**People's Democratic Republic of Algeria**
**Ministry of Higher Education and Scientific Research**

**University M'Hamed BOUGARA – Boumerdes**


**Institute of Electrical and Electronic Engineering**

**Department of Electronics**

Final Year Project Report Presented in Partial Fulfilment of the Requirementsfor the Degree of

# MASTER

In **Electrical and Electronic Engineering**

Option: **Telecommunications**

Title:

# Hidden Biometrics for identification using ECG and EEG Signals

Presented by:

- **Chaouki Adjerid**
- **Billal Boukerma**

Supervisor:

**Dr.Dalila CHERIFI**

Registration Number:...../2016

# ABSTRACT

Security concerns increase as the technology for falsification advances and biometrics provides airtight security by identifying an individual based on the physiological and/or behavioral characteristics. Physiological hidden biometrics represented by ECG and EEG biomedical signals are highly confidential, sensitive, and hard to steal and replicate, and also hold great promise to provide a more secure biometric approach for user identification and authentication.

This work proposes the human heartbeat as a characteristic to be used for identity recognition. An ECG-based biometric identification system is developed, a method based on autocorrelation (AC) in conjunction with the discrete cosine transform (DCT) proposed for feature extractions from the pre-processed ECG signal.

Also studied is the scenario where the proposed system deals with intruder signals in our database. For this goal, a study is performed to adjust the parameters allowing the system to avoid detection failure of false identification and false rejection scenarios.

In addition, human brain activities represented by EEG are studied for biometric system purposes. In this study, an EEG-based biometric system is represented by performing a pre-processing stage on the EEG signals, with the features extractions completed using a wavelet packet decomposition and a classification.

Further, the effect of myocardial infractions on our ECG-based identification system is also studied, by taking into account a disease, namely Atrial fibrillation, as an example for the study. Finding ECG recordings for people with Atrial fibrillation is possible, however, finding a sample of their ECG signals prior to sickness is not possible, unfortunately, so we therefore revert to developing a program able to simulate an ECG signal of a healthy person, and then make it resemble the ECG signal of a person suffering of Atrial fibrillation. By doing so, we will then test our identification system to determine whether the people tested can still be recognised.

## *ACKNOWLEDGEMENT*

The authors would like to take this opportunity to thank the following persons for the assistance extended to them:

To Dr.Cherifi, supervisor; for her valuable comments , suggestions and sharing of knowledge to make this report more meaningful as well as for her psychological support by encouraging us and spending time in providing instructions during the last months .

To Professor RachidNakib, teacher at Colorado University , for reviewing this work and providing advice concerning the format and the language.

To Doctors Karim Boukerma and Chemseddine Ouitis for sharing their knowledge with us.

To our supervisor in our training at CDTA, Dr.Bengherabi, who was the first to introduce us to biometry .

To our beloved parents for the moral and financial support in making this workbook possible.

And most of all, to our Almighty Allah, for giving his endless blessings, knowledge and strength in order to make this modular workbook possible.

Once again, we thank all those who have encouraged and helped us in preparing this workbook, and who have extended to us so much understanding , patience , and support.

# *THESIS ORGANIZATION*

# LIST OF FIGURES

# *LIST OF TABLES*

# *INTRODUCTION*

The first chapter of this project introduces the theme of our study by providing some generalities in addition to the objectives and motivations to pursue the proposed work. A summary of the proposed work organization is also included.

Automatic and accurate identity verification and validation is becoming increasingly critical in many aspects of today's world, namely in financial transactions, travelling, and access control to airports and other protected zones. Traditional strategies to automatic identity recognition include items such as Pin numbers, tokens, passwords and ID cards. According to the latest US Federal Commission Report for 2009 and 2010, identity theft was the number one complaint category in the United States with a total of 721,418 cases of consumer complaints. Unlike traditional forms of authentication (e.g., passwords and tokens), biometric recognition provides a highly reliable link between the human individual and his associated identity. The use of biometrics as means of security will therefore offer a more robust tool of identification which will help prevent the problems and deficiencies associated with more conventional identification and verification techniques.

The term "biometrics" is derived from the Greek words "bio" meaning life, while the term "metric" means to measure. Biometrics is then defined as the measurement and statistical analysis of people's physical and behavioral characteristics. This relatively recent technology is used for identification and access control, which include the surveillance of suspected individuals by law enforcement as well as for the protection of highly sensitive securitized areas [1]. A biometric system generally consists of three main processes, independently of the human characteristics involved and the application being used. Basically, a biometric system is composed of three phases as shown in figure 1-1: (1) a pre-processing stage, (2) a features extraction stage and (3) a classification phase.

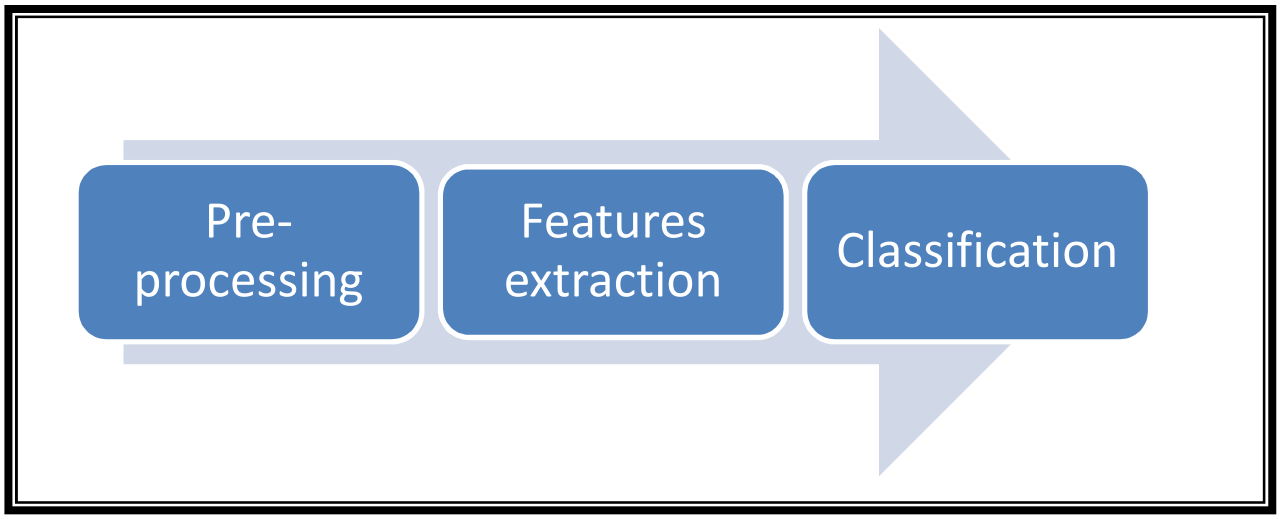



**Figure 1-1:** *General block diagram of the biometric system.*

The fundamental premise of biometric authentication is based on the biological/genetic uniqueness of each human individual who can therefore be identifiable through intrinsic traits and characteristics. Biometrics authentic identification of human beings is then related to characteristics which present various specific properties such as uniqueness and persistency, which make these characteristics suitable for the differentiation/identification of human beings. Traditionally; as introduced in figure 1-2, these characteristics have been categorized into two major groups: physical and behavioural.

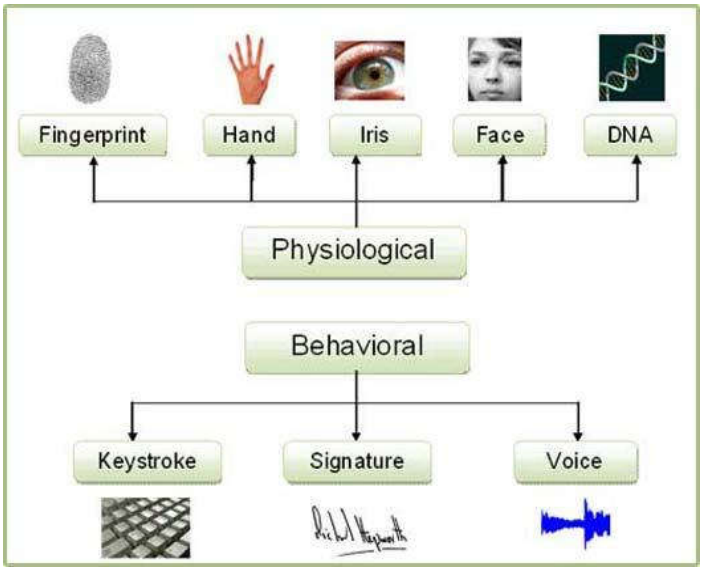


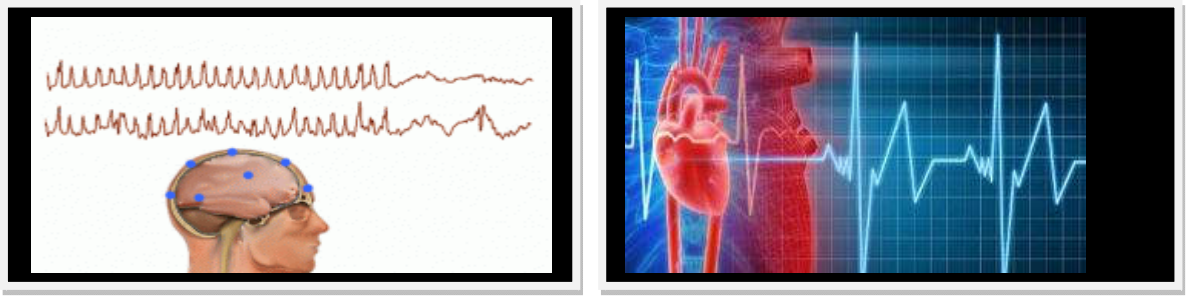

**Figure 1-2:** *Some physical and behavioural characteristics [2].*

- Behavioural characteristics relate to individual behaviour, and are associated to behavioural patterns, among them: signature recognition, voice/speaker verification/authentication, as well as typing recognition.

- Physical characteristics, on the other hand, are concerned with the shape or composition of the human body, and are subdivided into two types: visual and hidden characteristics:

  - Visual characteristics are those seen and visualized by the human eye. Some examples are: fingerprints recognition, ear shape recognition, eyes-iris recognition, eyes-retina recognition, and face recognition.
  - Hidden characteristics are invisible to the eye, requiring the use of highly sophisticated machinery and sensitive sensors for their detection. Examples are: DNA matching, ECG, and EEG [1].

Since each biometric characteristic has its own advantages and disadvantages, the question of which biometric characteristics should be used for authentication application will depend solely on the application requirements. Traditional biometrics has been crucial in providing a security mechanism used to authenticate and provide access to secured facilities, however, they can also be prone to forgery, unfortunately. In order to avoid this security breach issue, ECG and EEG biomedical signals have been selected in this research (figure 1-3). Firstly, there are an axe of research and new biometric models in recent research proposals that have suggested the possibility of using biomedical data as a biometrics modality for human identity recognition. Secondly, since ECG and EEG are biodynamic signals by nature, they offer unique advantages over traditional biometric techniques by ensuring the aliveness of the individual. This feature offers a superior and robust advantage over traditional methods which utilize static physical characteristics, which are highly prone to the risk of signal duplication. For example, a human face can easily be replaced by a photograph in front of a security system camera, and fresh fingerprints could easily be collected from a surface that has just been touched, and then forged with artificial fingerprints. As for spoofing, it is relatively difficult to achieve with ECG and EEG signals, yet not impossible. This work focuses on ECG and EEG technologies, particularly for the reason that these two emerging technologies possess the highly desirable capability to act as biological signatures suitable for dynamic continuous monitoring. Given the fact that every human individual has a unique brain configurations, spontaneous EEG signals would differ from one person to another, and the same applies to the heart, where shape differs from person to person, and so does the individual heart electric activity of each person. Nonetheless, in order to capture the complexities and size of such differences, a defined number of sensors must be placed at different areas, namely, on the human scalp to detect EEG signals, and at different areas of the human body in order to detect ECG signals.



(a) (b)

**Figure 1-3:** *Hidden biometrics; (a) EEG [3]; (b) ECG [4].*

The proposed research raises three fundamental questions:

*(1) How accurate can a verification of identity ECG-based system be?*

*(2) How accurate can an EEG-based system be?*

*(3) How can the limitations of performance of the two proposed systems be managed?*

In order to answer the stated questions and to develop the proposed biometric systems a research was conducted. This research is organized into four chapters as follows:

In Chapter 2, a general introduction to heart and electrocardiogram science is formulated, by defining the necessary elements for an adequate understanding of the cardiovascular system. The proposed ECG biometric system is then highlighted by providing a detailed description of the methodology used to build a simulation algorithm as well as the various mathematical and technical tools needed in this study. The implementation of the various processes of the proposed system is also included within this chapter.

Chapter 3 is organized in the same manner as Chapter 2. General definitions of the brain activities and the electroencephalogram function are described. Then, detailed descriptions of the methodology and implementation of the proposed EEG biometric system are introduced with a complete description of the necessary definitions of the mathematical and technical tools being used.

Chapter 4 explores the experimental investigations of this study, with statistical analyses performed for various proposed biometric system scenarios, as well as their respective performances. The evaluations and conclusive results obtained are discussed and compared with related existing works in order to compare and bring to light the benefits and advantages, but also the possible limitations of the biometric systems proposed in this study.

# ECG-BASED BIOMETRICS

This chapter presents an introduction with overall descriptions of the anatomy of the human heart as well as some general definitions of ECG-based technology. The methodology and implementation of the proposed ECG system is also described.

## 2.1 Introduction:

In order to determine the potential use of ECG as a biometric authentication and verification tool, it would be necessary to evaluate how ECG satisfies the requirements for biometric characteristics. Thus, a "perfect" biometric characteristic should be:

- Universal, i.e., each individual possesses this characteristic,
- Easily measured, i.e., quite easy technically and convenient for an individual to obtain the characteristic,
- Unique, i.e., there are no two individuals with identical characteristics, and
- Permanent, i.e., the characteristic does not change over time.

"Good" biometric characteristics can, to a greater or lesser extent, satisfy these requirements, depending on the purpose and application of the biometric system.

The ECG is a universal characteristic, since the heart beat is a necessary sign of life, and can be recorded with minimum inconvenience to the human individual. But in order to evaluate its uniqueness, and particularly, its persistence, the task becomes much more difficult. It is plausible to assume that an ECG is an almost unique human characteristic because morphology and amplitudes of recorded cardiac complexes are governed by multiple individual factors, in particular by the shape and position of the heart, and the presence and nature of pathologies, among other factors. As a result, QRS complexes possess a significant variety of configurations and metrics. The analysis of electrocardiogram (ECG) as a tool for clinical diagnostics has been an active research area in the past two decades or so, but only recently have researchers started looking at the potentiality of the bio-signal (ECG) as a biometric model for identification and verification purposes. And their results showing strong indications that the human heartbeat can indeed be used for identity recognition.

The existing solutions for biometric recognition from electrocardiogram signals are based on temporal and amplitude distances between detected fiducial points, with such methods relying heavily on the accuracy of the fiducial point detection. This type of solution still encounters a problem due to the difficulty in localizing the wave's boundaries with desirable accuracy. In this chapter, a new approach based on autocorrelation (AC) in conjunction with discrete cosine transform (DCT) is proposed, and in order to enhance the description of the proposed method, general descriptions of the heart and Electrocardiogram system are provided.

## 2.2 Overview of Heart and Electrocardiogram:

### 2.2.1 The Cardiovascular system:

The cardiovascular system is made up of: the heart; lungs; arteries and veins, which are all under the control of the autonomic nervous system (sympathetic and parasympathetic). In a healthy individual with a healthy heart, heart rate is dictated by the body's needs. If an individual is resting, then organ muscles and tissues require a reduced amount of blood and oxygen. This results in a reduction in blood pressure, a heart rate slowdown, and lower breathing. When the individual is more active, then the organs, muscles and tissues will all need an increase in blood and oxygen, resulting in a rise in blood pressure and an increase in heart rate and breathing. These responses are involuntary, and under the direct control of the autonomic nervous system. Note that the above scenarios apply only to reasonably healthy individuals with no cardiac complications.

### 2.2.2 Normal Function of the Heart:

The heart can simply be described as a pump, since the human heart plays a role quite similar to the role played by a central heating system. The heart pumps blood through the arteries and veins to the organs, muscles and tissues, in a manner similar to the central heating pump, forcing hot water to radiators through pipes.

### 2.2.3 The Heart:

The heart is made up of four parts or chambers, two atria and two ventricles. De-oxygenated blood returns to the right side of the heart via venous circulation. Blood is then pumped into the right ventricle and then to the lungs where carbon dioxide is released and oxygen is absorbed. The oxygenated blood will travel back to the left side of the heart into the left atria, then into the left ventricle from where it is pumped into the aorta and arterial circulation.

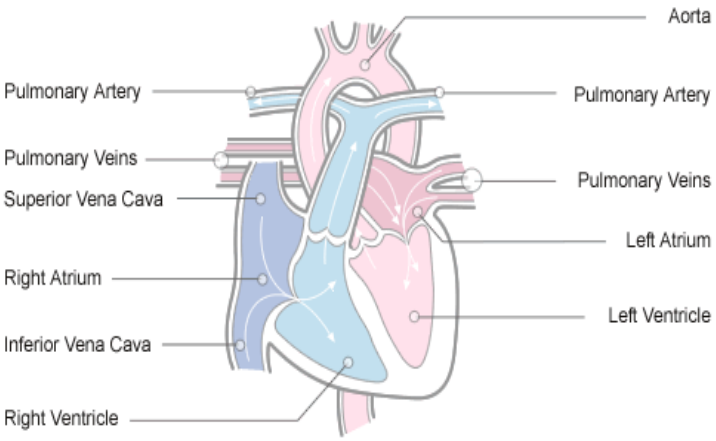



**Figure 2-1***: The passage of blood through the heart.[5]*

The pressure created in the arteries by the contraction of the left ventricle is known as the systolic blood pressure. Once the left ventricle has fully contracted, it begins to relax and refill with blood from the left atria. The pressure in the arteries falls while the ventricle refills. This function is known as the diastolic blood pressure. The two sides of the heart never communicate directly. Blood travels from the right side to the left side via the lungs only. However, the chambers themselves work together. The two atria contract simultaneously, and the two ventricles contract simultaneously also. The passage of blood through the heart is illustrated by figure 2-1. The description above is related to the heart anatomy, but the causes that lead the chambers to contract are what remain of most importance to our study.

### 2.2.4 Cardiac Conduction System:

Going back to the analogy of the central heating system, the pump, pipes and radiators are of no use unless connected to a power supply. The pump needs electricity in order to work. The human heart has a similar need for a power source, and also uses electricity. Thankfully, humans won't need to plug themselves into an electrical main, since the heart creates its own electrical impulses and controls the routes taken by the impulses via a specialized conduction pathway.

The pathway in question is made up of five elements:

- The sino-atrial (SA) node.
- The atrio-ventricular (AV) node.
- The bundle of His.
- The left and right bundle branches.
- The Purkinje fibres.

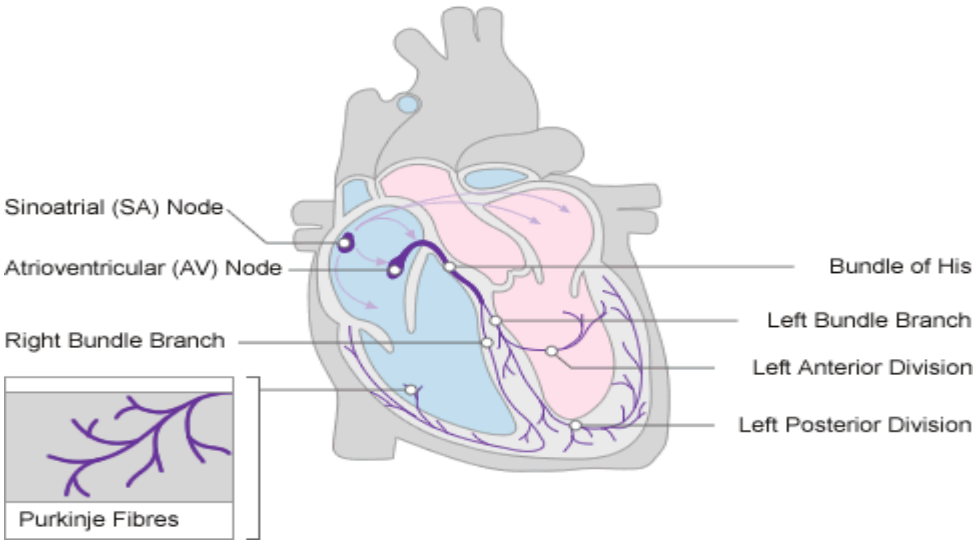



**Figure2-2:** *The cardiac conduction system.[5]*

The SA node is the natural pacemaker of the heart. It releases electrical stimuli at a regular rate, with the rate being dictated by body needs. Each stimulus passes through the myocardial cells of the atria, creating a wave of contractions which spread rapidly through both atria. The heart is made up of approximately half a billion cells, as shown in Figure2-2 where the difference in muscle mass of the various chambers is shown. The majority of the cells make up the ventricular walls. The rapidity of Atrial contraction is such that approximately 100 million myocardial cells contract in less than one third of a second, appearing almost instantaneous. The electrical stimulus from the SA node eventually reaches the AV node and is delayed briefly so that the contracting atria have enough time to pump all the blood into the ventricles. Once the atria are empty of blood the valves between the atria and ventricles close. At this point, the atria begin to refill and the electrical stimulus passes through the AV node and Bundle of His into the Bundle branches and Purkinje fibers. In this way, all the cells in the ventricles receive an electrical stimulus causing them to contract. Around 400 million myocardial cells making up the ventricles contract in less than one third of a second. As the ventricles contract, the right ventricle pumps blood to the lungs where carbon dioxide is released and oxygen is absorbed, whilst the left ventricle pumps blood into the aorta from where it passes into the coronary and arterial circulation. At this point, the ventricles are empty, the atria are full and the valves between them are closed. The SA node is now about to release another electrical stimulus and the process is about to repeat itself. However, there is a 3rd section to this process. The SA node and AV node contain only one stimulus. Therefore every time the nodes release a stimulus, they must recharge before they can do it again.The SA node recharges whilst the atria are refilling, and the AV node

recharges when the ventricles are refilling. In this way, there is no need for a pause in heart function. Again, this whole process takes less than one third of a second. The times given for the three different stages are based on a heart rate of 60 beats per minute, or one beat per second. The term used for the release (discharge) of an electrical stimulus is "depolarisation", while the term used for recharging is known as "repolarisation".

So, the three stages of a single heart beat are:

- Atrial depolarisation.
- Ventricular depolarisation.
- Atrial and ventricular repolarisation.

As the atria repolarise during ventricular contraction, there is no wave representing Atrial repolarisation as it is buried in the QRS.

The different heart elements responsible of the cardiac conduction system are shown in figure 2-2 .

### 2.2.5 Sinus Rhythm:

Sinus rhythm is the name given to the normal rhythm of the heart where electrical stimuli are initiated in the SA node, and are then conducted through the AV node and bundle of His, bundle branches and Purkinje fibres. Depolarisation and repolarisation of the atria and ventricles show up as 3 distinct waves on ECG. A unique labelling system is used to identify each wave. Although the diagram shows 5 waves, we will be concentrating on 3 waves. One does not always see a Q wave or an S wave on an ECG. This is why only 3 waves are emphasized.

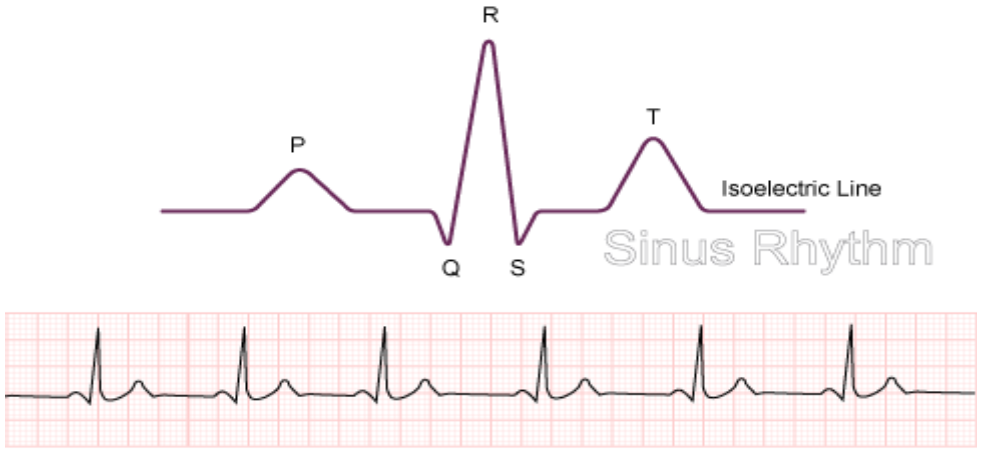



**Figure2-3:***P,QRS& T waves.[5]*

The more cells there are, the more voltage is required. If one takes a look at Figure 2-1 and Figure2-2, one sees that the walls of the atria are much thinner and smaller than those of the ventricles.

Now we will go through our main purpose of studying the heart's anatomy, the Electrocardiogram wave.

### 2.2.6 ECG basics:

An Electrocardiogram signal describes the electrical activity of the heart, itself related to the impulses travelling through the heart while providing information about the heart rate, rhythm, and morphology. Normally, ECG is recorded by attaching a set of electrodes on the body surface such as a chest, neck, arms, and legs.

A typical ECG wave of a normal heartbeat consists of a P wave, a QRS complex, and a T wave. Figure2-3 depicts the basic shape of a healthy ECG heartbeat signal. The P wave reflects the sequential depolarization of the right and left atria. It usually has positive polarity and its duration is less than 120 milliseconds. The spectral characteristic of a normal P wave is usually considered to have low frequency, below 10-15 Hz. The QRS complex corresponds to depolarization of the right and left ventricles. It lasts for about 70-110 milliseconds in a normal heartbeat, and has the largest amplitude of the ECG waveforms. Due to its step slopes, the frequency content of the QRS complex is considerably higher than that of the other ECG waves, and is mostly concentrated in the interval of 10 to 40 Hz. The T wave reflects ventricular repolarization and extends about 300 milliseconds after the QRS complex. The positive of the T wave is strongly dependent on the heart rate, becoming narrower and closer to the QRS complex at rapid rates.

### 2.2.7 Variation of ECG waveform within time:

The persistence of an individual's ECG characteristics over time is now considered. Since the shape of ECG waveforms is determined primarily by human anatomical features, it can be assumed that the ECG signal will have slow and gradual variations within time,at least for individuals with no traumatic history, such as myocardial infractions. In order to evaluate the tendency of ECG variability over time, a set of ECG records of one individual candidate were collected, first, within one hour, and then within six months. For each record, the acquisition procedure including attached electrodes was repeated from the beginning. Figure 2-4 shows the variability of ECG cycles recorded within one hour (Figure 2-4 (a)) and within six months (Figure 2-4 (b)). The figures show that the variations associated to one hour are almost identical to the variations of the six months period. Obviously, these results represent only empirical observations, and a six month period is considered too short a period to judge the variability of the ECG over much longer periods; however, these results do support the hypothesis of a slow variation. [6]

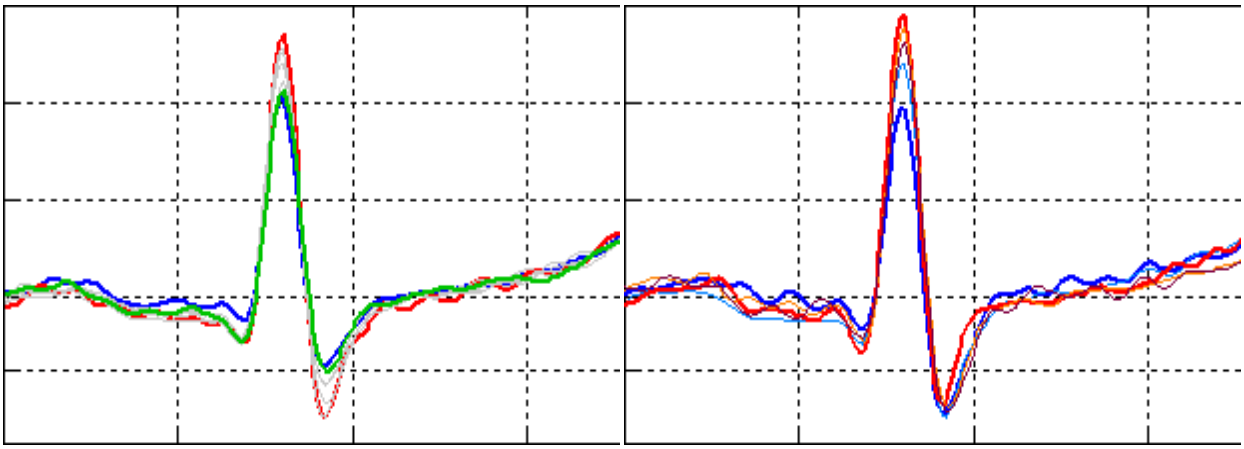



**Figure 2-4:** *Variation of ECG of one individual recorded at different times; (a) within one hour; (b) within six month.[6]*

## 2.3 ECG-Based Biometrics Methodology and Implementation:

The development of a biometric system based on ECG signal for user identification and authentication is presented in this chapter. Throughout this work, the Pre-processing related to De-noising of the ECG signal is first presented. Then, the extraction phase of the desired features from the signal using a new method known as the AC/DCT method is introduced. This method, unlike previous techniques, is not based on fiducial points. At the

end phase, an identification system is suggested based on a method that will be explained below.

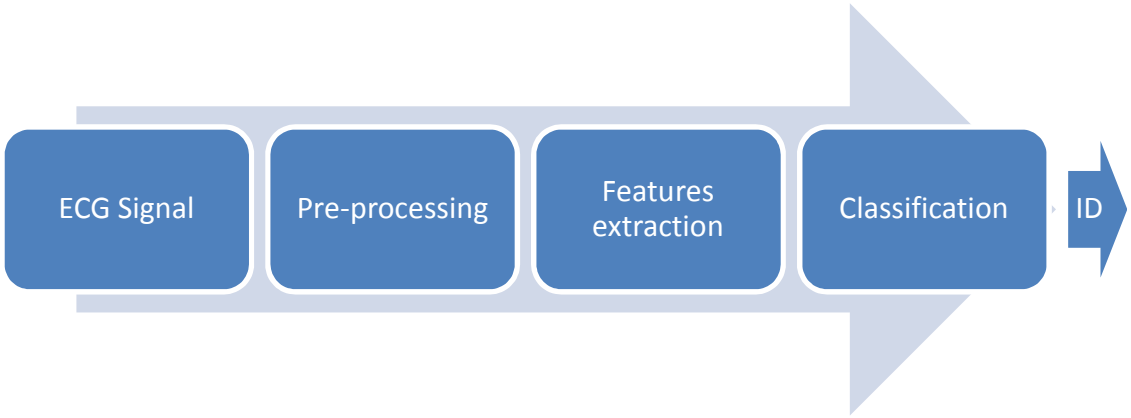


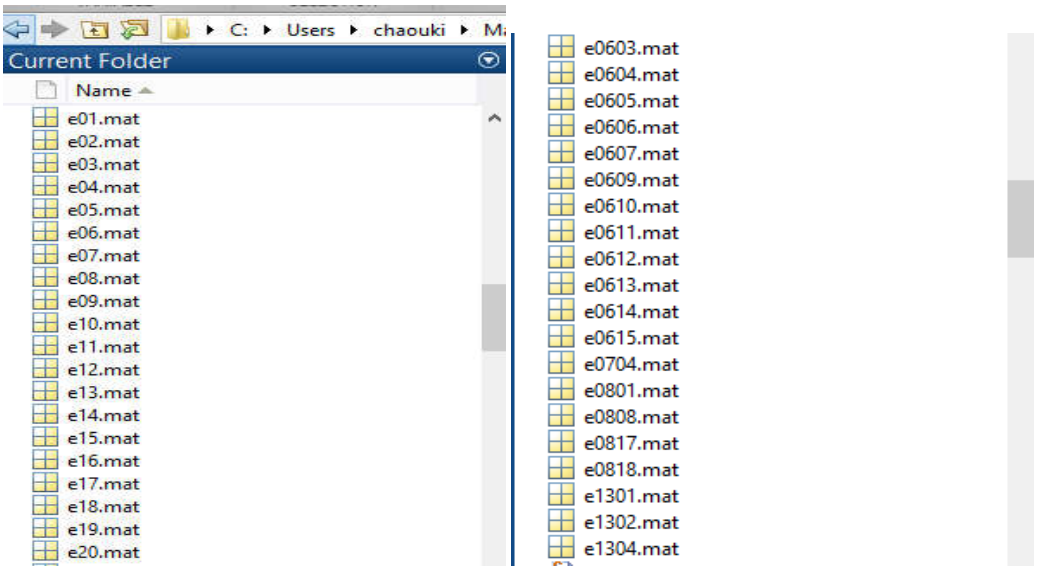

**Figure2-5**:*Block diagram of the proposed ECGbiometric system.*

**Note:** All algorithms were implemented using MATLAB 2012b.

### 2.3.1 ECG Database:

All the recordings were extracted from European ST-T Database and MIT-BIH Arrhythmia Database which are available on 'Physionet.com' website in public domain. In this chapter, our database consists of 20 ECG recordings of healthy persons taken from the European ST-T database. Each recording is of 1 minute duration, first 40 seconds for training, and 20 seconds for testing, and for better perception, each recording was attributed to a person of 20 students selected from IGEE institute. Names and their respective pictures were then added. Each recording was sampled at 256 samples (bits) per second. 20 more European ST-T Database ECG recordings were downloaded for use as intruders in the second part of the study.

MIT-BIH Arrhythmia Database will be used only in Chapter 4 for performance comparison.




(a) (b)

**Figure 2-6:** *ECG database ((a) 20 ECGs of local database; (b) 20 ECGs of intruders).*

**a) European ST-T Database:**

The European ST-T Database is shared to be used for evaluation of algorithms for analysis of ECG waves. This database consists of 90 annotated excerpts of ambulatory ECG recordings from 79 subjects, the subjects being 70 men aged between 30 and 84, and 8 women aged between 55 and 71.Myocardial ischemia was diagnosed or suspected for each subject; additional selection criteria were established in order to obtain a representative selection of ECG abnormalities in the database. Each record is of two hours in duration and contains two signals, each sampled at 250 samples per second with 12-bit resolution over a nominal 20 millivolt input range. The sample values were rescaled after digitization with reference to calibration signals in the original analogous recordings, in order to obtain a uniform scale of 200 ADC units per millivolt for all signals, (The calibration signals are not included in the signal files). Over half (48 of 90 complete records, and reference annotation files for all records) of this database was contributed to PhysioNet in 2000, and the remainder was contributed to November 2009. A list of all 90 record names is available. [7]

**b) MIT-BIH Arrhythmia Database:**

Since 1975, the laboratories at Boston's Beth Israel Hospital and MIT have supported a research into arrhythmia analysis and related subjects. One of the first major products of that effort was the MIT-BIH Arrhythmia Database, which was completed and began to be distributed in 1980. The database was the first generally available set of standard test material for evaluation of arrhythmia detectors, and has been used for that purpose as well as for basic research into cardiac dynamics at more than 500 sites worldwide. The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) at Boston's Beth Israel Hospital; the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that would not be well-represented in a small random sample. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. [8]

### 2.3.2 Pre-processing:

The recorded ECG data usually contain different type of noises which include low and high frequency components. Generally, the presence of such artifacts will corrupt the signal and make the features extraction and classification less accurate. Thus a de-noising procedure is so important to minimize the negative effects of the noise.

**a) ECG noise sources:**

While extracting ECG, many sources of noise are to be encountered unfortunately, with frequency content of some of them overlapping with ECG. While selecting appropriate filters, the final solution will depend on both the types of noise, and on our goals.

The most common types of noise that affect the ECG signal are considered herein:

- Power line interference, it is a 60 Hz and multiples; it goes up to 50% of QRS amplitude.
- Electrode contact noise, Loose contact, motion artifact, baseline drift due to respiration; typically <0.5 Hz signals.

- EMG from the chest wall, its amplitude is about 10% of ECG amplitude; its frequency goes from 20 to 1000 Hz.
- Instrumentation noise.
- Electrosurgical noise.

**b) Filtering /De-noising:**

Since the ECG signals in our study are noisy, some pre-processing needs to be applied in order to eliminate all signals of frequencies less than 1 Hz and higher than 40 Hz. To achieve this goal, a Butterworth band-pass filter as a de-noising filter for ECG signal is selected .

Butterworth filter is a high order design, but it can be realized by using simple first order cascaded together to achieve the desired order. Butterworth has a flat response; its equation is given as follows:

$$H(z) = \frac{b_0+b_1 z^{-1}+b_2 z^{-2}+.....+b_N+z^{-N}}{1+a_1 z^{-1}+a_2 z^{-2}+.....+a_M z^{-M}} \qquad \text{eq (2-1)}$$

Its frequency response is given as follows:

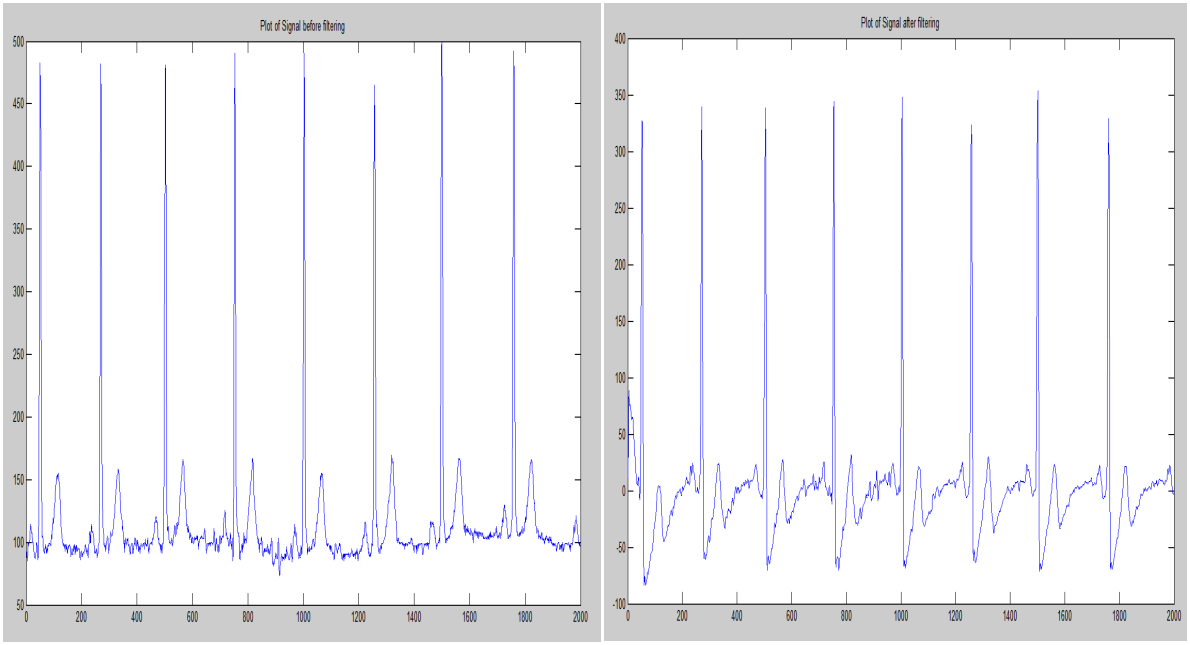
$$|(H(jw))|^2 = \frac{H_0}{1+(w/w_0)^{2n}} \qquad \text{eq (2-2)}$$

On Matlab, the appropriate ECG filter with a second order Butterworth band-pass filter is designed:

```
%filtering the  ECG signal
sampleRate = 256; % Hz
cutOffFreq1 = 1; % Hz
cutOffFreq2 = 40; % Hz
filterOrder = 2; % Filter order (e.g., 2 for a second-order Butterworth filter)
ftype='bandpass';
[b, a] = butter(filterOrder, [cutOffFreq1/(sampleRate/2),cutOffFreq2/(sampleRate/2)],ftype); %
Generate filter coefficients
filtered_signal = filter(b, a, signal); % Apply filter to data using zero-phase filtering
```

The filter is second order because the date was already filtered by the hardware , so increasing the order of the filter would not have given a significant change. The performance of this filter is shown as applied to the 1000 first bits of our first ECG recording ('e01.mat'). Notice a

significant difference between the recordings before and after the filtering step, and the different P, QRS, T waves can easily be observed.




(a) (b)

**Figure 2-7:** *(a) ECG signal plot before filtering; (b) ECG signal plot after filtering.*

**Note:** From now forward, processes will only use filtered signals.

### 2.3.3 Features extraction:

To build an efficient human identification system, the extraction of features can truly represent the distinctive characteristics of a person in a challenging problem. We have methods for ECG-based identity recognition use attributes that are temporal and amplitude distances between detected fiducial points.

Firstly, by focusing only on fiducial points, the representative of discriminate characteristics of ECG signals might be inadequate. Secondly, their methods rely heavily on the accurate localization of wave boundaries, which is generally very difficult.Our study suggests an analytic-based method, an appearance based-features extraction method, which captures the holistic patterns in the heart-beat signal, and where only peak detection is necessary.The proposed approach depends on estimating and comparing the significant coefficients of the Discrete Cosine Transform (DCT) of the auto-correlated heart-beat signals.

The *Features extraction* process of the method used in this project is realized after several steps. As a first step, the highest peaks are localized, constituted by the R peaks in each period of the signal.

**a) R-Peaks Detection:**

```
%Peaks detection
PeaksDetected=ecgdemowinmax(filtered_signal, 256);
Indexes = find(PeaksDetected>0);
%
```

Peak detection is a crucial step in this process; the localization of peaks will allow the capture of a fixed region of the signal. In addition, this will also allow us to capture any desired parts of the (PQRST frames) signal. The peaks in question are called Indexes.

For demonstration, the 6 first R-peaks are plotted starting from the second Index.

The MATLAB commands utilized to perform the task are presented below:

```
%plot of 6 first PQRST waves of our signal
signal=filtered_signal((Indexes(2)-127):(Indexes(7)+128),1);
plot(signal)
%
```

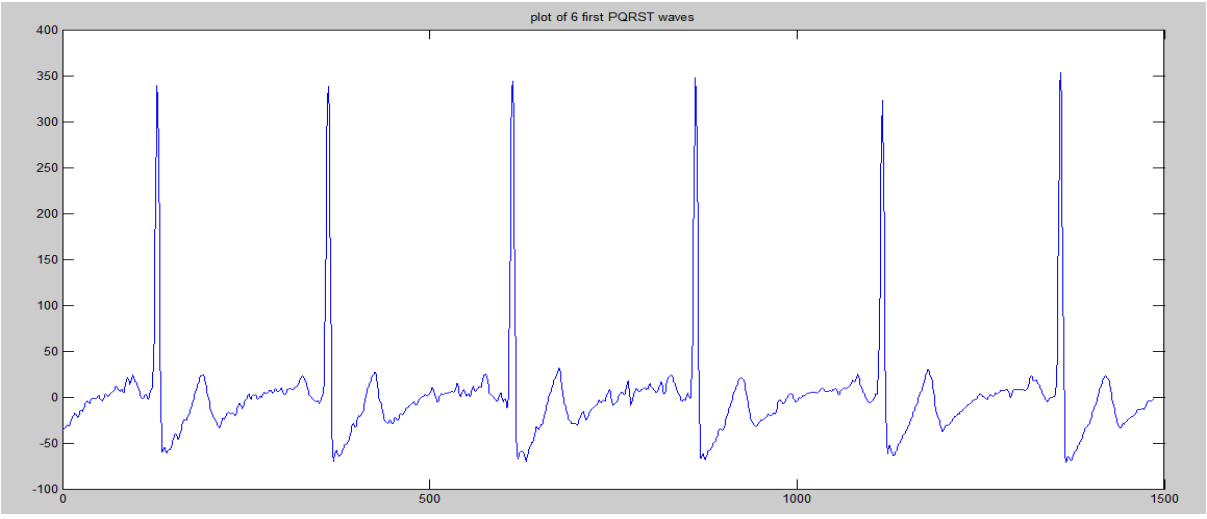



**Figure 2-8:** *Plot of the 6 first PQRST waves.*

In the next step of the Features Extraction phase, various mathematical tools are used in order to obtain the necessary information.

**b) Application of Autocorrelation/Discrete Cosine Transform:**

Autocorrelation is used to eliminate the similarities between observations as a function of the time lag between them. For this purpose, a MATLAB function stored in DSP system toolbox under the name "Autocorr" was utilized. The function is presented in the following syntax:

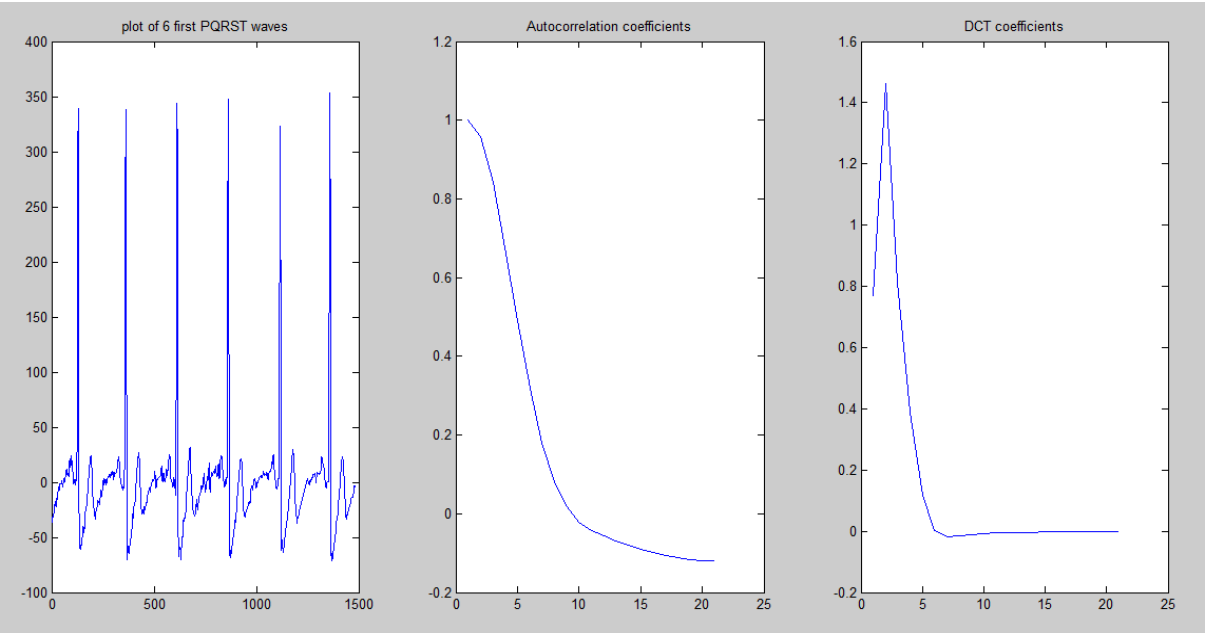
```
Y=AutoCorrelation(X, lags)
```

With X being the input signal, Y the output coefficient and "Lags" being an optional option. Note that lags were not fixed, with 21 lags being taken by default. Next, the obtained (21) AC coefficients are transformed using the DCT function stored in the signal processing toolbox. Its syntax is:

```
Y=dct(X)
```

With X being the input coefficients and Y the Output coefficients.

The AC and DCT coefficients of the same subject as in the previous step are derived using the stated MATLAB functions.




(a) (b) (c)

**Figure 2-9:** *(a) Filtered ECG of subject A; (b) 21 AC coefficients A; (c) DCT coefficients.*

The previous procedure is repeated for the 20 inputted signals; each signal consisting of 23 (complete periods (PQRST)) frames, then the DCT coefficients of each signal are stored in one Matrix for use as reference while testing later. It is called Strain.

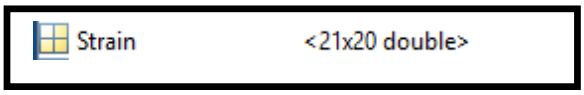



**Figure 2-10:** *The training data that contains the 20 subject's DCTs.*

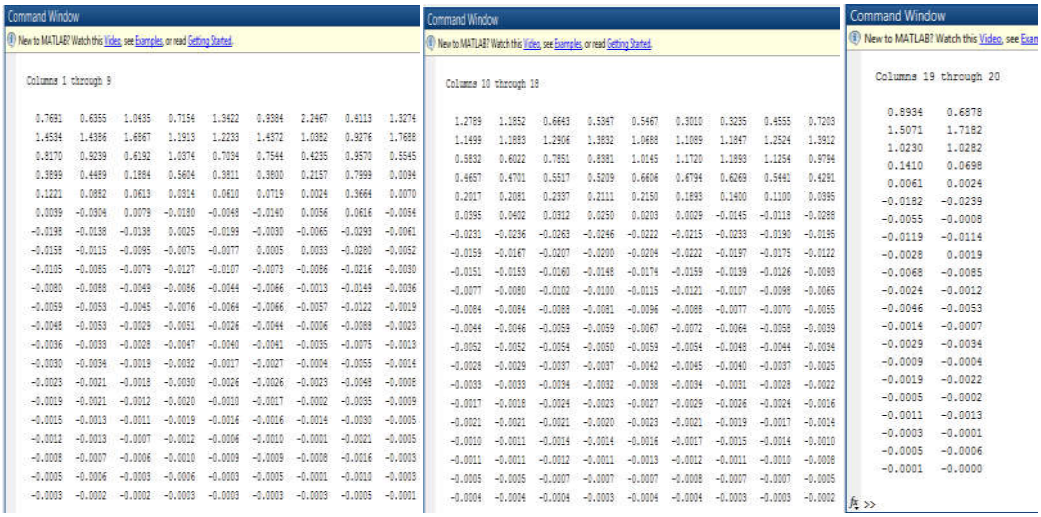



**Figure 2-11:***The 20 subject DCT coefficient vectors.*

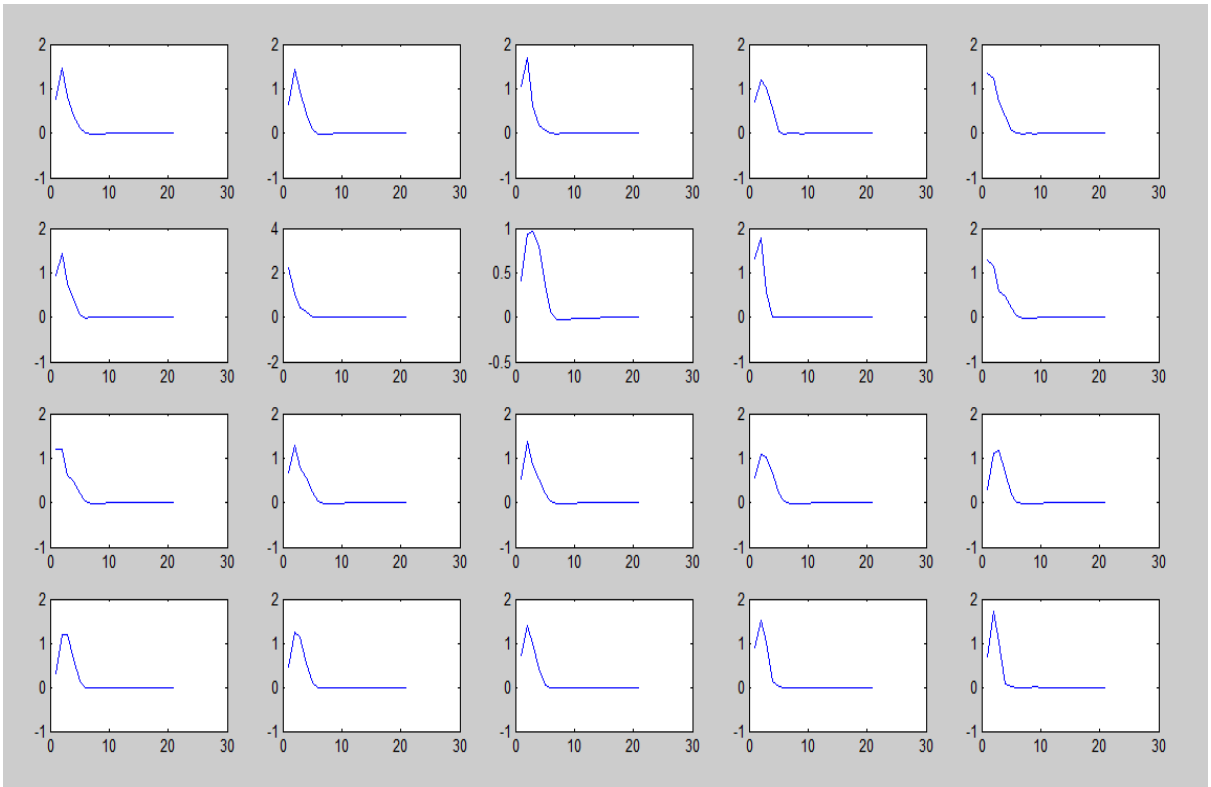


**Figure 2-12:** *Plots of all DCT matrix elements of our Database.*

**c) Thresholds computation:**

In some cases, and due to various reasons, different waves of the same ECG signal differ. One can take advantage of this fact to calculate the differences between each wave of the signal associated to the same person using the Euclidean Distance. This can be used as a Threshold while testing, to avoid the inclusion of “external” people from being identified as subjects. An algorithm that computes the distances between different DCT coefficients of different parts of the same signal was developed; taking Thresholds maxima for each ECG signal. The procedure is further described in the diagram below:

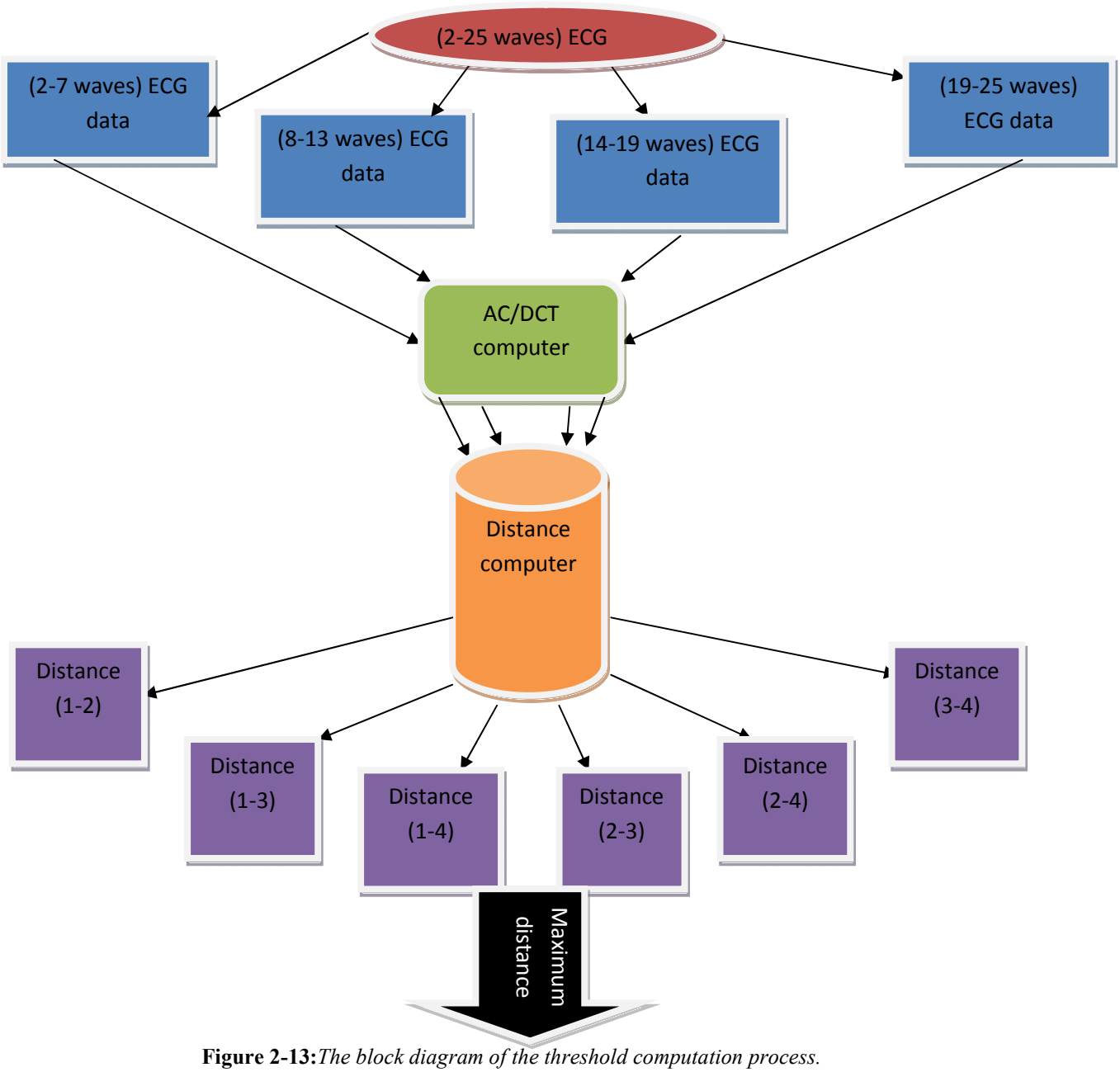



**Figure 2-13:***The block diagram of the threshold computation process.*

Thresholds are stored in one vector under the name ‘Thr’ for use for testing later. The 'Thr' vector is composed of 20 elements; with each element representing the Threshold corresponding to a particular person.

```
>> Thr

Thr =

  Columns 1 through 9

    0.0045    0.0123    0.0027    0.0273    0.0080    0.0204    0.0176    0.0050    0.0514

  Columns 10 through 18

    0.0159    0.0046    0.0058    0.0426    0.0100    0.0175    0.0122    0.0099    0.0063

  Columns 19 through 20

    0.0083    0.0146

>>
```

**Figure 2-14:** *The stored threshold used for testing.*

### 2.3.4 Classification:

Classification is done using the Euclidean distance ,and the nearest neighbor is used as the classifier . The normalized Euclidian distance between two feature vectors x1 and x2 is defined as:

$$D(x_1, x_2) = \frac{1}{N}\sqrt{(x_1 \quad x_2)^T(x_1 \quad x_2)} \qquad \text{eq (2-3)}$$

where N is the dimensionality of the feature vectors , which is the number of DCT coefficients in the proposed method

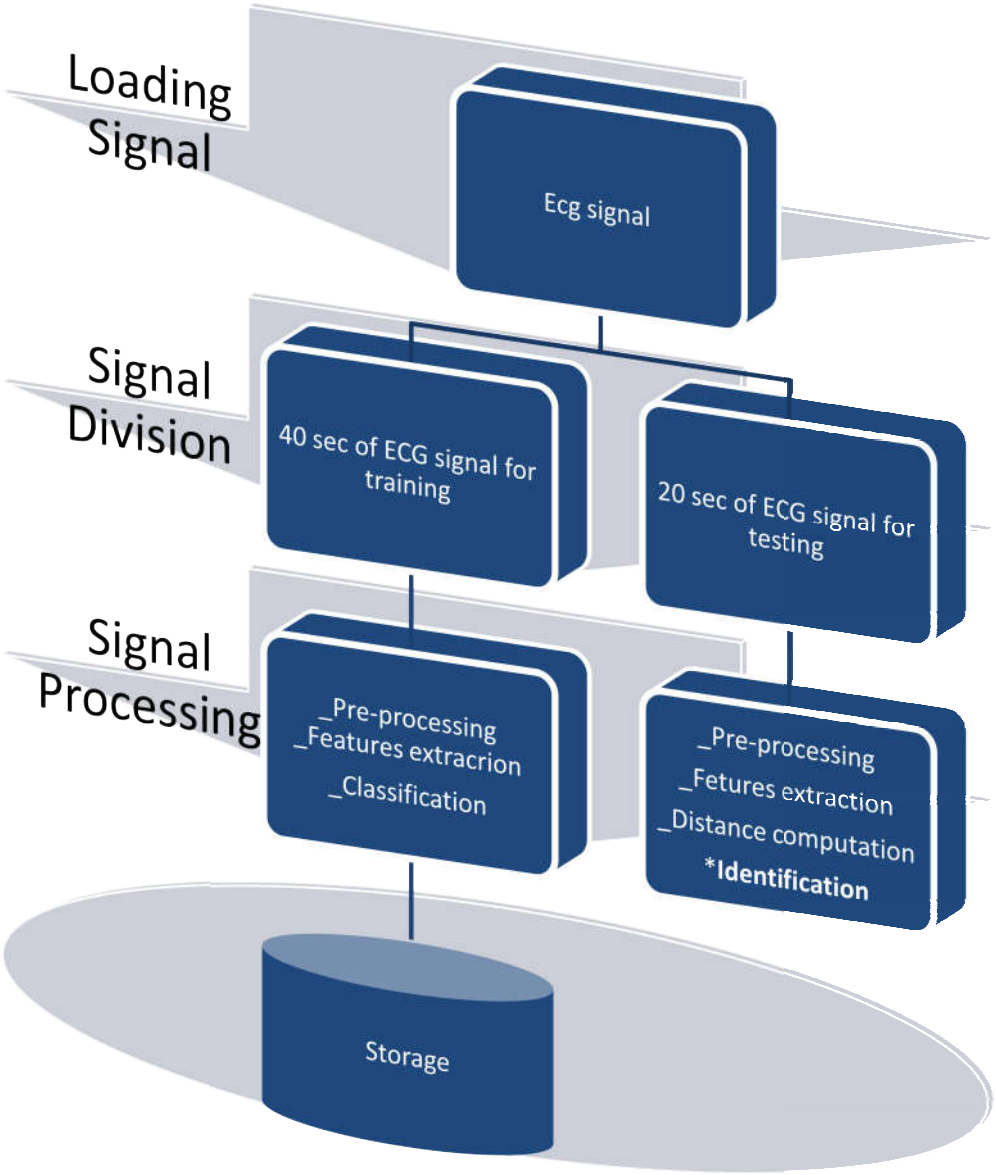



**Figure 2-15:** *The general methodology of the proposed system.*

**Summary :**

In this chapter, a brief introduction to ECG-based biometrics and some generalities about heart and electrocardiogram was provided. Then some ECG basics and the variation of ECG waveform within time were also introduced.

Also described was our proposed model of ECG-based biometrics , and we then explained the procedures followed in order to build the database that will be taken later as reference for identification .

# EEG-BASED BIOMETRICS

This chapter's introduction provides an overview of the brain anatomy and some basics of the electroencephalogram, followed by a methodology and implementation of the proposed EEG system.

## 3.1 Introduction:

Electroencephalography (EEG) records the brain's activities by measuring the voltage fluctuation on the scalp; although, the signals recorded can be influenced by mood, stress, as well as the mental state of the individual being tested. EEG has been used for more than a century in the medical field and also as the basis for brain computer interfaces and brain machine interfaces for assistance, rehabilitative, and entertainment applications. Only recently has EEG been proposed as a biometric trait with potentialities allowing identity recognition. Furthermore, brain signals are related to the subject's genetic information, making them unique for each individual and stable over time. Therefore, EEG holds great promise to provide a far more reliable and secure biometric approach for user identification and authentication. In conventional scalp EEG, recordings are obtained by placing electrodes on the scalp according to the 10-20 international system. Scalp EEG activity shows oscillations at a variety of frequencies, mainly in the [1, 40] Hz range. Several of these oscillations show characteristic frequency content and spatial distributions associated to different states of brain functioning that can be investigated as potential distinctive traits for the purpose of user recognition.

In this chapter, an EEG-based biometric system is presented. Specifically, a noise reduction technique is proposed, by passing the EEG signal over a pre-processing stage, then frequency features will be extracted using the wavelet packet decomposition technique. Finally, a classification process is performed. Before introducing these mathematical and technical steps, a general introduction to brain anatomy science and EEG concepts is presented.

## 3.2 EEG Overview:

### 3.2.1 Definition of the brain:

The brain is the main organ of the human central nervous system. It is located in the head, protected by the skull. It has the same general structure as the brains of other mammals, yet has a more developed cerebral cortex [9]. The brain is the most complex organ in a vertebrate's body. In a typical human being, the cerebral cortex (the largest part) is estimated to contain 15–33 billion neurons, each connected by synapses to several thousand other neurons. The neurons communicate with one another by means of long protoplasmic fibers called axons, which carry trains of signal pulses called action potentials to distant parts of the brain or body targeting specific recipient cells [10].

Mainly, the brain can be divided into three separate areas (figure 3-1), that are the cerebrum, the cerebellum and the brain stem.

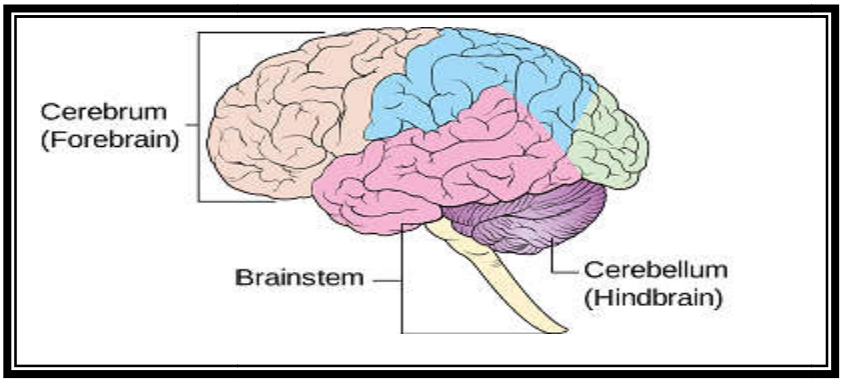



**Figure 3-1:** *Brain's areas and anatomy [11].*

### 3.2.2 The brain's activity:

For this study, the interest area is the cerebrum since the electrical activity recorded by the Electroencephalograph is extracted from the rich nerve fibers situated in the cerebrum area. Neurons, which are also known as nerves cells differ widely but share some basic characteristics. All neurons contain a cell body, a nucleus, and an axon. High levels of positively charged potassium ions (K+) are contained inside the neurons, whereas (K+) levels are low outside the neurons. The opposite is true for positively charged sodium ions (Na+), so there are low levels of (Na+) inside the neurons and high levels outside. The membrane of the neuron is permeable and positively charged potassium (K+) ions are able to leak out the neuron and similarly positively charged sodium (Na+) ions are able to leak into the neuron. The membrane of the neuron is selectively more permeable to the positively charged potassium ions than the positively charged sodium ions, that is to say it is easier for the positively charged potassium ions to leave the neuron than it is for the positively charged sodium ions to enter the neuron. This activity results in an eventual loss of positive charge within the neuron. Once the inside of the neuron reaches –70 mV, the permeability of the neuron membrane changes and no longer allows the escape of positively charged potassium ions. It is at this stage that there is a level of equilibrium found, and this is known as the neuron's resting potential. One of the human body's most mysterious functions is the electrical activity in the brain. The electrical activity in the brain can be measured and defined as five distinct electrical levels, with the brainwaves known as Beta, Alfa, Theta, Delta and Gamma (table 3-1). The electrical activity is defined in "Hz", directly related to the physical and mental activity of the human.

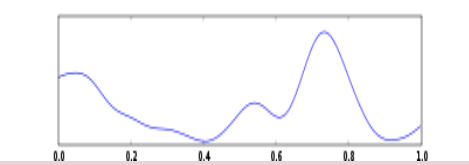
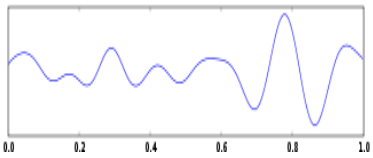
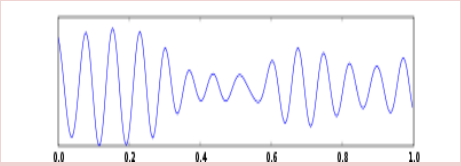
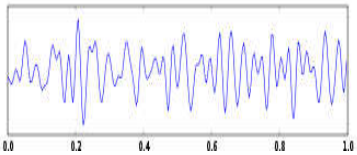
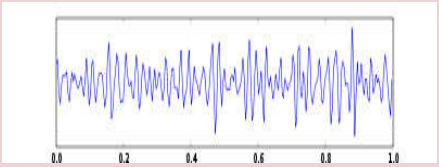
| The wave | Frequency range | Characteristics |
|---|---|---|
| **Delta rhythm**  | Up to 4 Hz | The highest in amplitude and the slowest waves. It is seen normally in adults in slow wave sleep and in babies. |
| **Theta rhythm**  | (4-8) Hz | Theta is seen normally in young children, in drowsiness or arousal in older children and adults |
| **Alpha rhythm**  | (8-15) Hz | It was the first rhythmic EEG activity seen by Hans Berger, It emerges with closing of the eyes and with relaxation, and attenuates with eye opening or mental exertion. |
| **Beta rhythm**  | (15-30) Hz | It is often associated with active, busy or anxious thinking and active concentration and it is the dominant rhythm in patients who are alert or anxious or who have their eyes open. It may be absent or reduced in areas of cortical damage and is generally attenuated during active movements. |
| **Gamma rhythm**  | Greater than 30 Hz | The highest in frequency and associated with rapid eye movement (REM) during human REM sleep. |

**Table 3-1:** *The different brainwave types each with its specific characteristics.*

### 3.2.3 EEG definition:

EEG is the abbreviation of "Electro-Encephalo-Gram"; it is a recording of the electrical changes occurring in the brain produced by electrodes placed on the scalp and amplifying the electrical potential developed as demonstrated in figure (3-2) [12].

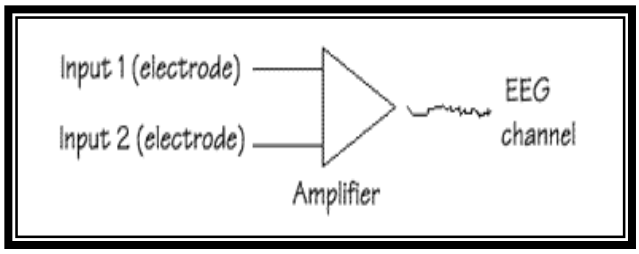



**Figure 3-2:** *The derivation of an EEG channel [12].*

### 3.2.4 Multi-channel EEG:

An EEG channel represents an activity captured by an electrode. Electrode placement on the head adheres to a formal standard called the 10/20 system or International 10/20 system (figure 3-3).

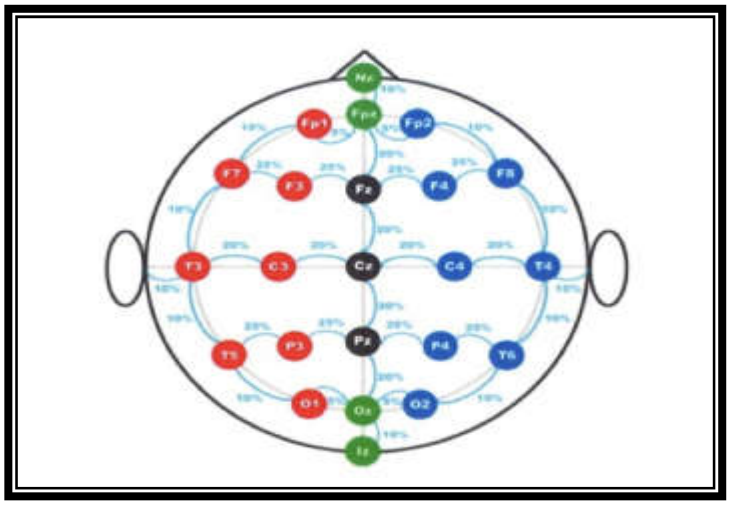



**Figure 3-3:** *Multi channel EEG electrode placement [32].*

Various parts of the brain serve different functions. Placement systems like 10/20 attempt to standardize the exact positioning of each electrode (and assign symbols to each location such as Cz or P3) to access brainwave data at a specific brain location that serves specific brain functions. For example, placements Fp1, Fpz and Fp2 collect important prefrontal cortex data, and placements at F7, F6, F5, etc., capture frontal lobe data. Typical EEG systems can have from as few as a single channel to as many as 256 channels. But using larger electrode arrays (dense array EEG) and expanding to more channel EEG systems provide several advantages, among them:

- A Multi-channel setting does a better job avoiding the loss of any crucial data (caused when electrode distances grow further apart when fewer are deployed). Biophysical analyses have shown that information is lost unless an inter-sensor distance of one to two centimeters is achieved with EEG sampling.
- Detecting important clinical signals. Medical use cases need higher resolution EEG systems (larger sensor networks) to get the job done *[13]*.

**Here's an important and relevant question: does an increase in the number of channels provide better results when dealing with EEG based biometrics?**

## 3.3 EEG Based Biometrics Methodology and Implementation:

In this chapter, a biometric system based on EEG signal for user identification and authentication is developed figure (3-4). A pre-processing stage which reduces the signal noise by a low pass filter is first presented, then the needed features are extracted using wavelet packet decomposition as well as some statistical parameters (mean, standard deviation, and entropy). Finally a Classification based on absolute distance and decision is proposed using the NN (nearest neighborhood) method.

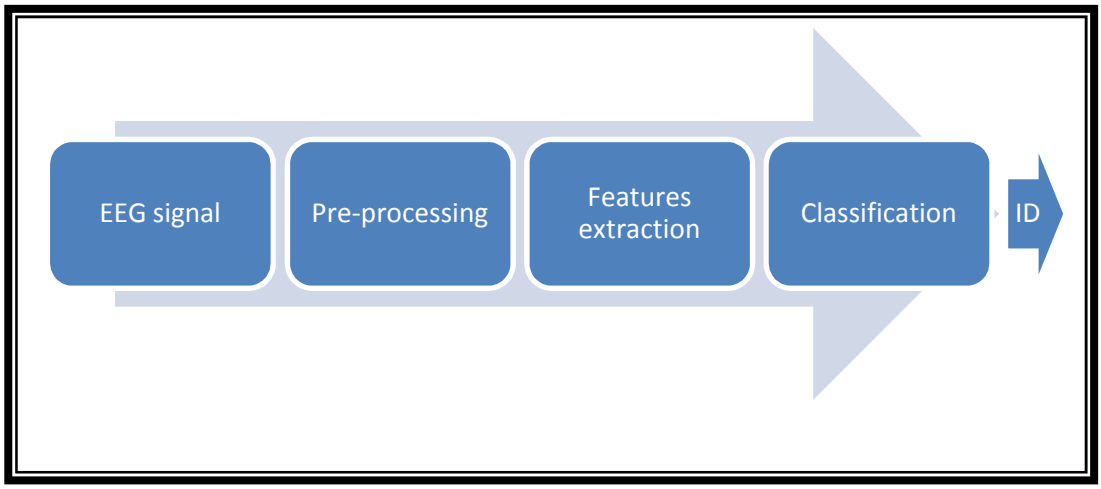



**Figure 3-4:** *Block diagram of the proposed EEG biometric system.*

### 3.3.1 EEG Database:

To realize the proposed biometric system, our research was developed based on two sets of database, one public and the other private.

**a) Public Database:**

Public database consists of a total of 40 EEG recordings obtained from twenty volunteers each with a two minute length. This means that two recordings are associated to each volunteer, one reserved for the training part, while the other is for testing. All 40 recordings were collected from the EEG Motor Movement/Imagery Dataset described below, knowing that they are all from the same task (Task 3), which will also be described later.

**EEG Motor Movement/Imagery Dataset:**

The EEG Motor Movement/Imagery dataset was created and contributed to PhysioNet by the developers of the BCI2000 instrumentation system, which they used in making these recordings. This dataset consists of over 1500 one-minute and two-minute EEG recordings, obtained from 109 volunteers. Subjects performed different motor/imagery tasks while 64-channel EEGs were recorded, each sampled at 160 samples per second using the BCI2000 system. Each subject performed 14 experimental runs: two one-minute baseline runs (one with open eyes, the other with shut eyes), and three two-minute runs of each of the four following tasks:

1) A target appears on either the left or the right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes.

2) A target appears on either the left or the right side of the screen. The subject imagines opening and closing the corresponding fist until the target disappears. Then the subject relaxes.

3) A target appears on either the top or the bottom of the screen. The subject opens and closes either both fists (if the target is on top) and both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes.

4) A target appears on either the top or the bottom of the screen. The subject imagines opening and closing either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes [14].

**b) Private Database:**

Our private Database consists of a total of twelve EEG recordings obtained from twelve volunteers each of one minute length. This means that the first 30 seconds is reserved for the training part and the second 30 seconds is for the testing stage. All the 12 recordings are real EEG data of healthy patients who attended an EEG session at the office of a

Neurology specialist located in Algiers. Subjects were tested with 19-channel EEG system recording data at a sampling rate of 500 samples per second (500 Hz).The first 18 channels recorded brain activities while channel 19 recorded the heartbeat (ECG) of the subject, noting that for every record of each subject, both EEG and ECG records are obtained.

**Note:**

From now forward, all programs and analyses shown will be conducted on the public database (EEG Motor Movement/Imagery Dataset) as shown in figure (3-5), the reason being that this research can be as reference for further work, thus providing the possibility of comparing our obtained results with future ones. For better visualization, each recording was attributed to a person of 20 students chosen from the IGEE institute. Names of subjects and their corresponding photos were also added.

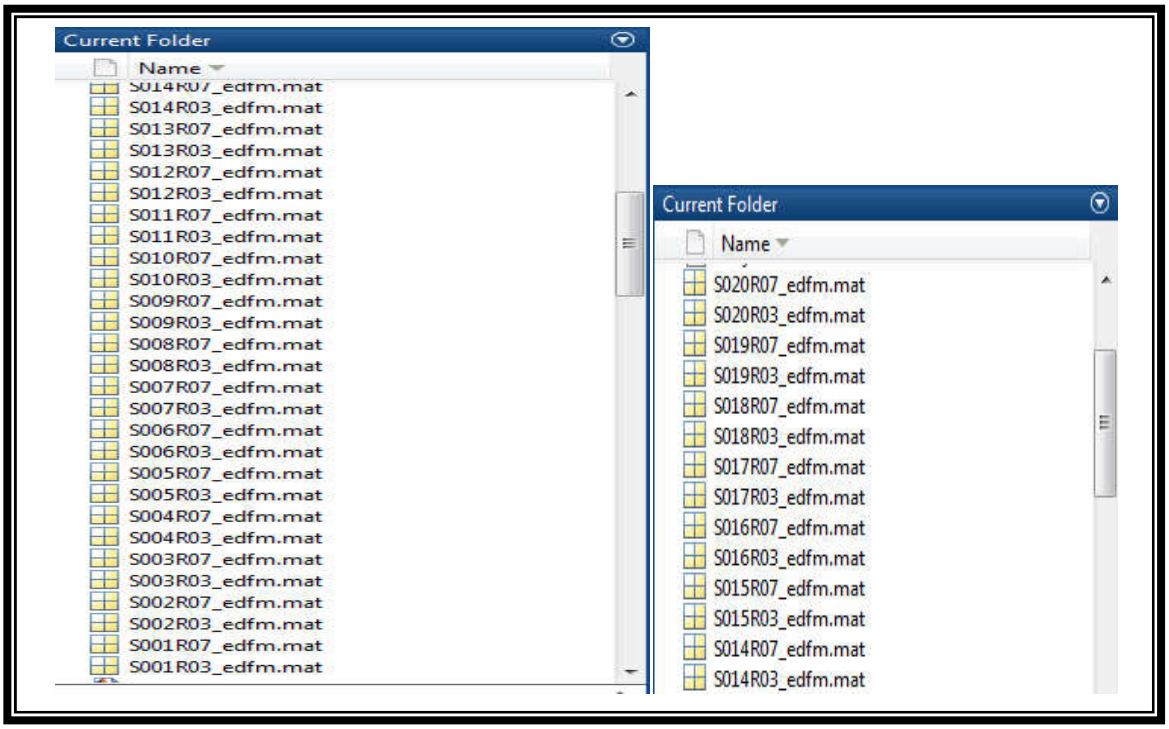



**Figure 3-5:** *EEG database (R03 EEGs are for local database; R07are for intruders).*

### 3.3.2 Pre-processing:

The EEG is usually analyzed by physicians and medical specialists particularly for the detection of neural rhythms. However; it is generally contaminated with various noise sources and mixed with other biological signals whose common artifact sources are the power line interference (50 or 60 Hz), in addition to the ECG and EOG signals. To correct, or remove the artifacts from the EEG signal without losing necessary data, a Butterworth filter has been adopted, as explained in a previous chapter.

Depending on the types of EEG noises, and taking into consideration the frequency bands of the five rhythms that contain the necessary features, this yields the use of a low-pass Butterworth filter with a frequency band of 60 Hz. For this purpose, and using MATLAB, a second order low-pass filter was designed as follows:

```
%filtering the signals
Sample_Rate = 160; % Hz
cut_Off_Freq = 60; % Hz
filter_Order = 2; % Filter order (e.g., 2 for a second-order Butterworth filter)
[b, a] = butter(filter_Order, cut_Off_Freq/(sample_Rate/2)); % Generate filter coefficients
filtered_signal = filtfilt(b, a, eeg_signal); %Apply filter to data using zero-phase filtering
```

The performance of this filter is shown when applied to the 160 first samples of our first EEG recording ('S001R03_edfm.mat').One can observe from figure (3-6) a significant difference between the recording before and after the filtering step.

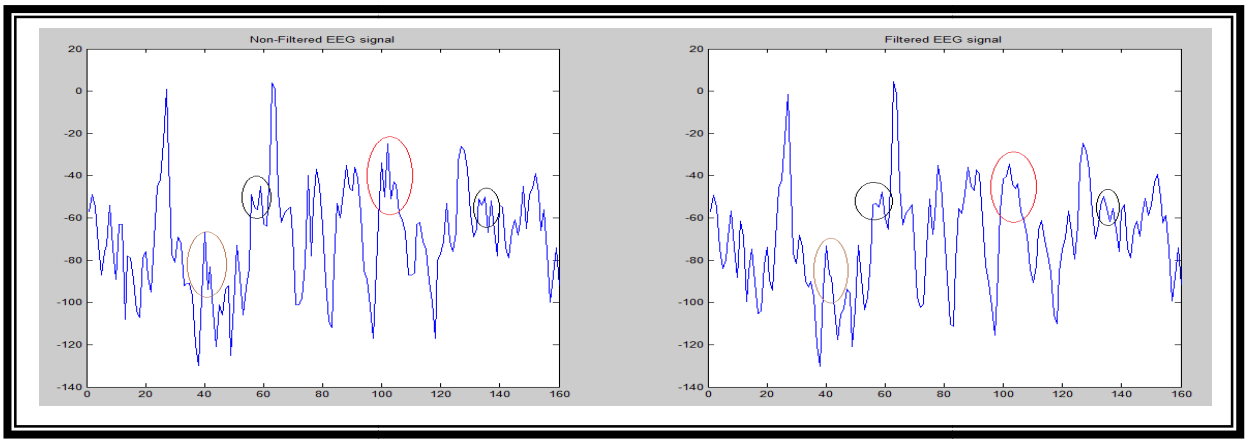



**Figure 3-6:** *Portion of an EEG signal before and after the filtering process.*

### 3.3.3 Features extraction:

After the pre-processing stage, a filtered EEG signal suitable for extracting the needed features was obtained. In our work a method was used merging two kinds of feature extraction techniques, dealing with both the time domain and frequency domain features extraction. Statistical features, such as mean, standard deviation and entropy, belong to the time domain features extraction, whereas frequency domain features extraction uses the wavelet packet decomposition (WPD), which analyzes the frequency distribution of the EEG signal.

**a) Wavelet Packet Decomposition:**

Originally known as the Optimal Sub-band Tree Structuring (SB-TS), it is also called Wavelet Packet Decomposition (WPD) [15]. WPD is a down-sampling process where the discrete-time (sampled) signal is passed through multi-level filters to analyze the time-frequency information.

**b) The mean:**

The mean is the average of the numbers, in other words it is the sum divided by the count. To calculate it, all the numbers are added up, then divided by the sum of numbers added [16].

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \text{eq (3-1)}$$

**c) Standard deviation:**

In statistics, the standard deviation (SD, also represented by the Greek letter sigma σ or s) is a measure used to quantify the amount of variation or dispersion of a set of data values [17].

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i \quad \mu_x)^2} \qquad \text{eq (3-2)}$$

**d) Entropy:**

Entropy is a statistical measure of randomness used to characterize the texture of the input data. Entropy is defined as:

$$\varepsilon(x) = \quad \sum_t x^2(t)\log\left(x^2(t)\right) \qquad \text{eq (3-3)}$$

For different individuals, the energy distribution of the frequency components between individuals is quite different and this option makes it possible to adapt those frequency components as features to represent the EEG signals. Improvising from this fact, a 4-Level WPD is applied to each channel of the pre-processed EEG signal for each subject (table 3-2), and only coefficients from nodes{(1,1),(2,1),(3,1),(4,0) and (4,1)} representing the frequency bands of the main 5 EEG rhythms {Gamma, Beta, Alpha, Delta, and Theta} are extracted (figure 3-7). This particular study was performed with the help of a MATLAB function stored in the wavelet toolbox called (wpdec). This MATLAB function returns a wavelet packet tree (T) corresponding to the wavelet packet decomposition of vector X at level L. Extraction of the five nodes from tree (T) was then conducted by another MATLAB function (wpcoef) of

the wavelet toolbox. (wpcoef) returns the coefficients associated with the node N of the wavelet packet tree (T).

<table>
<tr><td colspan="16">(0,0) 0-60 Hz</td></tr>
<tr><td colspan="8">(1,0) 0-30 Hz</td><td colspan="8">(1,0) 30-60 Hz “Gamma”</td></tr>
<tr><td colspan="4">(2,0) 0-15 Hz</td><td colspan="4">(2,1) 15-30 Hz “Beta”</td><td colspan="4">(2,2) 30-45 Hz</td><td colspan="4">(2,3) 45-60 Hz</td></tr>
<tr><td colspan="2">(3,0)<br>0-8 Hz</td><td colspan="2">(3,1)<br>8-15 Hz<br>“Alpha”</td><td colspan="2">(3,2)<br>15-23 Hz</td><td colspan="2">(3,3)<br>23-30 Hz</td><td colspan="2">(3,4)<br>30-38 Hz</td><td colspan="2">(3,5)<br>38-45 Hz</td><td colspan="2">(3,6)<br>45-53 Hz</td><td colspan="2">(3,7)<br>53-60 Hz</td></tr>
<tr><td>(4,0) 0-4 Hz “Delta”</td><td>(4,1) 4-8 Hz “Theta”</td><td>(4,2)8-12Hz</td><td>(4,3)12-15Hz</td><td>(4,4)15-19Hz</td><td>(4,5)19-23Hz</td><td>(4,6)23-27Hz</td><td>(4,7)27-30Hz</td><td>(4,8)30-34Hz</td><td>(4,9)34-38Hz</td><td>(4,10)38-42Hz</td><td>(4,11)42-45Hz</td><td>(4,12)45-49Hz</td><td>(4,13)49-53Hz</td><td>(4,14)53-57Hz</td><td>(4,15)57-60Hz</td></tr>
</table>

**Table 3-2:** *The 4-Level wavelet packet decomposition tree and the extracted nodes [18].*

The following MATLAB code does the previously stated task:

```
channel_dec = wpdec(filtered_channel,4,'db1');%wavelet packet decomposition.

gamma_rhythm = wpcoef(channel_dec,2);%node(1,1)coefficient extraction.
beta_rhythm = wpcoef(channel_dec,4); %node(2,1)coefficient extraction.
alpha_rhythm = wpcoef(channel_dec,8); %node(3,1)coefficient extraction.
delta_rhythm = wpcoef(channel_dec,15); %node(4,0)coefficient extraction.
theta_rhythm = wpcoef(channel_dec,16); %node(4,1)coefficient extraction.
```

**Note:**

The node (0,0) can be represented by a number 0, the node (1,0) is represented by a 1, and all other nodes are represented similarly.

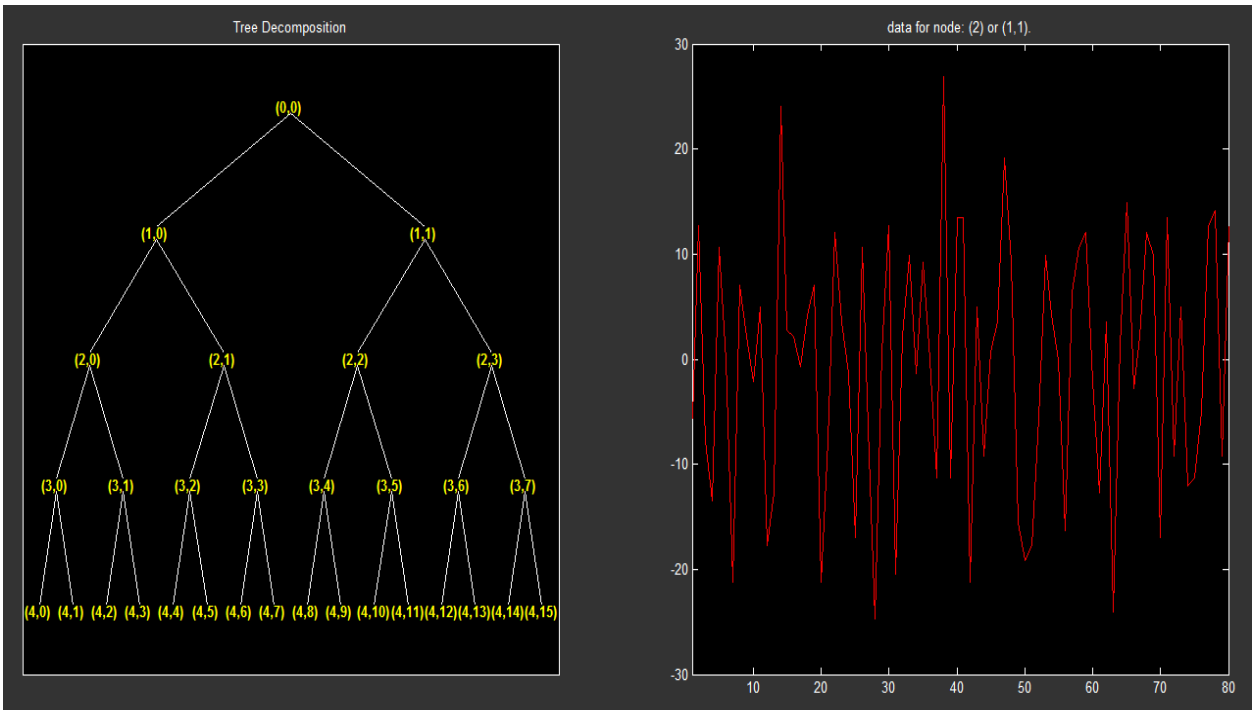



**Figure 3-7:** *MATLAB demonstration of the 4-level tree with the data of node (1, 1).*

To form the feature vectors: the mean, standard deviation and entropy were calculated for each node of every channel figure (3-9). Since there are three statistical parameters and five nodes, there are 3*5=15 features obtained for each channel associated to every subject. Finally The 64 channels data are stored, knowing that each data contains mean, standard deviation and entropy data of the five rhythms. These features were stored for future use as data for the training part while testing as shown in figure (3-8).

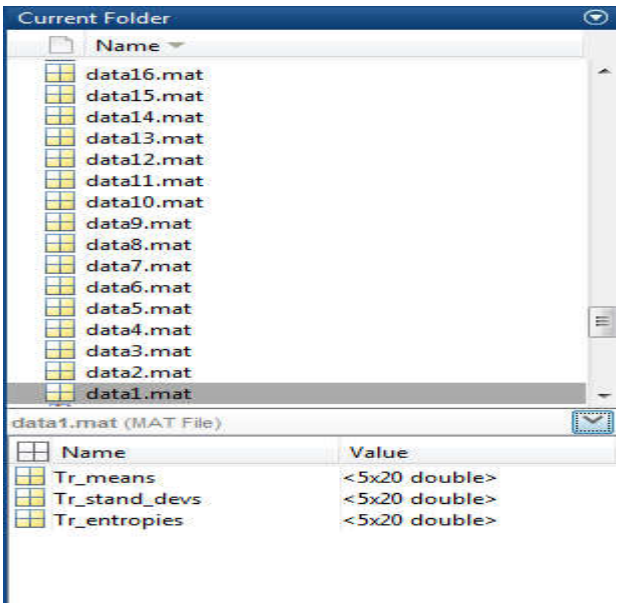



**Figure 3-8:** *The first 16 channels stored data.*

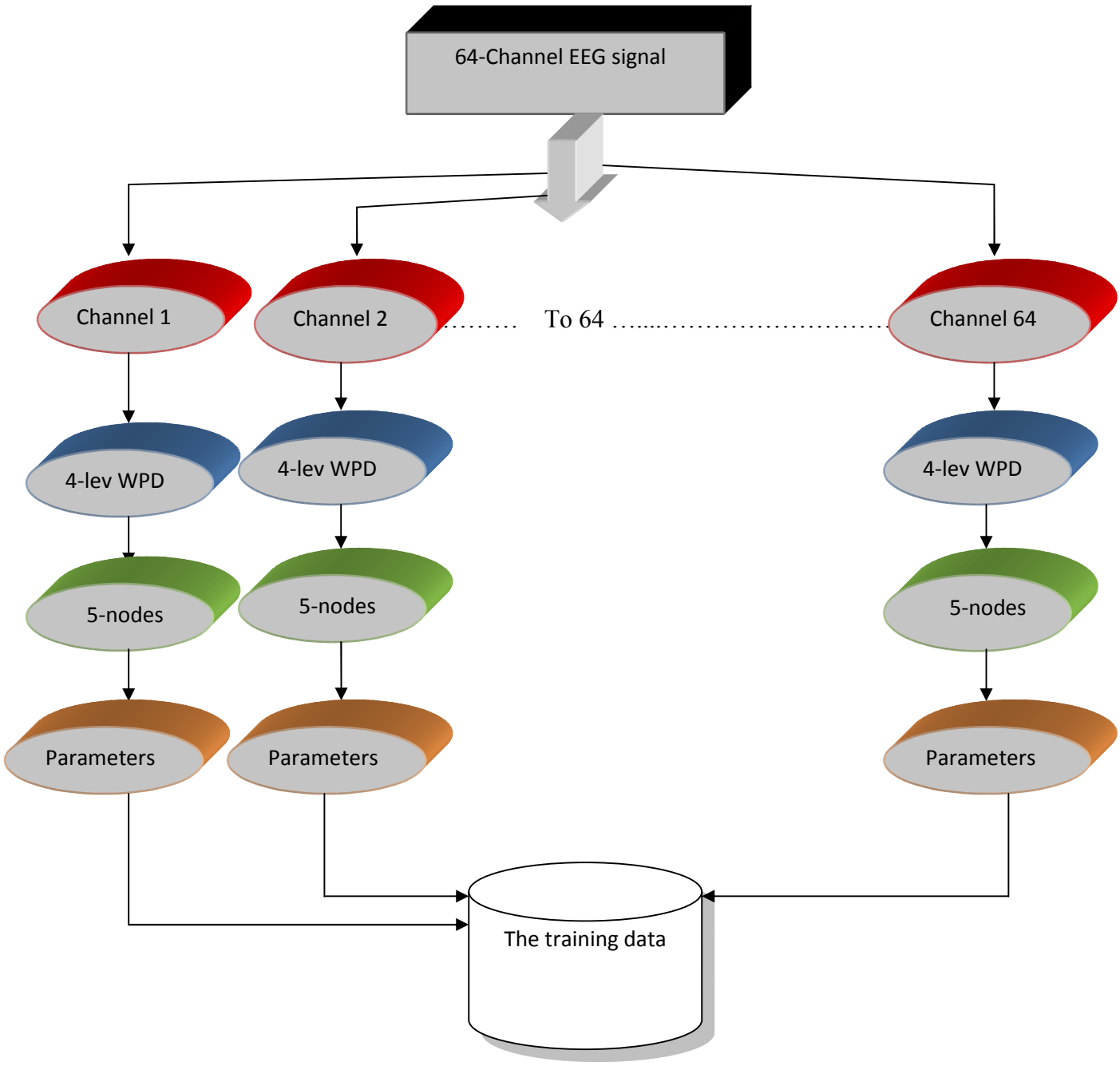



**Figure 3-9:** *General block diagram of the features extraction process.*

### 3.3.4 Classification:

In general, the classification process goal is to check the identity of an input vector to the feature vectors that has been stored in the data base. This work is based on the absolute distance, where the nearest neighborhood (NN) method is used as a classifier.

**a) Absolute distance:**

The absolute distance of two real numbers x, y is given by |x y|, the absolute value of their difference. This represents the distance of the real line between the points corresponding to x and y.

**b) Nearest neighborhood:**

Nearest neighborhood classifier is a method used in statistical classification (or pattern recognition). This method classifies objects based on closest training examples in the feature space [19].

To conduct this final stage, the absolute distance was proposed, calculating the real distance between the feature vectors of the 64 channels of the testing subject on one hand, and their corresponding feature vectors of the stored training data on the other. And so, the subject identification decision process is based on the nearest neighborhood method.

**Summary:**

In this chapter, a brief introduction of the brain's anatomy and some important electroencephalogram concepts were introduced. The methodology of the proposed EEG-based biometric system was detailed throughout explaining the mathematical and technical tools used for its realization, and finally, the implementation of the tools and processes used was also presented.

## *Experimental part:*

This chapter includes an experimental demonstration of the performance of thesystems proposed in this study. It also provides statistical aspects of the ECG and EEG based biometrics, in addition to comparisons of results, showing both advantages and limitations.

Experimental testing and evaluations of the performance of the proposed biometric systems are conducted herein. Analyses of results and efficiency were accomplished in order to compare this study with existing related works as well as to present relevant recommendations and suggestions for future research.

## 4.1 ECG based biometric system:

First, an experimental demonstration of the proposed ECG based biometric system was conducted by testing one local data; all steps will be shown:

### 4.1.1 Testing:

In the testing part, the signal of the person intended for testing is first loaded, and the remaining 5000 bits of the signal corresponding to the 20 seconds for testing is then taken.

In this example, the signal corresponding to the person number 11, with the name 'Ferhat' is first tested.

```
%loading signal for testing
load('e11.mat');
signal2=val(10001:15000)';
%
```

Procedures followed for this signal are similar to those used for the training part, the Pre-Processing, and Features extraction parts. At the end, DCT coefficients of the testing signal are therefore obtained.

The Euclidian Distance between these DCT coefficients and DCT coefficients of each DCT database stored in the system (Figure 2-11) are computed, thus obtaining 20 distances.

The algorithm chooses the smallest distance and it selects the most appropriate person as delegate. In addition, one must test whether this distance is below the threshold corresponding to subject under study. If the results distance is larger than the corresponding threshold, the identification should be rejected; otherwise, identification is confirmed.

After running the program, MATLAB displays the minimum distance , its location (in order to determine the appropriate person) , the threshold corresponding to the person chosen as delegate, then the decision whether access is accepted or denied , and finally a display of the name of the person selected.

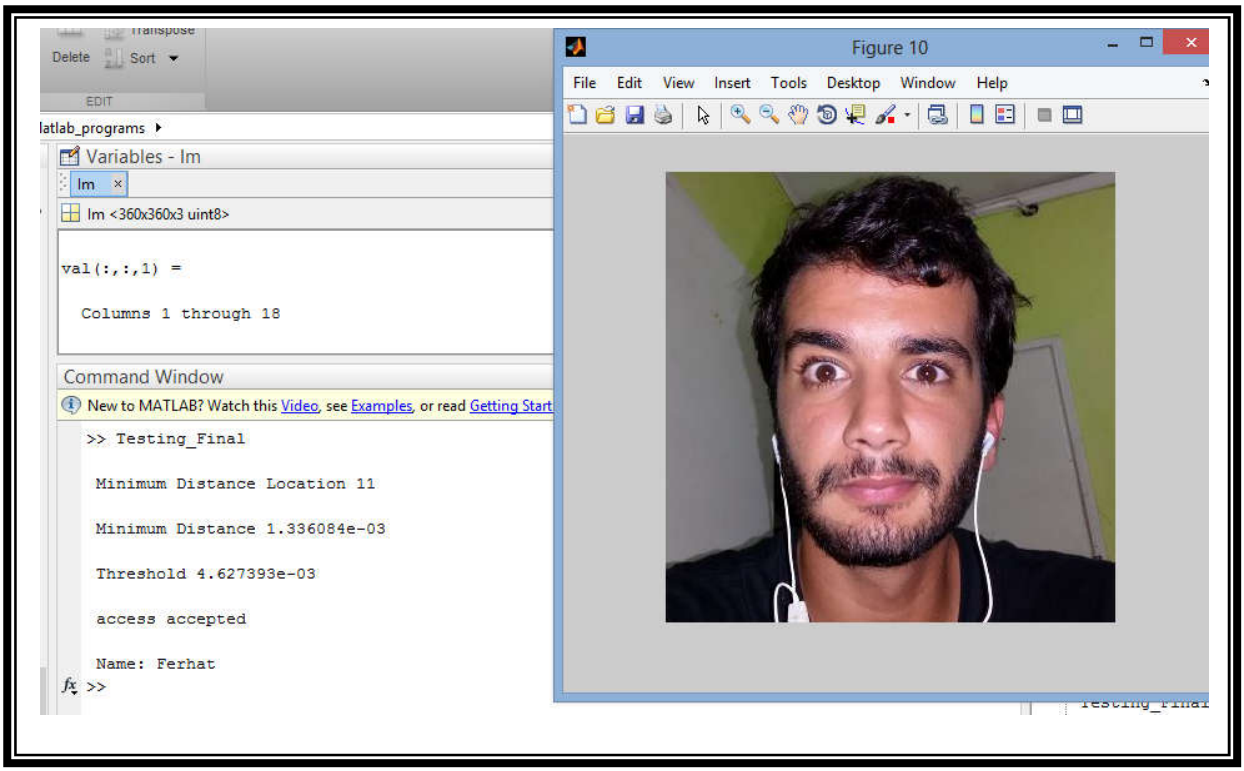



**Figure 4-1:** *The result of running the program when loading ('e11.mat') which is the data of 'Ferhat'.*

As can be seen when loading the ECG signal of person number 11, the minimum distance equals 1.33e-03, less than the threshold corresponding to the person with the Threshold Thr=4.62e-03, so the person is therefore recognized.

On the other hand,, when loading the 'e0704' signal which is not in our database, the minimum distance is 6.82e-03, higher than the Thr=2.69e-03 threshold, so access is therefore denied.

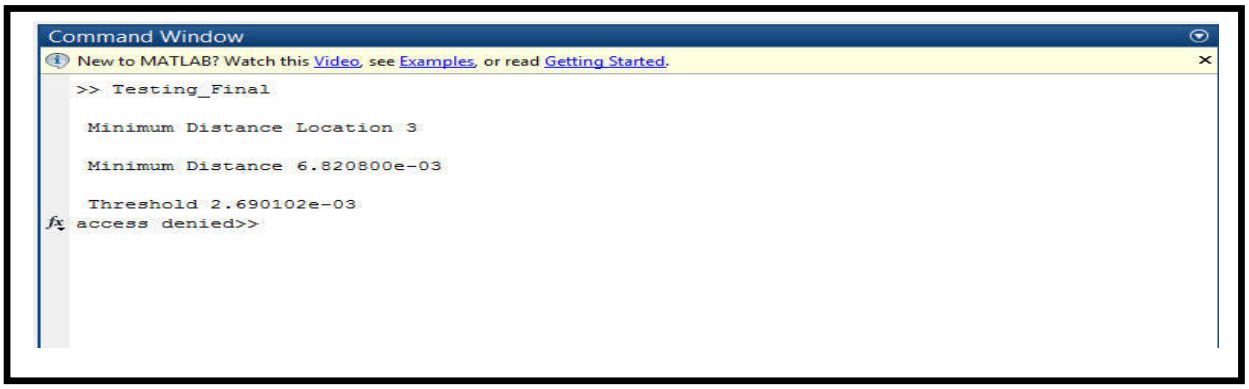



**Figure 4-2:** *The result of running the program when loading ('e0704.mat') which is not on the stored data.*

### 4.1.2 Performance evaluation:

**Experiment 1:**

In this part, a statistical study of our AC/DCT based algorithm is performed on both MIT-BIH and European ST-T databases. The study's interest is to test the overall identification rate for both MIT and European recordings as our own database. Training portions were 2/3 of the overall length of the database, while testing was only 1/3.

In order to test the effectiveness of our algorithm of identification, iterative execution of testing with the variation of the number of PQRST frames provides the following tabulated results and plots:

| Number of PQRST frames | Accuracy (%) | | |
|---|---|---|---|
| | European ST-T database With AC/DCT | MIT-BIH database With AC/DCT | MIT-BIH database with PCA |
| 2 | 93.75 | 53.84 | N/A |
| 3 | 96.81 | 57.08 | N/A |
| 4 | 97.50 | 57.27 | 96.91 |
| 5 | 98.33 | 60.50 | 97.50 |
| 6 | 98.75 | 61.66 | 97.56 |
| 7 | 98.57 | 65.62 | 97.32 |
| 8 | 97.50 | 67.14 | 97.39 |
| 9 | 98.00 | 70.083 | 97.22 |
| 10 | 97.50 | 73.00 | 97.08 |
| 11 | 95.00 | 72.50 | 97.34 |
| 12 | 95.00 | 75.00 | 97.22 |
| 13 | 100.00 | 72.50 | 97.43 |
| 14 | 100.00 | 70 | 97.32 |

**Table 4-1:** *Accuracy values for 2 different databases performed on 2 different algorithms with the variation of number of the PQRST frames.*

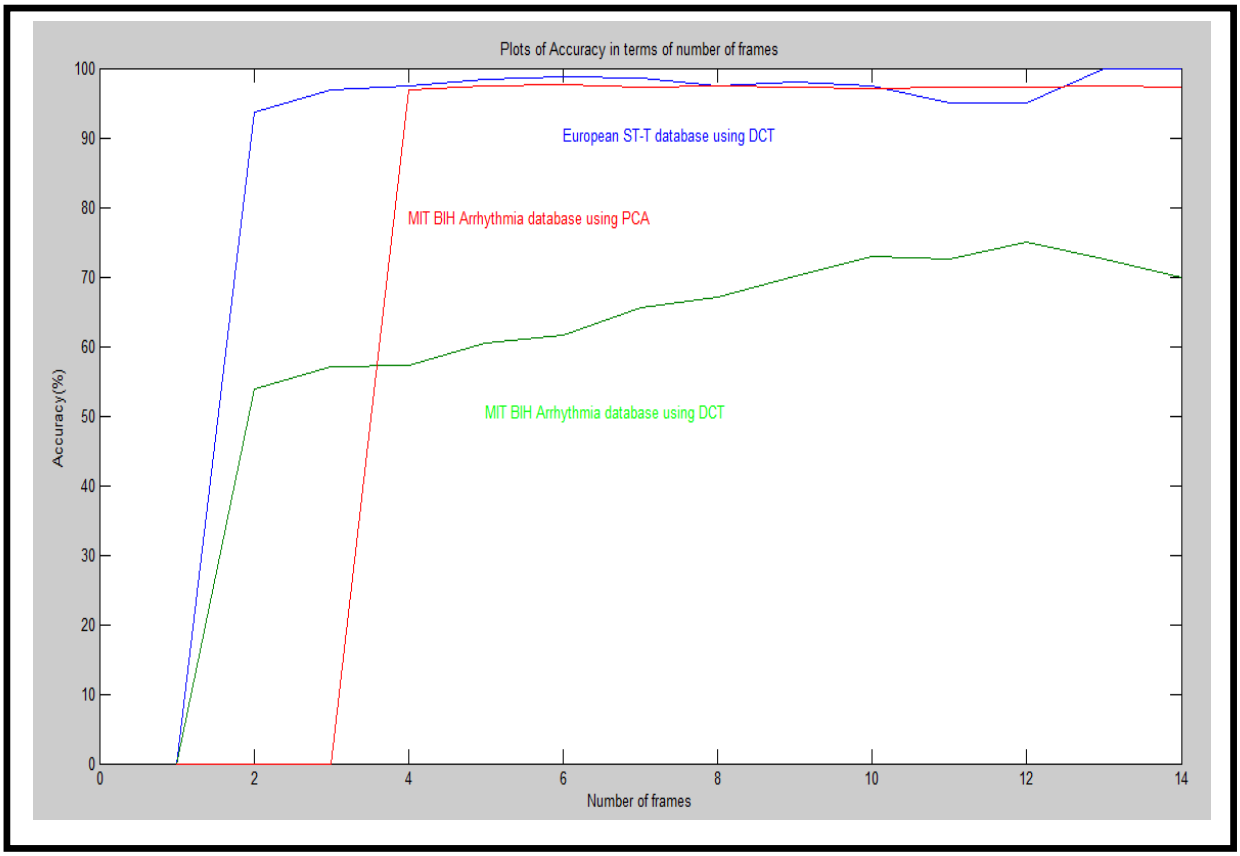



**Figure 4-3:** *Plot of accuracy for each case of the three tabulated cases in terms of Number of PQRST frames.*

**Note:** For comparison purposes, MIT-BIH Arrhythmia database with PCA was extracted from a related work under the name "Implementation of a Biometric identification system using Electrocardiogram ECG/EKG" which was conducted by a former IGEE student.

From the table 4-1 and figure 4-3, notice a significant difference between the results obtained when using AC/DCT and those using PCA methods on the MIT-BIH database. PCA works a lot better with this Database, which is logical given the fact that the data was obtained from patients with Cardiac diseases, i.e., which implies abnormal ECG with a shape modified in a way that alters the PQRST waves, thus significantly affecting the system based on the AC/DCT method as this method deals solely with normal and complete PQRST frames.

Comparing the results obtained from using AC/DCT on MIT-BIH database and while using the same method on the European ST-T database provides a clear advantage to AC/DCT with European ST-T database, for the same reasons mentioned above.

Comparing the results realized on European database using AC/DCT while using PCA method on MIT database, we notice that AC/DCT provides better results when taking less

than 3 frames, and when taking more than 11 frames; while the difference is negligible in-between.

From this experiment, we conclude that when dealing with the MIT database (i.e., the majority of persons are sick). PCA provides good results, because the PCA method focuses on fiducial points which are usually not affected by heart diseases. For this reason, when dealing with the European database (i.e., the majority of persons are healthy), DCT provides us with better results.

**Experiment 2:**

The performance of the proposed ECG identification/ authentication system in Chapter 2with the presence of intruder signals was evaluated by the following parameters [20], (note that identification is among N number of trials in the experiment) :

TP: correctly identified.

TN: correctly rejected.

FP: incorrectly identified.

FN: incorrectly rejected.

- FPIR: False Positive Identification Rate; refers to the Probability of a test sample falsely identified as a subject.

$$\mathrm{FPIR} = \frac{1}{\mathrm{N}} \sum_{\mathrm{n}=1}^{\mathrm{N}} \frac{\mathrm{FP}}{\mathrm{TP+TN+FP+FN}} \qquad \text{eq (4-1)}$$

- FNIR: False Negative Identification Rate; refers to the Probability of a test sample falsely identified as different subject.

$$\mathrm{FNIR} = \frac{1}{\mathrm{N}} \sum_{\mathrm{n}=1}^{\mathrm{N}} \frac{\mathrm{FN}}{\mathrm{TP+TN+FP+FN}} \qquad \text{eq (4-2)}$$

From FNIR, one can calculate another metric called True Positive Identification Rate (TPIR); representing the overall identification performance of the proposed biometric:

TPIR=1 - FNIR

Additionally, system performance over four additional metrics can also be calculated:

Precision, Recall, Sensitivity and Specificity.

The Precision metric represents the ratio of correctly identified positive samples with total number of identification in the experiment:

$$\text{Precision} = \frac{1}{N}\sum_{n=1}^{N}\frac{TP}{TP+FP} \qquad \text{eq (4-3)}$$

The Recall metric represents the ratio of correctly identified positive samples with total number of positive samples in the experiment:

$$\text{Recall} = \frac{1}{N}\sum_{n=1}^{N}\frac{TP}{TP+FN} \qquad \text{eq (4-4)}$$

Specificity represents the ratio of correctly identified rejected negative samples with total number of negative samples

$$\text{Specificity} = \frac{1}{N}\sum_{n=1}^{N}\frac{TN}{TN+FP} \qquad \text{eq (4-5)}$$

Sensitivity represents the ratio of correctly identified positive and negative samples with total number of positive and negative samples in the experiment:

$$\text{Sensitivity} = \frac{1}{N}\sum_{n=1}^{N}\frac{TP+TN}{TP+TN+FP+FN} \qquad \text{eq (4-6)}$$

**Note:**

In this part, only European ST-T Database is used due to the fact that this database matches better with our algorithm. Also, 20 recordings considered as intruders ECGs that were not considered in the training part are now added in order to test the response of the system to intrusion.

**a) Effect of the Threshold coefficient:**

The metrics listed above are now calculated for different threshold values. This entails the multiplication of the threshold values by coefficient numbers varying between 0.1 and 10. The process is repeated for three successive parts of the signal, with the length being 5 PQRST frames obtained from our ECG testing signal. It is expected that our system will not recognize any ECG in the case of coeff=0.1, and will accept all ECGs, including intruders, when coeff=10.

Table 4-2 summarizes the overall system performance in the terms mentioned above.

*Reminder:* Our Thresholds were calculated by selecting the maximum of the distances between different ECG parts of the same person for each threshold.

| **Coefficients** | **0.1** | **0.3** | **0.5** | **1** | **2** | **3** | **5** | **10** |
|---|---|---|---|---|---|---|---|---|
| **FPIR** | 0 | 0.0417 | 0.075 | 0.1667 | 0.3250 | 0.4 | 0.4417 | 0.5 |
| **FNIR** | 0.475 | 0.3167 | 0.1917 | 0.025 | 0 | 0 | 0 | 0 |
| **TPIR** | 0.525 | 0.6833 | 0.8083 | 0.975 | 1 | 1 | 1 | 1 |
| **Precision** | 0.8213 | 0.8071 | 0.7986 | 0.7402 | 0.6076 | 0.5558 | 0.5313 | 0.5 |
| **Recall** | 0.05 | 0.3667 | 0.6167 | 0.95 | 1 | 1 | 1 | 1 |
| **Specificity** | 1 | 0.9167 | 0.85 | 0.6667 | 0.35 | 0.2 | 0.1167 | 0 |
| **Sensitivity** | 0.525 | 0.6417 | 0.7333 | 0.8083 | 0.6750 | 0.6 | 0.5583 | 0.5 |

**Table 4-2:***Metrics versus threshold multiplication coefficient.*

According to the definitions of each metric, FPIR, FNIR are negative parameters that need to be reduced as much as possible; while all the other parameters are positive parameters that need to be increased as high as possible, and doing so without affecting the system.

The selection of any appropriate threshold coefficient needs to be justified. To do so, an algorithm has been coded, plotting the summation of each column (while taking inverse probabilities for both FPIR and FNIR), determining the location corresponding to the highest value as that of the column of the appropriate case. Result values are shown as follows:

```
best_matching =

    4.5250    5.0571    5.5402    5.9485    5.3076    4.9558    4.7646    4.5000
```

Note the location of the highest value (5.9485) listed in the fourth column. This means that the Threshold coefficient'**1**' is of the most appropriate case.

These results indicate that our proposed biometric system can effectively identify the subjects with **a high overall identification performance of 0.975,** and on the other hand, **a False Positive and Negative Identification Rates** of**0.1667**and **0.025,** respectively**.**

**b) Effects of number of PQRST frames taken for testing:**

In order to achieve system enhancement, previous results are used to calculate the system's parameters for different sizes of recordings that are used for testing.

Firstly, the threshold coefficient is set to unity, and metrics are calculated for different lengths of our testing ECG signal, ranging from 3 to 9.

| **Frames** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|---|
| **FPIR** | 0.175 | 0.1667 | 0.1375 | 0.1250 | 0.1333 | 0.1417 | 0.1500 |
| **FNIR** | 0.0813 | 0.0417 | 0.0250 | 0.0250 | 0.0083 | 0.0083 | 0.0083 |
| **TPIR** | 0.9187 | 0.9583 | 0.9750 | 0.9750 | 0.9917 | 0.9917 | 0.9917 |
| **Precision** | 0.7022 | 0.7326 | 0.7767 | 0.7913 | 0.7867 | 0.7769 | 0.7662 |
| **Recall** | 0.8375 | 0.9167 | 0.9500 | 0.9500 | 0.9833 | 0.9833 | 0.9833 |
| **Specificity** | 0.65 | 0.6667 | 0.7259 | 0.7500 | 0.7333 | 0.7167 | 0.7000 |
| **Sensitivity** | 0.7438 | 0.7917 | 0.8375 | 0.8500 | 0.8583 | 0.8500 | 0.8417 |

**Table 4-3:** *Metrics versus number of samples.*

Repeating the same procedure for this experiment provides the values listed below after summations:

```
best_matching =

    5.5959    5.8576    6.1026    6.1663    6.2117    6.1686    6.1246

>>
```

Results show that the most appropriate case would be to take 7 PQRST frames for testing, while exhibiting significantly small differences when taking more than 7 PQRST frames.

From these results, it can be concluded that the proposed biometric system can effectively identify the subjects with **a high overall identification performance of 0.9917,** and on the other hand, **a False Positive and Negative Identification Rates** of **0.1333**and **0.0083,** respectively**.** Comparison with previous results demonstrates a significant improvement in the performance of our system.

In regard to 7, 8 and 9 frames, one observes that the **high overall identification performance** remains constant, while **False Positive Identification Rate value** increases, corresponding to the percentage of 'incorrectly identified' events. This result indicates that an increase in the number of frames tends to also increase the intrusion risk of the system, while not affecting system performance in regard to **subjects whose recordings are stored in the database** .

To obtain the best result in the case scenario of intrusive presence within the system; an acceptance threshold must be chosen as the highest distance between one's different parts of his Electrocardiogram, and at various periods of time, in order to prevent the intruder from being recognized as subject, with the size of the signal (number of PQRST frames taken for testing) not below 7 frames in order to obtain the best overall performance.

## 4.2 EEG based biometric system:

The same procedure applied to ECG biometric system will be conducted for the EEG system; thus, starting with a demonstration of the proposed system by testing one local data and showing all the various steps.

### 4.2.1 Testing:

The EEG signal data of the person being identified is again loaded. Note that this current data is different from the one already stored and used in the training part, which was also recorded at different periods of time. Below is the MATLAB code that performed this command:

```
load('S007R07_edfm.mat');
eeg_channel = val;
```

Consequently, the same procedures are performed, meaning that the data is filtered and features are extracted for each channel, ending up with a feature vector.

The absolute distance is calculated from the obtained features vector and the stored data for the 20 persons, while the smallest distance is used for identifying the person.

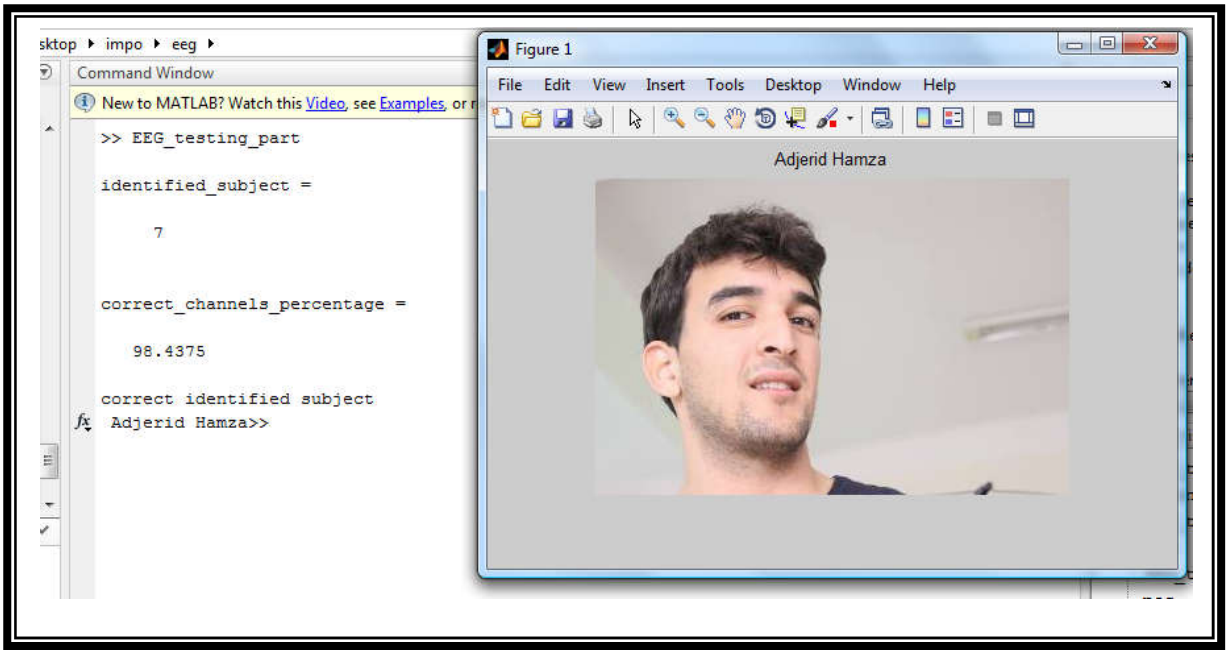



**Figure 4-4:** *The result of running the program when loading (*'S007R07_edfm.mat'*) which is the data of 'Hamza'.*

### 4.2.2 Performance evaluation:

In order to test the performance of our system, our algorithm is evaluated using all the N data, by analyzing it for various variable parameters.

The results are displayed for two variables, statistical parameters and number of channels. For each of the three given cases, results are produced for a variation of the number of channels between 1 and 64, taking only the mean for the first case; mean and standard deviation for the second case, and mean, standard deviation and entropy for the third case.

| Number of used channels | 1 | 5 | 10 | 20 | 30 | 40 | 50 | 64 |
|---|---|---|---|---|---|---|---|---|
| Accuracy (%) | 0 | 0 | 10 | 20 | 20 | 20 | 20 | 20 |

**Table 4-4:** *Accuracy versus number of used channels when taking only the mean.*

| Number of used channels | 1 | 5 | 10 | 20 | 30 | 40 | 50 | 64 |
|---|---|---|---|---|---|---|---|---|
| Accuracy (%) | 5 | 25 | 45 | 90 | 90 | 90 | 90 | 90 |

**Table 4-5:** *Accuracy versus number of used channels when taking the mean and the standard deviation.*

| Number of used channels | 1 | 5 | 10 | 20 | 30 | 40 | 50 | 64 |
|---|---|---|---|---|---|---|---|---|
| Accuracy (%) | 45 | 70 | 85 | 90 | 90 | 95 | 100 | 100 |

**Table 4-6:** *Accuracy versus number of used channels when taking the mean, the standard deviation and the entropy*

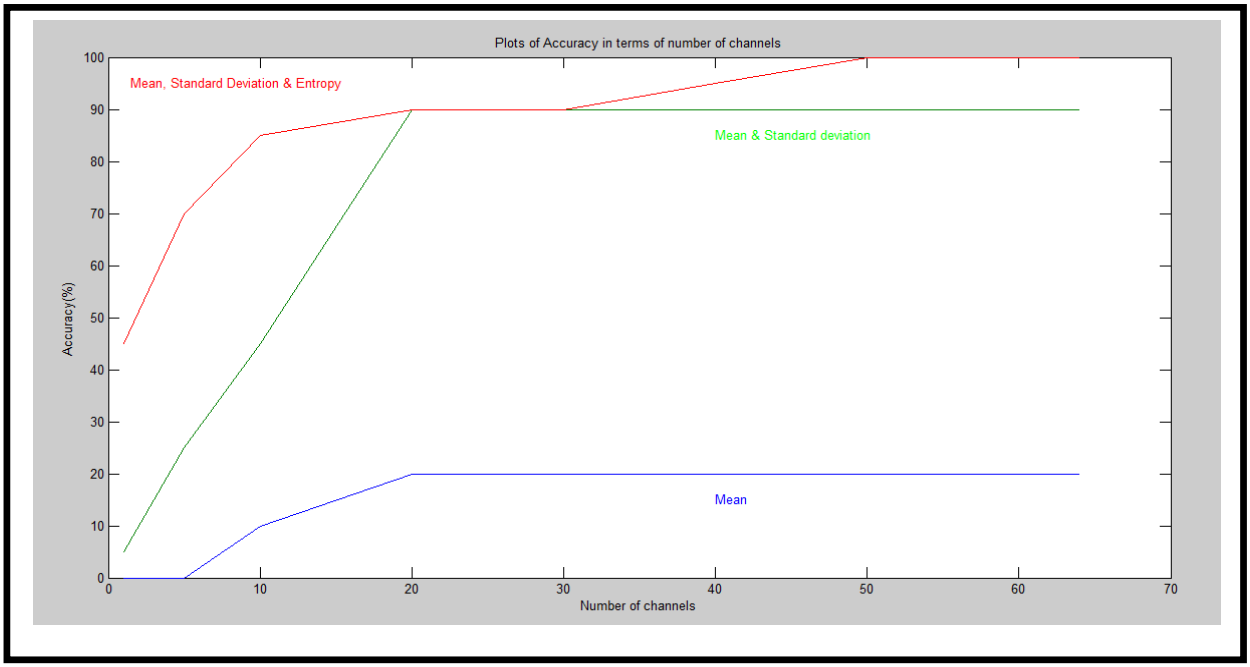



**Figure 4-5:** *Plot of accuracy for each three cases in terms of Number of channels.*

Notice that when using only the mean as parameter, the accuracy percentage ranges from 0% to only20% when increasing the number of channels used, so this scenario won't present enough efficiency and performance for an identification system.

When adding standard deviation, notice a significant change when using more than 10 channels. Accuracy reached 90%, presenting an acceptable performance for an identification system. When the entropy parameter is added, and combined with the two other parameters, the accuracy percentage for the case of 10 channels or moreexceeds85%, reaching100% when using 64 channels. In conclusion, the system proposed is 100%performantwhen using the combination of the mean, the standard deviation and the entropy for a 64 channel system.

To answer the question asked in Section 3.1.4 of Chapter 3, since we sought the effect of the use of multi-channels in biometric systems, and after the realization and analysis of the performance of such system, it can be concluded that it is highly important to utilize multi-channels when dealing both with biometric systems as with other EEG applications.

## 4.3 Effect of Diseases on ECG-based identification:

*"How does ECG based biometrics account for natural variation due to individual rhythm or pathology? Does the device need to be recalibrated every so often?*

This question was asked in a paper published by the American College of Cardiology under the name "The Heart Knows-ECG-Based Biometrics / Cardio Source World News" on January 27, 2014.

Motivating by the fact that our developed ECG based biometrics algorithm is limited to work under the condition of healthy subjects only, meaning that our proposed system may not be efficient while dealing with ECG abnormalities. To answer the previous question; the following study is performed.

For the purpose of knowing the effect of heart diseases on our identification system, Atrial Fibrillations, considered to be one of the most popular heart diseases, was selected for our study.

### 4.3.1 Atrial Fibrillation:

Atrial fibrillations is a heart condition involving irregular heart rhythm, known as an arrhythmia. It is the most common type of arrhythmia affecting people with heart diseases. Generally, the risk of developing Atrial fibrillation increases with age and with other risk factors such as diabetes, high blood pressure and underlying heart disease.

-Some people affected by Atrial fibrillation may feel perfectly fine. They may not even know they have the condition until they take a routine test called Electrocardiogram.

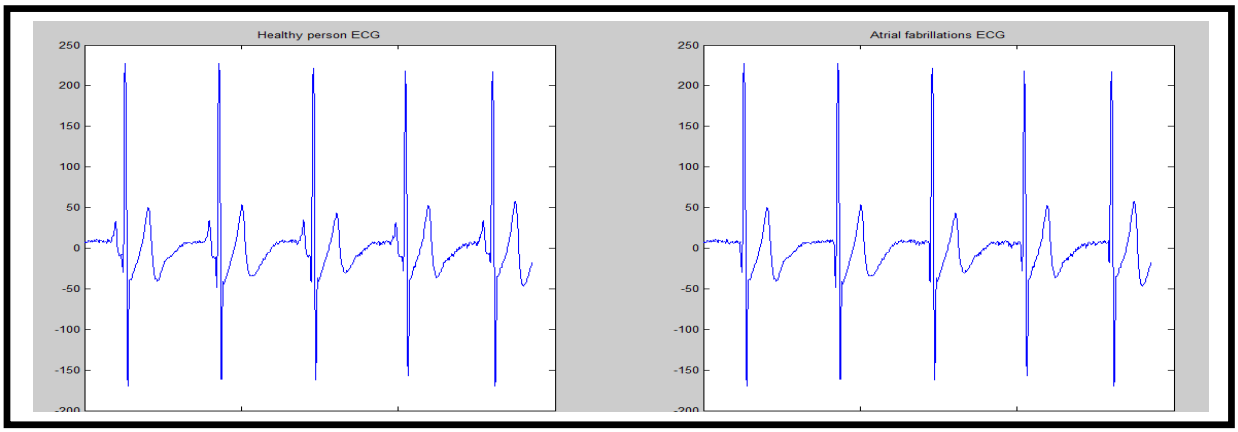
Electrocardiograms of people suffering from Atrial fibrillation is characterized by [21] :

- Irregular and fast heartbeat.
- Absence of normal P waves.
- Normal QRS complexes.

An algorithm has been developed in our study transforming the ECG of a healthy person by making it appear like an ECG of someone suffering of Atrial fibrillation.

## 4.3.2 Practical part:

The main work of the algorithm is to omit the P wave, which comes just before the QRS complex. To do this, the algorithm was coded by taking into account the ECG signal selected for testing, the algorithm localizing all the P waves in the signal making them look like any other parts of the ECG in a period of rest (any part of ECG other than P, QRS, and T waves).




(a) (b)

**Figure 4-6:** *ECG of the data ('e01.mat'); (a) ECG of healthy person; (b) ECG after simulation (Atrial fibrillations ECG).*

In Figure 4-6, notice the absence of the P wave .Now our simulated program is saved in our database under the name ('AF_1.mat')

```
%%% save p_removed signal


savefile='AF_1.mat';
save(savefile,'wave_s_p');
```

Following the same procedure, all the Electrocardiograms of our Databases are simulated and saved under succeeding names.

### 4.3.3 Testing healthiness:

In order to test the healthiness of our persons, an algorithm was coded for that purpose with the main task of looking for the wave that is supposed to be P wave, checking its amplitude and taking appropriate decision.

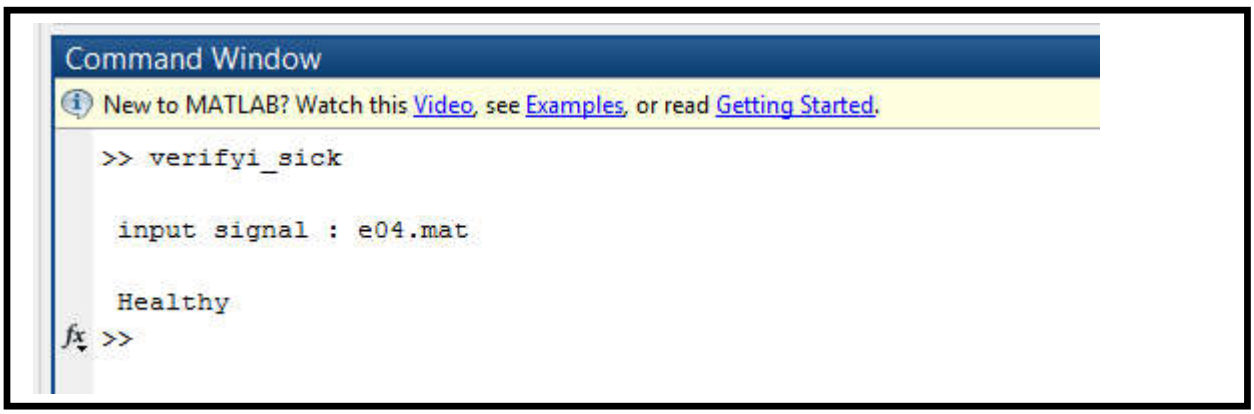
The algorithm locates the peak situated right before the QRS complex (R wave), a P wave being easily recognized because of its recognizable amplitude in the signal. The decision is taken as follows:

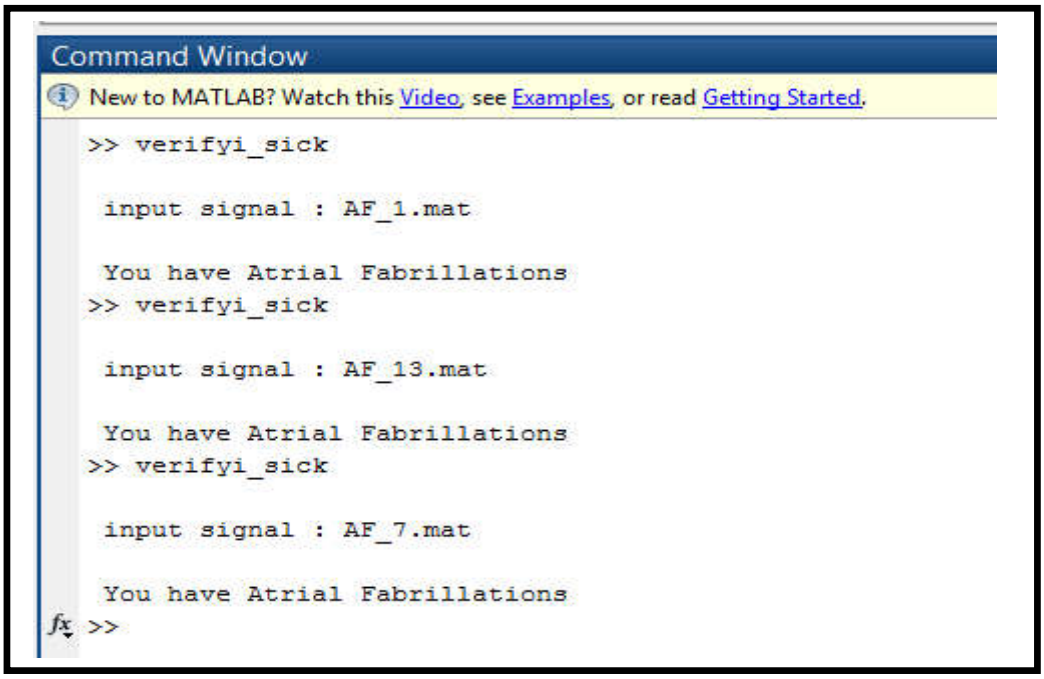
```
if p_value>10

fprintf('\n Healthy \n');
else
fprintf('\n You have Atrial Fabrillations \n');
end
```

### 4.3.4 Implementation:

In the initial step, the signal of a healthy person is input, while in the second step, the signal after simulation (saved under the name 'wave_s_p') is introduced. We are thus taking one healthy person's ECG from our database, in addition to three randomly simulated signals, as shown in the display the command window below:




Further, the performance of our program for identification using our simulated ECG signals for 20 people is therefore tested by loading the simulated signals one after the other.

No one of the persons was recognized.


ECG based biometrics can tolerate some natural changes as explained in Chapter 2; Such results were obtained by fixing a threshold for each subject, however, ECG based biometrics are highly sensitive to natural variations in individual pathology **when caused by diseases**. Our study results have determined that Identification is impossible in cases of Atrial Fibrillations disease. The ECG device would need frequent recalibrations in order to update the database in the event of subjects recently starting to suffer from Heart diseases; or, alternatively, the development of a stronger and more reliable Biometric identification system by using more advanced identification features. For this reason, we have proposed the EEG-based Biometrics option in our study.

**Summary:**

In this chapter, and experimental testing of our ECG and EEG based models was conducted , in addition to performance evaluation and statistical studies using algorithms developed for this research.

Firstly, performances of our ECG-based identification model for MIT-BIH and European ST-T databases were evaluated using PCA and AC/DCT method .

In a second experiment, the performance of our ECG-based identification model was evaluated for two variables , namely Threshold coefficients, and the number of PQRST frames. The evaluation was conducted with seven types of parameters, resulting with an ideal case suitable for our identification system.

In a third experiment, our proposed EEG-based biometric system was tested , then its performance was evaluated while adding additional parameters in each time.

In the fourth and last experiment, the effect of a heart disease (i.e., Atrial fibrillation) on the proposed ECG-based identification system was studied. A program was developed to simulate a healthy person's Electrocardiogram into the Electrocardiogram of a person sick with Atrial fibrillation. Finally, the performance of the proposed identification system in the case of a person affected with Atrial fibrillation was then tested..

## Conclusion:

Biometric technologies are based on the ways in which humans can be uniquely identified through various biological traits, which include fingerprints, eye retina patterns, voice waves, DNA and signatures. Biometric technologies are used to secure a variety of securitized areas, including enterprise security, online commerce, airport security, and banking.Biometricsystems are becoming increasingly critical and important in many aspects of today's world, as they are shown to offer the most robust and reliable tool of identification and verification techniques.

Our proposed ECG-based biometric system consists of extracting the necessary features using the AC/DCT method, as we verify the accuracy of our identification system against two databases, the MIT-BIH and the European ST-T database using both the AC/DCT method and the PCA method that was previously performed in an experiment done in a related work , for which we made a comparison of the accuracy of our identification system .We found that AC/DCT performs better when dealing with the European ST-T database due to the fact that most of them were recordings of healthy persons, whereas the PCA method provides better results with the MIT-BIH database . We also noticed that increasing the number of PQRST frames provided a higher accuracy overall.

A critical question was how our proposed system would respond to intrusion,To address this question, we added additional recordings that were not considered in the training part of the study, and we then evaluated the performance of our system using seven different parameters, witha variation applied to the acceptance threshold and the number of PQRST frames.Our results indicated that the ideal case scenario coincided with the selection of the acceptance Threshold as the highest distance between different parts of someone's ECG , and the number of PQRST frames being fixed at 7, at avoid false acceptance. This result is specific to the case of intruder presence only.

For our proposed EEG-based biometric system, a performance evaluation was completed by computing its accuracy while increasing the number of channels and adding more statistical parameters for each step. It was found that the ideal case is when 64 channels are selected, in combination with the three proposed statistical parameters (mean, standard deviation, entropy) .

This research also addressed and analyzed the manner in which our ECG-based biometrics models account for natural variations in individual rhythm or pathology for those suffering from myocardial issues. The analysis of our identification system using recordings of persons suffering of Atrial fibrillation diseases shown that identification is impossible in the case of the presence of subjects with heart diseases .

## Appendix :

- **Autocorrelation function:**

Autocorrelation is the cross-correlation of a signal with itself at different points in time (that is what the cross stands for). Informally, it is the similarity between observations as a function of time lag between them.

It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals.

**Syntax:**

Autocorrelation(x, Lags)

Where:

- x: discrete univariate real time series given as a Vector ,List , Matrix with one column, or Time Series object with one dataset.
- Lags: maximal lag to return, or range of lags to return. By default, all possible lags are returned, so it is an optional option while programming.

For a discrete time series x, the Autocorrelation command computes the autocorrelations $Rk = \frac{Ck}{C0}$ ; where:

$$Ck = \sum_{t=1}^{n-k}(X_t \quad \mu)(X_{t+k} \quad \mu) \qquad \text{(for k=0....n-1) and } \mu \text{ being the mean of x.}$$

- **Discrete Cosine Transform:**

Like many other transforms, the Discrete Cosine Transform DCT attempts to de-correlate the signal or image data. After de-correlation, each transform coefficient can be encoded independently without losing compression efficiency. This section describes the DCT and some of its important properties. In this project, DCT is employed to transform vectors only, leading to the enhancement of only One-Dimensional DCT. The most common DCT definition of a 1-D sequence of length N is: y=dct(x).

$$y(k) = w(k) \sum_{n=1}^{N} x(n).\cos\left(\frac{\pi}{2N}(2n \quad 1)\,(k \quad 1)\right), \qquad k = 1,2, \ldots \ldots, N$$

Where:

$$w(k) = \frac{1}{\sqrt{N}} \quad , \quad k = 1.$$

$$w(k) = \sqrt{\frac{2}{N}} \quad , \quad 2 \leq k \leq N.$$

Some properties of DCT are:

De-correlation: The principle advantage of Vector transforming is the removal of redundancy between neighboring -pixels- . It leads to uncorrelated transform coefficients which can be encoded independently.

Energy compaction: The efficacy of a transformation scheme can be directly gauged by its ability to pack input data into as few coefficients as possible. This allows the quantizer to discard coefficients with relatively small amplitudes without introducing visual distortion on the reconstructed vector. DCT exhibits excellent energy compaction for highly correlated vectors.

## References :


[1] http://searchsecurity.techtarget.com/definition/biometrics (14/06/2016)

[2]https://www.google.dz/search?q=biometric+characteristics&rlz=1C1GGGE_frDZ620DZ620&espv=2&biw=1366&bih=667&tbm=isch&imgil=I0fQXzYVljxSMM%253A%253BlvJz8aPYdesyMM%253Bhttp%25253A%25252F%25252Felectronicmaffia.weebly.com%25252Fbiometrics.html&source=iu&pf=m&fir=I0fQXzYVljxSMM%253A%252ClvJz8aPYdesyMM%252C_&usg=__449MUqLnf-GY0ZJXw4wMbtbF0PY%3D&ved=0ahUKEwiGs5HXiKjNAhXENhoKHelzDkIQyjcIPA&ei=P0JgV8aXOsTtaOnnuZAE#imgrc=I0fQXzYVljxSMM%3A (14/06/2016)

[3]https://www.google.dz/search?q=eeg&rlz=1C1GGGE_frDZ620DZ620&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiIrY6xi6jNAhVCLB4KHc9lA6cQ_AUICCgB&biw=1366&bih=667#tbm=isch&q=eeg+biometric&imgrc=3zwo5rX9IztgHM%3A (14/06/2016)

[4] https://www.google.dz/search?q=pulse_biometrics-1&rlz=1C1GGGE_frDZ620DZ620&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiI84qDi6jNAhUJJh4KHRZ5DzIQ_AUICSgC&biw=1366&bih=667#imgrc=GWIsejggoseu_M%3A (14/06/2016)

[5] http://www.nottingham.ac.uk/nursing/practice/resources/cardiology/function/anatomy.php (14/06/2016)

[6] http://www.physionet.org/pn3/ecgiddb/biometric.shtml (14/06/2016)

[7] https://www.physionet.org/physiobank/database/edb/ (14/06/2016)

[8] https://www.physionet.org/physiobank/database/mitdb/ (14/06/2016)

[9] //en.wikipedia.org/wiki/Human_brain/ (14/06/2016)

[10] https://en.wikipedia.org/wiki/Brain (14/06/2016)

[11]https://www.google.dz/search?q=brain_regions&rlz=1C1GGGE_frDZ620DZ620&espv=2&biw=1366&bih=623&source=lnms&tbm=isch&sa=X&ved=0ahUKEwi1-MzC5uzMAhUDOxoKHSnGAbYQ_AUIBigB#imgrc=BG0JXW0CMetQ2M%3A (14/06/2016)

[12] http://medical-dictionary.thefreedictionary.com/EEG (14/06/2016)

[13] http://neurosky.com/2015/07/multi-channel-eeg-bci-devices/ (14/06/2016)

[14] https://www.physionet.org/physiobank/database/eegmmidb/ (14/06/2016)

[15] https://en.wikipedia.org/wiki/Wavelet_packet_decomposition (14/06/2016)

[16] https://www.mathsisfun.com/mean.html (14/06/2016)

[17] https://en.wikipedia.org/wiki/Standard deviation (14/06/2016)

[18] Qiong Gui, Zhanpeng Jin and Wenyao Xu. Exploring EEG-based biometrics for user identification and authentification.

[19] https://en.wikipedia.org/wiki/Nearest_neighbour_classifiers (14/06/2016)

[20] Haque, Mohammad Ahsanul; Nasrollahi, Kamal; Moeslund, Thomas B. " Heartbeat Signal from Facial Video for Biometric Recognition", Aalborg Universitet, *Published in* 19th Scandinavian Conference on Image Analysis (SCIA) *(part 3)*

[21] http://www.mayoclinic.org/diseases-conditions/atrial-fibrillation/home/ovc-20164923 (14/06/2016)

[22] Syed Ali Khayam "The Discrete Cosine Transform (DCT): Theory and Application1", Department of Electrical & Computer Engineering, Michigan State University, *March 10th 2003 (page 4-14)*

[23] YongjinWang, Foteini Agrafioti, Dimitrios Hatzinakos, and Konstantinos N. Plataniotis, Analysis of Human Electrocardiogram for Biometric Recognition, *Research Article, The Edward S. Rogers Sr., Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4*, Correspondence should be addressed to YongjinWang, ywang@comm.utoronto.ca, Received 3 May 2007; Accepted 30 August 2007 (*page 1-9*)

[24] Foteini Agrafioti, Jiexin Gao and Dimitrios Hatzinakos, Heart Biometrics: Theory, Methods and Applications, *The Edward S. Rogers Sr. Department of Electrical and Computer Engineering University of Toronto 10 Kings College Road, Toronto*, *Canada (page 200-209)*

[25] The Heart Knows—ECGBased Biometrics | CardioSource WorldNews, AmericanCollege of Cardiology, Jan 27, 2014

[26] https://en.wikipedia.org/wiki/Sensitivity_and_specificity (14/06/2016)

[27] http://www.techexams.net/forums/security/49256-true-positive-v-s-true-negative.html (14/06/2016)

[28] http://msp.ucsd.edu/techniques/v0.11/book-html/node147.html (14/06/2016)

[29]http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.5052135/k.2C86/Heart_disease__Atrial_fibrillation.htm (14/06/2016)

[30] https://www.nhlbi.nih.gov/health/health-topics/topics/af (14/06/2016)

[31] http://link.springer.com/referenceworkentry/10.1007/978-3-642-27733-7_9145-2#page-1 (14/06/2016)

[32]http://www.maplesoft.com/support/help/Maple/view.aspx?path=Statistics%2FAutoCorrelation (14/06/2016)

[33] http://www.mathworks.com/help/signal/ref/dct.html (14/06/2016)

[34] http://www.intechopen.com/books/biomedical-engineering/eeg-based-personal-identification (14/06/2016)

[35]http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=7019605&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel7%2F7005967%2F7019573%2F07019605.pdf%3Farnumber%3D7019605 (14/06/2016)

[36] http://www.biometricsinstitute.org/pages/types-of-biometrics.html (14/06/2016)

[37] http://www.ncbi.nlm.nih.gov/pubmed/25836061 (14/06/2016)

[38] Implementation of a Biometric identification system using Electrocardiogram ECG/EKG, Badreddine Zebbiche , IGEE Institute .