end,:), V03C(2:end,:), V04C(2:end,:),...
    V05C(2:end,:), V06C(2:end,:), V07C(2:end,:), V08C(2:end,:), ...
    V09C(2:end,:), V10C(2:end,:)};

stair = {V01D(2:end,:), V02D(2:end,:), V03D(2:end,:), V04D(2:end,:),...
    V05D(2:end,:), V06D(2:end,:), V07D(2:end,:), V08D(2:end,:), ...
    V09D(2:end,:), V10D(2:end,:)};

% Build the structure.
Struct = struct('supine',supine,'dorsiflexion',dorsiflexion,'walking',...
    walking,'stair',stair);

% TODO: is it necessary now?
% Pre-allocate this structure for the next step
%StructureDFT = struct('supine',{},'dorsiflexion',{},'walking',{},'stair',{});

for i=1:10
    % - Create a new signal as the sum of the three existing components and
    %   append at the other components.
    Struct(i).supine = ...
        [Struct(i).supine(:,4), Struct(i).supine(:,1:3), sum(Struct(i).supine(:,1:3),2)];
    Struct(i).dorsiflexion = ...
        [Struct(i).dorsiflexion(:,4), Struct(i).dorsiflexion(:,1:3),
sum(Struct(i).dorsiflexion(:,1:3),2)];
    Struct(i).walking = ...
        [Struct(i).walking(:,4), Struct(i).walking(:,1:3),
sum(Struct(i).walking(:,1:3),2)];
    Struct(i).stair = ...
        [Struct(i).stair(:,4), Struct(i).stair(:,1:3), sum(Struct(i).stair(:,1:3),2)];

    % - Plot of each volunteer's activity/position signal and boxplot
    %   to better show patterns, time features, probability distribution.
    if showPlots
        % full screen figure
        figure('units','normalized','outerposition',[0 0 1 1]);

        subplot(2,3,1);
        plot(Struct(i).supine(:,1),Struct(i).supine(:,2:5));
        title(strcat('Supine - volunteer ',num2str(i)));
        legend('sensor1','sensor2','sensor3','sum');
        xlabel('time [s]');
        ylabel('pressure [ohm]');

        subplot(2,3,2);
        plot(Struct(i).dorsiflexion(:,1),Struct(i).dorsiflexion(:,2:5));
        title(strcat('Dorsiflexion - volunteer ',num2str(i)));
        legend('sensor1','sensor2','sensor3','sum');
        xlabel('time [s]');
        ylabel('pressure [ohm]');

        subplot(2,3,4);
        plot(Struct(i).walking(:,1),Struct(i).walking(:,2:5));
        title(strcat('Walking - volunteer ',num2str(i)));
        legend('sensor1','sensor2','sensor3','sum');
        xlabel('time [s]');
        ylabel('pressure [ohm]');

        subplot(2,3,5);
        plot(Struct(i).stair(:,1),Struct(i).stair(:,2:5));
        title(strcat('Stairs - volunteer ',num2str(i)));
        legend('sensor1','sensor2','sensor3','sum');
        xlabel('time [s]');
        ylabel('pressure [ohm]');

        subplot(2,3,[3 6]);
        var = [Struct(i).supine(:,5)' Struct(i).dorsiflexion(:,5)'...
            Struct(i).walking(:,5)' Struct(i).stair(:,5)'];
        grp = [zeros(1,length(Struct(i).supine)), ...
            ones(1,length(Struct(i).dorsiflexion)), ...
            2.*ones(1,length(Struct(i).walking)), ...
            3.*ones(1,length(Struct(i).stair))];
        boxplot(var,grp,'Labels',{'supine','dorsiflexion','walking','stair'});
    end
end
```

A.2 SPLIT.M

```
%% 02 Data splitting - split.m
% - We need to split every signal to find out the least temporal interval
%   that is necessary and sufficient to recognize the position/activity.

% N.B. We are using a sampling period of ~82ms (12.2Hz).
% Number of samples based on the interval size:
% N1 = 12.2Hz * 3sec = 36.6;     %3sec
% N2 = 61;                       %5sec
% N3 = 122;                    %10sec
% N4 = 144.4;                  %12sec
% N5 = 183;                      %15sec
% N6 = 244;                           %20sec
% N7 = 366;                           %30sec
% N8 = 732;                           %1min
% N9 = 1464;                     %2min
% N = 2196;                      %3min - NO SIGNAL HAS A 3min LONG TRACE
N = [37 61 122 145 183 244 366 732 1464];

% The index only need to use one choice among the upper showed.
index = 7;
clear A B C D;
A = Struct(1).supine(1:N(index), 2:5);
B = Struct(1).dorsiflexion(1:N(index), 2:5);
C = Struct(1).walking(1:N(index), 2:5);
D = Struct(1).stair(1:N(index), 2:5);

% If the piece of signal cannot fill entirely the number of samples
% needed, it wont be used:
% "length(Struct(i).<activity>)/N(j)" is used just for its integer part so
% to truncate incomplete pieces of signal.
for i=1:10
    for k=2:length(Struct(i).supine)/N(index)
        A = [A Struct(i).supine(1+N(index)*(k-1):k*N(index), 2:5)];
    end
    for k=2:length(Struct(i).dorsiflexion)/N(index)
        B = [B Struct(i).dorsiflexion(1+N(index)*(k-1):k*N(index), 2:5)];
    end
    for k=2:length(Struct(i).walking)/N(index)
        C = [C Struct(i).walking(1+N(index)*(k-1):k*N(index), 2:5)];
    end
    for k=2:length(Struct(i).stair)/N(index)
        D = [D Struct(i).stair(1+N(index)*(k-1):k*N(index), 2:5)];
    end
end
```

A.3 CLEANING.M

```
%% 03 Data cleaning - cleaning.m
% Clean the signal from noise by mean of a smoothing filter
% (Savitzky-Golay Filter), sgolayfilt(X,K,F):
%  - K=3, third-order polynomial;
%  - F=7, just an odd value greater than the piece of signal;

clear smoothA;
smoothA = sgolayfilt(A,3,7);
% plot(1:N(j),[A(:,1:4) smoothA(:,1:4)])

clear smoothB;
smoothB = sgolayfilt(B,3,7);

clear smoothC;
smoothC = sgolayfilt(C,3,7);

clear smoothD;
smoothD = sgolayfilt(D,3,7);

% Use Z-score normalization for each signal, in order to be able to compare
% signals.
% figure,
% plot(1:N(index),zscore(smoothA(:,1:4)));
% figure,
% plot(1:N(index),smoothA(:,1:4));
smoothA = zscore(smoothA);
smoothB = zscore(smoothB);
smoothC = zscore(smoothC);
smoothD = zscore(smoothD);

if showPlots
    figure, plot(1:size(smoothA),smoothA(:,[1:4 21:24 101:104]))
    figure, plot(1:size(smoothB),smoothB(:,[1:4 21:24 101:104]))
    figure, plot(1:size(smoothC),smoothC(:,[1:4 21:24 101:104]))
    figure, plot(1:size(smoothD),smoothD(:,[1:4 21:24 101:104]))
end
```

A.4 FEATUREEX.M

```
%% 04 Feature extraction - featureEx.m
% Obtain a set of temporal feature starting from the N(j) samples of the
% signal.
clear fftA fftB fftC fftD;

clear featuresA;
featuresA = [max(smoothA); min(smoothA); skewness(smoothA); ...
    kurtosis(smoothA)];
clear featuresB;
featuresB = [max(smoothB); min(smoothB); skewness(smoothB); ...
    kurtosis(smoothB)];
clear featuresC;
featuresC = [max(smoothC); min(smoothC); skewness(smoothC); ...
    kurtosis(smoothC)];
clear featuresD;
featuresD = [max(smoothD); min(smoothD); skewness(smoothD); ...
    kurtosis(smoothD)];

%% TODO: Autocorrelation
% computed at lags 0,1,2, ... T= min[20,length(y)-1]
clear RxxA;
RxxA = my_autocorr(smoothA); % See my_autocorr.m
clear RxxB;
RxxB = my_autocorr(smoothB);
clear RxxC;
RxxC = my_autocorr(smoothC);
clear RxxD;
RxxD = my_autocorr(smoothD);

if showPlots
    figure, plot(1:size(RxxA),RxxA);
    figure, plot(1:size(RxxA),RxxB);
    figure, plot(1:size(RxxA),RxxC);
    figure, plot(1:size(RxxA),RxxD);
end

% first value, autocorr with lags=0, is always equal to 1 (it's irrelevant)
featuresA = [featuresA; RxxA(2:end,:)];
featuresB = [featuresB; RxxB(2:end,:)];
featuresC = [featuresC; RxxC(2:end,:)];
featuresD = [featuresD; RxxD(2:end,:)];

%% Frequential features:
% - Fundamental frequency f0
% - Power Spectral Density PSD

% Define the frequency domain
f = 12.2*(0:N(index)/2-1);

% Compute the single-sided spectrum
fftA = my_fft(smoothA,N(index)); % See my_fft.m
fftB = my_fft(smoothB,N(index));
fftC = my_fft(smoothC,N(index));
fftD = my_fft(smoothD,N(index));

if(showPlots)
    figure, plot(f,fftA);
    figure, plot(f,fftB);
    figure, plot(f,fftC);
    figure, plot(f,fftD);
end

% The max amplitude should be a good approximation for the fundamental
% frequency
[amp, x] = max(fftA);
% [~, x2] = findpeaks(fftA, 'MinPeakProminence', 0.7*max(fftA));

% PSD
PSD = sum(fftA.^2);
% figure
% plot(f,fftA(:,[1 11 111]));
% hold on;
% plot(f(x([1 11 111])),y([1 11 111]),'rv');
featuresA = [featuresA; f(x); amp; PSD];

clear amp x;
[amp, x] = max(fftB);
featuresB = [featuresB; f(x); amp; sum(fftB.^2)];
clear amp x;
[amp, x] = max(fftC);
featuresC = [featuresC; f(x); amp; sum(fftC.^2)];
clear amp x;
[amp, x] = max(fftD);
featuresD = [featuresD; f(x); amp; sum(fftD.^2)];

%% Rotate matrix
% Since we use 4 columns to represent 3 sensor signals + 1 "virtual" sensor
% (the sum), we move all the features of the same piece of signal on the same
% column.
newFeaturesA = rotate_features(featuresA);
newFeaturesB = rotate_features(featuresB);
newFeaturesC = rotate_features(featuresC);
newFeaturesD = rotate_features(featuresD);
```

A.5 FEATURESEL.M

```
%% 05 Feature Selection
% Sequential feature selection

% Classify each activity with a different class:
% Y = 0 => supine
% Y = 1 => dorsiflexion
% Y = 2 => walking
% Y = 3 => stairs
X = [newFeaturesA                    newFeaturesB ...
     newFeaturesC                    newFeaturesD]';
Y = [zeros(size(newFeaturesA,2),1);  ones(size(newFeaturesB,2),1); ...
     2*ones(size(newFeaturesC,2),1); 3*ones(size(newFeaturesD,2),1)];
%
% % Randomly order activities features
% perm = randperm(size(Y));
% X = X(perm,:);
% Y = Y(perm);
%
% % Divide dataset in training/test datasets
% w = 0.7*size(Y);
% XT = X(1:w, :); Xt = X(w+1:end, :);
% YT = Y(1:w); Yt = Y(w+1:end);

f = @(xtrain, ytrain, xtest, ytest) sum(ytest ~= classify(xtest, xtrain, ytrain));
opts = statset('display','iter');

[fs, history] = sequentialfs(f,X,Y,'nfeatures',10,'options',opts);
```