CEREN DURSUN

M.S. Thesis

2017

BIOMETRIC RECOGNITION USING BIO-SIGNALS

CEREN DURSUN

IŞIK UNIVERSITY

2017

BIOMETRIC RECOGNITION USING BIO-SIGNALS

CEREN DURSUN
B.S., Electronics Engineering, IŞIK UNIVERSITY, 2014

Submitted to the Graduate School of Science and Engineering
in partial fulfillment of the requirements for the degree of
Master of Science
in
Electronics Engineering

IŞIK UNIVERSITY
2017

IŞIK UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

BIOMETRIC RECOGNITION USING BIO-SIGNALS

CEREN DURSUN

APPROVED BY:

Assoc. Prof. Hakan Gürkan Işık University ____________
Thesis Supervisor(Thesis Co-Advisor: Assoc. Prof. Ümit Güz)

Assist. Prof. Sakip Önder Işık University ____________

Prof. Dr. Serdar Özoğuz Istanbul Technical University ____________

APPROVAL DATE: 14/04/2017

BIOMETRIC RECOGNITION USING BIO-SIGNALS

Abstract

The main objective of the project is to increase the recognition rate by establishing a multimodal biometric recognition system that uses two different biometric characteristics, as bio-signals.

Today, institutions use biometric recognition systems quite often to provide security for many areas such as information security and physical security. The importance of these systems increases day by day in the direction of technological development and increasing demand. Recognition systems based on biometric characteristics are more reliable, because of the possibility of forgetting or losing knowledge in the recognition systems based on knowledge (eg: password) and the possibility of being stolen or guessed by third persons in the biometric recognition based on possessed (eg: card). However, the fraud techniques are also developing in the direction of technological developments and biometric characteristics cannot be renewed in case of imitation, hence the use of multiple biometrics recognition system may be a solution to this problem. At the same time, the use of multiple biometrics increases in the security of systems.

In this thesis, a biometric recognition system, which uses the Electrocardiogram (ECG) and speech signals of the person, was created. Since there was not enough time and possibility, an artificial database was generated with obtaining these signals from various sources. First, the MIT-BIH Arrhythmia Database was used for ECG signals. This database consists of 48 ECG signals, which belong to 22 females and 26 males. In accordance with this database, a database was created for the speech signals, which were obtained from the website, given in [1].

The features of the biometric signals were extracted by AC / DCT (Autocorrelation/Discrete Cosine Transform) method for ECG signals and by Mel Frequency Cepstrum Coefficients (MFCCs) method for speech signals. The data, which were obtained from the feature extraction, were then classified by the Gaussian Mixture Model (GMM) method. The scores, which were obtained from the classification process, were fused as a single individual's data, and the decision-making step was passed. Recognition rates were obtained in the decision making step.

The recognition rate for the ECG signal was %87.50 and 42 persons were matched correctly. The recognition rate for 2 seconds speech signals was %58.33 and 28 persons were matched correctly.

Normalization was applied before the fusion of these two datasets. The recognition rate after the fusion was %70.83 and 34 persons were matched correctly. However, when the recognition rates are considered, it has been observed that the recognition rate, which obtained after the fusion, is lower than recognition rate of the ECG signals. Therefore, instead of 2 seconds speech signals, 10 seconds speech signals were used. In this case, the recognition rate of the speech signals was %97.9 and 47 persons were matched correctly. Then, normalization was applied again and two datasets were fused. After the fusion, the rate of recognition reached %95.8 and 46 persons were matched correctly.

Keywords: GMM, MFCC, AC/DCT, ECG signals, Speech signals, Multimodal Biometric Recognition System

BİO İŞARETLER KULLANILARAK BİYOMETRİK TANIMA

Özet

Projenin temel amacı kişinin iki farklı biyometrik karakteristiğini kullanan bir çoklu biyometrik tanıma sistemi kurarak tanıma oranını yükseltmektir.

Günümüzde kurumlar bilgi güvenliği ve fiziksel güvenlik gibi birçok alanda güvenliği sağlamak için biyometrik tanıma sistemlerini oldukça sık kullanmaktadır. Teknolojinin gelişimi ve artan talep doğrultusunda da bu sistemlerin önemi gün geçtikçe artmaktadır. Kişinin bildiği (ör: şifre) veya sahip olduğu (ör: kart) bilgiye veya nesnenin varlığına dayanan tanıma sistemlerinde, kişinin bunları unutabilme, kaybedebilme veya üçüncü kişiler tarafından çalınabilme veya tahmin edilebilme olasılığından dolayı biyometrik karakteristiklere dayanan tanıma sistemleri daha güvenilirdir. Fakat teknolojik gelişmeler doğrultusunda dolandırıcılık teknikleri de gelişmektedir, dolayısıyla biyometrik karakteristikler taklit edilmesi durumunda yenilenemeyeceği için tanıma sistemlerinde birden fazla biyometriğin kullanılması bu sorunun çözümü olabilir. Aynı zamanda birden fazla biyometriğin kullanılması sistemlerin güvenliğini daha artırmaktadır.

Bu tezde kişinin Elektrokardiyogram (EKG) ve ses işaretlerini kullanan biyometrik tanıma sistemi oluşturulmuştur. Yeterli zaman ve imkan olmadığından dolayı bu işaretler çeşitli kaynaklardan elde edilerek yapay bir veri tabanı oluşturulmuştur. Öncelikle, EKG işaretleri için MIT-BIH Arrhythmia Database'i kullanılmıştır. Bu veri tabanı 22 kadın ve 26 erkek olmak üzere toplamda 48 kişinin EKG işaretlerinden oluşmaktadır. Bu veri tabanına uygun olarak ses işaretleri için internet sitesinden [1] elde edilen ses kayıtları ile bir veri tabanı oluşturulmuştur. Bu tezde kullanılan biyometrik işaretlerin öznitelikleri ise EKG işaretleri için ÖK/AKD (Özilinti Katsayıları/Ayrık Kosinüs Dönüşümü) yöntemi ve ses işaretleri için Mel Frekans Kepstral Katsayları (MFKK) kullanılarak çıkarılmıştır. Öznitelik çıkarımından elde edilen veriler daha sonra Gaussian Karışım Modeli (GKM) yöntemi ile sınıflandırılmıştır. Sınıflandırılma işleminden elde edilen veriler tek bir bireyin verileriymiş gibi birleştirilerek karar verme adımına geçilmiştir. Karar verme adımında tanıma oranları elde edilmiştir. EKG işareti için tanıma oranı %87.50 olup, 42 kişi doğru eşleşmiştir. 2 saniyelik ses işaretleri için tanıma oranı

%58.33 olup, 28 kişi doğru eşleşmiştir. Bu iki veri seti birleştirilmeden önce normalizasyon yapılmıştır. Birleştirme sonrası elde edilen tanıma oranı %70.83 olup, 34 kişi doğru eşleşmiştir. Fakat tanıma oranlarına bakıldığında birleşmeden sonra elde edilen tanıma oranının EKG işareti için elde edilen orana göre düşük olduğu gözlemlenmiştir. Bundan dolayı 2 saniyelik ses işaretleri yerine 10 saniyelik ses işaretleri kullanılmıştır. Bu durumda ses işaretlerinin tanıma oranı %97.9 olup, 47 kişi doğru eşleşmiştir. Daha sonra tekrar normalizasyon yapılıp iki veri seti birleştirilmiştir. Birleşme sonrası elde edilen tanıma oranı ise %95.8 ulaşmış ve 46 kişi doğru eşleşmiştir.

Anahtar kelimeler: GKM, MFKK, ÖK/AKD, EKG işaretleri, Ses işaretleri, Çoklu Biyometrik Tanıma Sistemi

Acknowledgements

The research work was supported by Coordination Office for Scientific Research Project, ISIK University, Project Number: 14A203, Project Title: Fingertip ECG Signal Based Biometric Recognition System.

Firstly, I warmly thank Assoc. Prof. Hakan Gürkan, who is my project supervisor, and Assoc. Prof Ümit Güz, who is my co-advisor, to always support me with their intimate knowledge in the all phases of this project. Thanks to them, i never feel hopelessness about the completion of this project.

Also, I want to thank my special friends, who always make me feel never alone with their real friendship, help and patience during the academic years.

Finally, I thank my parents Mustafa Dursun and Halime Dursun, who always support me with their love and patient.

To My Family...

Table of Contents

List of Tables

List of Figures

List of Abbreviations

MFCC	**M**el **F**requency **C**epstrum **C**oefficient
GMM	**G**aussian **M**ixture **M**odel
AC	**A**uto**c**orrelation
DCT	**D**iscrete **C**osine **T**ransform
ECG	**E**legtro**c**ardio**g**ram

Chapter 1

Introduction

1.1 Biometric Recognition System

Nowadays, using bio signals becomes prevalent for public intuitions, companies etc. to preserve systems from danger and risks. For example, biometric recognition is used generally for ID control, entrance and exit control to buildings, access control to systems, databases etc. Thanks to the biometric recognition system, security and confidentiality of systems are improved, but today the biometric recognition systems, which use bio signals to identify or verify person, are still being developed. Therefore, targets of biometric recognition may be defined as identification and verification. There are 2 main processes, which are named as training and test, for the recognition systems. In the training phase, bio signal is enrolled into the system. In the test phase, the systems work to identify or verify the person. The system cannot identify or verify person in the test phase, if his/her bio signal is not modelled and enrolled in training phase, as mentioned in [2]

When the biometric recognition system works for identification, there is an unknown person to be researched for matching, according to comparison with his/her bio signal and other bio signals in database. In other words, identification is to detection that an unknown bio signal belongs to which certain person in the database. Identification may be classified into 2 sub-classes, which are open set and close set. In the close-set, the unknown bio signal belongs to the one of person in the database. In the open-set, the unknown bio signal may not belong to the one of person in the database.

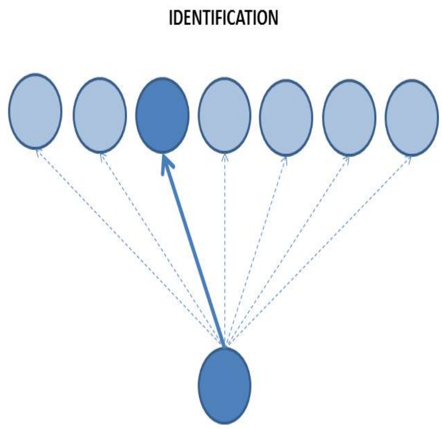

Figure 1.1: Identification Logic

When the biometric system works for verification, there is a person, who is unknown, to be accepted or rejected according to comparison with his/her bio signal and other bio signals in the database. In other word, verification is to determine whether or not that an unknown bio signal belongs to the claimed person.

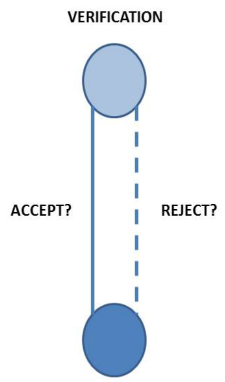

Figure 1.2: Verification Logic

The decision can be made according to the comparison, which may be acception, rejection, exist or not exist based on identification or verification. By help of the decision, the system provides person to access, enter, and exit system, database, building etc. "Biometric recognition can be described as automated methods to accurately recognize individuals based on distinguishing physiological and behavioral traits." [3]. So, the bio signals are classified as the physiological and behavioral characteristics of person. The physiological characteristics are fixed physical characteristic as parts of body like face, finger, fingerprints etc. The behavioral characteristics are based on performance of persons such as speech, gait, signature etc.

The physiological or behavioral characteristics may be assessed with their attributes, which are universality, uniqueness, permanence and collectability. The definitions of attributes, which are clarified below as mentioned in [3].

- *Universality* may be defined as that every person should have this characteristic.

- *Uniqueness* may be defined as that the same characteristic should be different for every person.

- *Permanence* may be defined as that the characteristics should be invariant in certain period of time, position and conditions.

- *Collectability* may be defined as that the characteristic should be measureable.

At the present time, face, fingerprint, hand geometry, iris, voice are preferable than other bio signals, thanks to be more practical. However, all bio signals have vulnerability against frauds, failures. For example, voice may be imitated.

Nevertheless, the bio signals do not require keeping password in mind or carrying cards or tokens to be identified or verified. And also, the bio signals are never forgotten. Hence, using bio signals provides more advantages to people, systems, companies, public institutions. However, if the bio signals are seized by any way, they do not have any chance to be regenerated like cards, tokens, passwords. This problem can be solved by using multimodal biometric recognition systems. The multimodal biometric recognition system uses multiple bio signals to identify and verify person.

1.1.1 History of Biometric Recognition System

First automated biometric recognition system performed as semi-automated for speaker recognition system in the 1940s. In the 1960s, semi and fully automated biometric recognition based on fingerprint, handwriting and face appeared. In the 1970s, first fully automated hand-geometry and fingerprint recognition system were constituted for commercial aim. Larger pilot projects for banking and government applications became popular in the 1980s. By the 1990s, the fully automated biometric recognition systems performed with many different technologies based on including iris and face for both government and commercial applications, as mentioned in [4].

1.2 Multimodal Biometric Recognition System

The Multimodal biometric recognition systems perform with fusion of multiple sources such as multiple biometrics, sensors, units etc. Multimodal systems have more advantages than unimodal and also, the multimodal system may be a solution of unimodal system's problems.

The unimodal biometric recognition system have 5 main problems, because of it is based on single source, such as single biometric, single sensor, single unit etc. These 5 main problems are called as in sensed data, intra-class variations, distinctiveness, non-universality, spoof attacks, as mentioned in [5].

- *Noise in sensed data:* This problem occurs, when the sensors operate to sense biometrics. Information, which are called as sensed data, may be noisy or distorted, which causes to prevent the biometric recognition system to perform succesfully.

- *Intra-class variations:* This problem occurs because of that person interacted incorrectly with the sensor. For example, when the ECG signal is obtained from person, who moved during the sense operation.

- *Distinctiveness:* When the huge amount of persons' biometrics are used in the biometric recognition system, there may be some similarities in the database.

- *Non-universality:* It may be occurred by possibilities of some person, who may not possess some biometrics or details of biometric, which are used to extract the feature.

- *Spoof attacks:* This type of attack is mostly considered against to specially signature or voice.

Multimodal may solve non-universality, because multiple biometrics provide sufficient population coverage. By help of the structure of multimodal systems, which requires that all biometrics is processed simultaneously, attacker cannot spoof the all biometrics simultaneously.

1.2.1 Scenarios of Multimodal Biometric Recognition Systems

Using multiple biometrics in the recognition system is assessed as multimodal biometric recognition systems. However, using multiple sensors, units, snapshots, matching algorithms may be assessed also as multimodal biometric system, as mentioned in [5]. Other biometric recognition systems scenarios:

- *Single biometric and multiple sensors:* Information are gathered by using multiple sensors for single biometric.
- *Single biometric and multiple units:* The clearest example is fingerprint and iris for this scenarios. It may be clarified by that information are gathered from 2 or more fingers or both of irises of a same person.
- *Single biometric and multiple snapshots:* Information are obtained from multiple instance of the same biometric. It may be clarified by good examples, which are multiple samples of the voice, multiple images of the face for this scenario.
- *Single biometric and multiple representations and matching algorithms:* Fusion of results of multiple matching and feature extraction algorithms may be used for the same biometric.

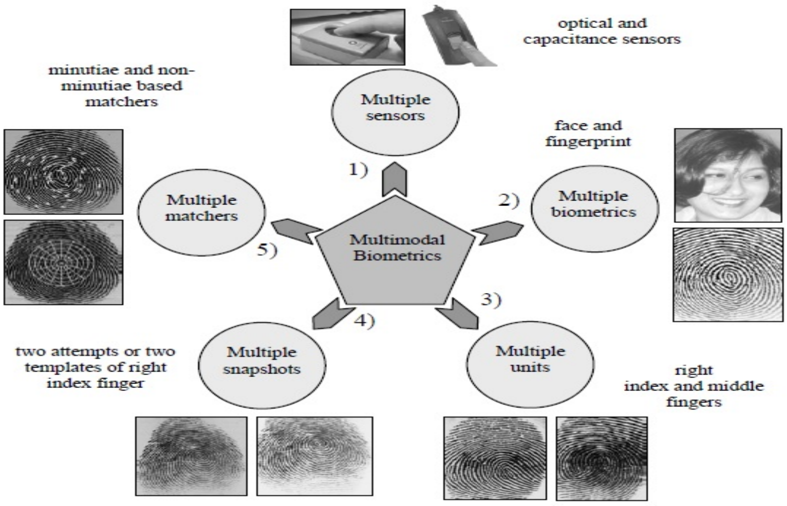

Figure 1.3: Scenarios of Multimodal Recognition System [5]

1.2.2 Modes of Multimodal Biometric Recognition Systems

Multimodal biometric recognition systems may perform in three different modes, which are called serial mode, parallel mode, or hierarchical mode, as mentioned in [6].

- *The parallel mode:* result is obtained from multiple classifiers, which perform at the same time.
- *The serial mode:* in opposition to parallel mode, multiple classifiers perform consecutively to feed the next classifier with their result.
- *The hierarchical mode:* classifiers are combined into a tree scheme.

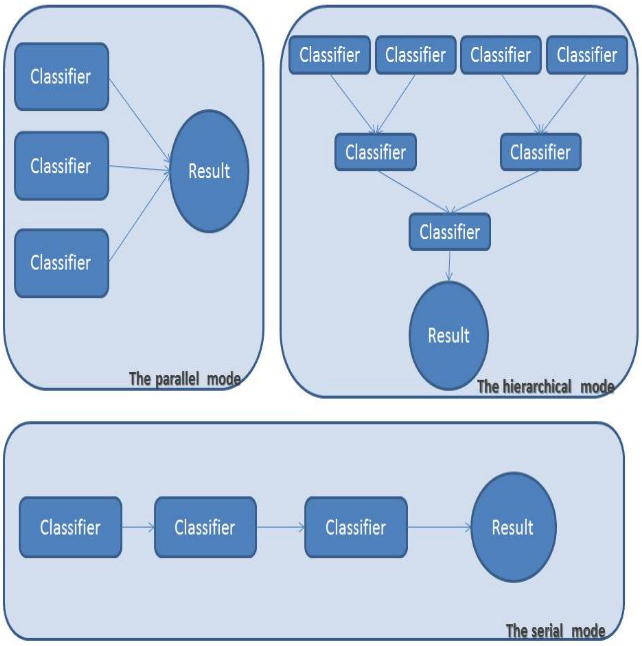

Figure 1.4: Operational modes of multimodal biometric systems.

1.2.3 Fusion for Multimodal Biometric Recognition Systems

Fusion is main section of the multimodal biometric recognition systems. However, fusion operation may be performed at different levels of the multimodal biometric recognition systems. Hence, fusion may be classified into three different groups, as following.

- *Fusion at the Feature Extraction Level*
- *Fusion at the Matching Score Level*
- *Fusion at the Decision Level*

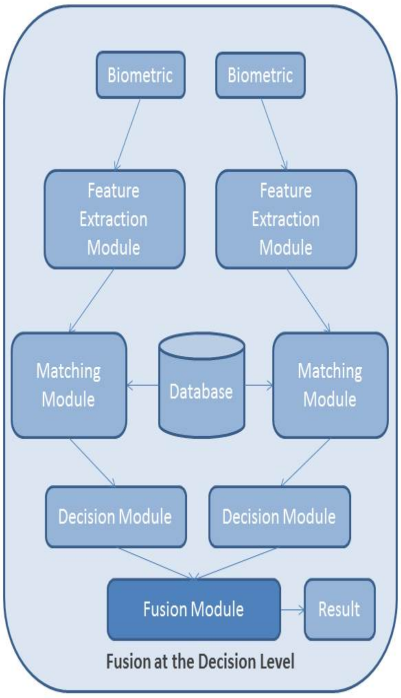

Figure 1.5: Block Diagram of Fusion at the Extraction Level

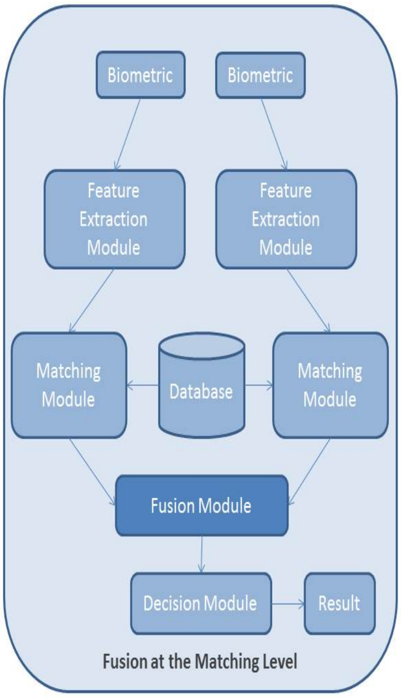

Figure 1.6: Block Diagram of Fusion at the Matching Score Level

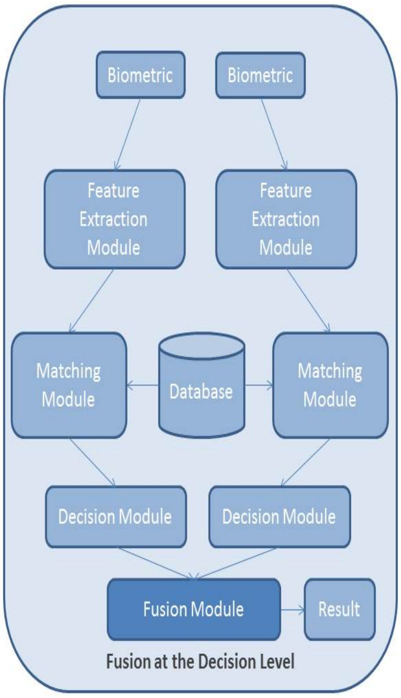

Figure 1.7: Block Diagram of Fusion at the Extraction Level

1.3 In This Project

For this thesis, the multimodal recognition system was constituted with the fusion of two unimodal systems at the matching score level and this multimodal performs on parallel mode for human identification. The speech signals and the ECG signals were used separately for these unimodal systems, which are speaker recognition system and ECG based recognition system, and also, artificial database was generated. In this thesis, the speaker recognition system performs in text-independent group and also these two unimodal systems perform in close set. The aim of this project is that the unimodal systems contribute the recognition rate with fusion of them.

Chapter 2

The Speech Signals

Sound is a vibration that propagates as typically audible mechanical wave of pressure and displacement, through a medium such as air or water[7]. In other words, sound may be defined as vibration of molecules of matters. Voice is one of the kinds of sound, which is produced by people. Vocal fords are vibrated by air, which comes from lungs, thus voice may be produced by people thanks to vibration of the vocal fords. People can convert the voice to speech with using teeth, mouth, lip and muscles of larynx. Speech signals are not perceived by human beings, because speech signal is conversion of voice in digital system to transmit, use or save the voice.

People have different voice; reason of this difference is physical structure of voice track. So, by help of the difference, people are able to recognize each other. The speech recognition systems like people use the difference and also similarity of voice to recognize people.

2.1 Speaker Recognition System

Like all biometric recognition systems, speaker recognition system has two phases to perform. These phases are called as training and test. In the training phase, instances of all users are gathered to constitute reference models. In the test phase, instance of claimed user is compared with reference models to verify and

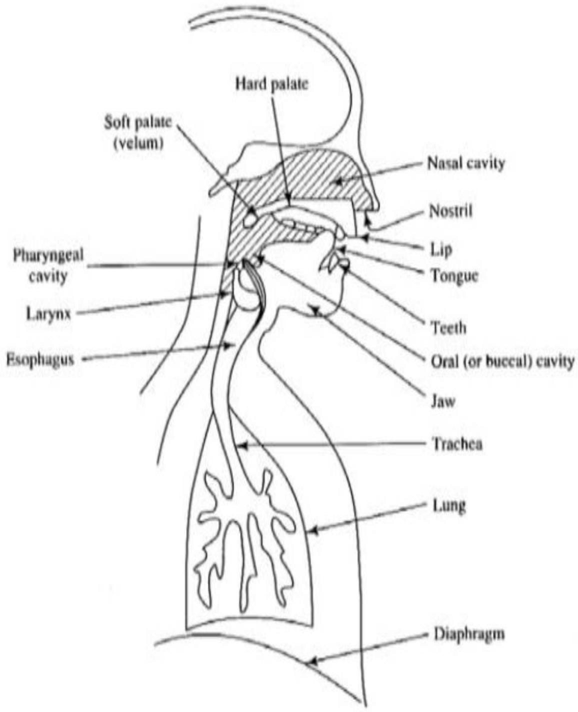

Figure 2.1: Human Speech Production Illustration [8].

identify. There are two main targets of speaker recognition system, which have been described previously. According to these targets, speaker recognition system may be grouped as following blocking.

Text-independent and text-dependent are sub-groups of the speaker recognition. Text-independent, which is easily understood from its name, may be defined as that user is free whatever he/she wants to say. Text-dependent may be defined as that user says certain words, sentences or text in contrast with Text-Independent.

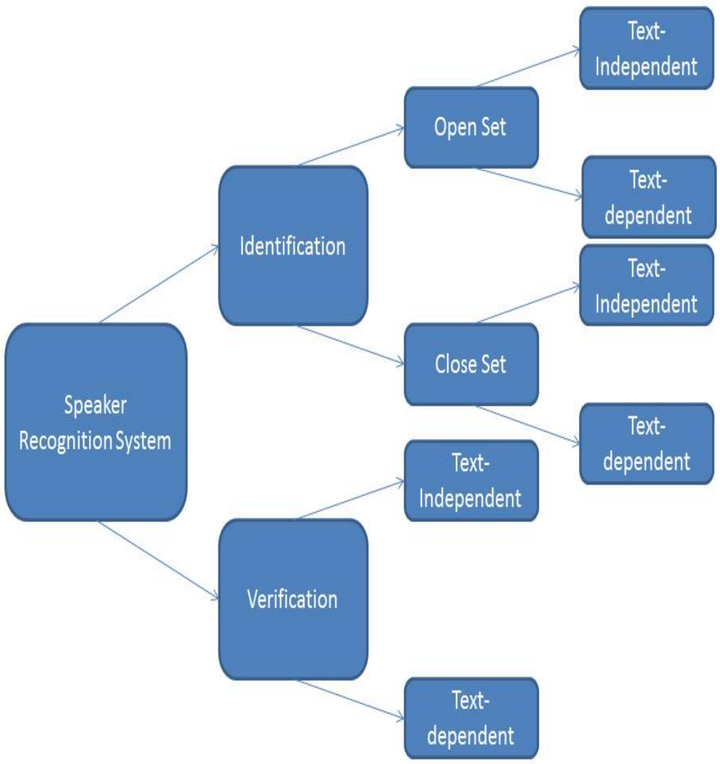

Figure 2.2: Grouping Structure of the Speaker Recognition System.

2.2 History of Recognition System based on Speech Signal

There are some historical improvements of recognition system based on speech signal are summarized below as mentioned in [9].

1950s 1960s:

- At 1952, the first system, which was constituted by Bell Labor, could perceive only digits.
- At 1960, the source-filter model of speech production was improved by Gunnar Fant.
- At 1962, Shoebox, which was introduced by IBM, could perceive 16 English words.

- At the late 1960s, continuous speech recognition was firstly interested by Raj Reddy.
- Dynamic Time Warping algorithm was created by Soviet researchers, who exploited this to constitute recognition system, which might perform with 200 words.

1970s:

- At 1971, Harpy, which was introduced by DARPA, could perceive 1011 English words.

1980s:

- At the mid-1980s, Tangora, which was constituted by IBM, was a voice activated typewriter and it could perform for 20,000.

1990s:

- At 1992, The Sphinx-II system was first system to be able to perform speaker-independent with huge amount of words as continuous speech recognition and also, its performance was the best until that this year.
- At 1990s, speech recognition systems were introduced for commercial.

2000s:

- Speech recognition ability was ensured to Siri with software, which was licensed by Apple from Nuance.
- Google performed for speech recognition first at 2007. Today, Google voice search may be supported in over 30 languages.

Chapter 3

The ECG Signals

The Electrocardiogram may be defined as illustration of the electrical activity of the heart. ECG signal are generally used by health workers to make assessments and medical diagnosis for their patients. In other words, it may be used to monitor and diagnose heart diseases but today, ECG signals may be used for also recognition system as biometric input, because length, weight, age and gender of human and also, position, structure and anatomy of the heart provide the ECG signals have distinctive features for person by person. The using of electrocardiogram as a biometric has been observed to generate successful recognition rate for biometric recognition systems, as mentioned in [10].

3.1 Details of the ECG Signal

The ECG wave consists of 3 basic components, which are named as P, T waves and QRS complex. In a normal ECG wave, firstly P wave occurs and then, QRS complex and T wave occurs, respectively.

- The P wave is occurred by atrial depolarization. Its duration is less than 120 milliseconds and it is assessed as low frequency, which is below 10–15 Hz.
- The QRS complex is occurred by ventricular depolarization. R is the highest and sharpest peak in the QRS. The frequency of QRS complex may be

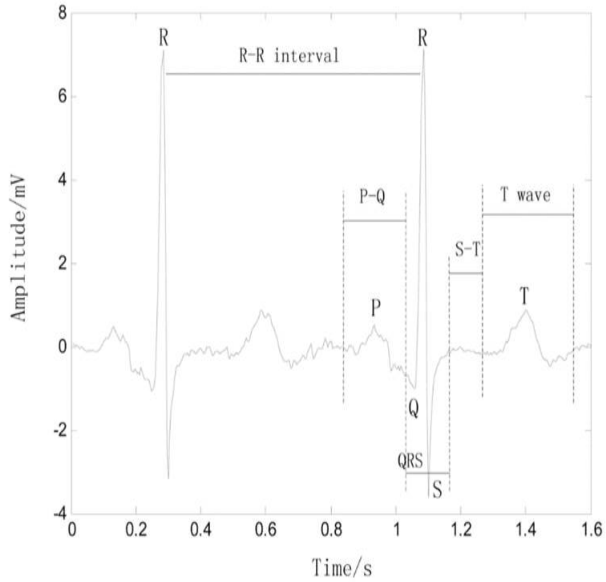

Figure 3.1: ECG wave consists of 3 basic components; P Wave, QRS Complex and T Wave. [11]

higher than other components of ECG signals. Its range is between 10 and 40 Hz and duration is about 70–110 milliseconds, as determined in [12].

- The T wave is occurred by ventricular repolarization and its duration is about 300 milliseconds.

The ECG wave components may be split into the PR and QT intervals and ST segment. These intervals are determined as measurement of time in the certain range.

- PR interval covers from the beginning of the P wave to the beginning of the QRS complex. Duration of the PR interval may be 120–200 milliseconds. The PR interval is defined as the starting depolarization of the atria to the starting depolarization of the ventricles and reflects a physiological delay in AV conduction imposed by the AV node, as mentioned in [13].

- QT interval covers between the beginning of the QRS complex and end of the T wave. Duration of the QT interval may be about 420 milliseconds. This interval may be determined as duration of ventricular depolarization and repolarization and depends on heart rate. Also, the QT interval may be used for detection of patients, who has a risk of ventricular arrhythmia.

- ST segment covers between S wave and the beginning of T wave. This interval may be determined as the ventricular depolarization.

The sampling rate, sampling precision, number of leads, recording time etc. may enlarge the amount of data, which obtained from the ECG signal. Clearly, record of enormous amount of the ECG signals necessitates high storage capacity and also wide transmission band for the remote monitoring activities. Thus, compression and modelling methods may meet these requirements for the ECG signal. These methods may be divided into the three main groups, which are named as the direct time-domain methods, the transform-based methods and the parameter extraction methods, as mentioned in [14] [15].

- The direct time-domain methods may be used to decrease redundancy in the samples of the ECG signal.

- The transform-domain methods create a coefficient sequence, which decrease the amount of data required to show the real signal, and then, these methods is performed to restore the actual signal with acceptable error.

- The parametric extraction methods create a set of parameters, which is obtained from the actual signal.

3.2 ECG Measurement Techniques

The commonly used ECG recording is the 12-lead ECG, which provides information about the heart function. The 12-lead ECG is used for recording the electrical activity of heart in waveform with 10 electrodes, which procure 12 perspectives. These 12 perspectives may be gathered by placings electrodes on the skin. The 12-lead ECG examines the heart in the two planes, which are frontal and horizontal planes. The 12 leads are classified into three different classes, which are named as standard limb leads, augmented leads and transverse leads. The standard limb leads and augmented leads give information about frontal plane of the heart and the transverse leads give information about horizontal plane of the heart. Lead I, lead II and lead III belong to the standard limb leads class and these leads exploit the comparison of electrical potential differences between two electrodes to give information about electrical activity of the heart. One of these two electrodes is negative and other one is positive, thus the standart limb leads are bipolar, as mentioned in [16] [17].

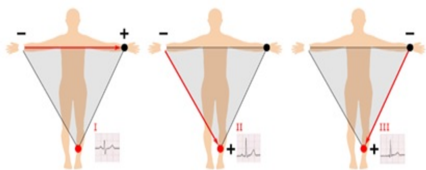

Figure 3.2: Illustration of lead I, lead II and lead III, which belong to the standard limb leads class [16] .

Augmented vector left (aVL), augmented vector right (aVR) and augmented vector foot (aVF) belong to the augmented leads class. These leads are unipolar, because they exploit only one electrode, which is positive.

The V1, V2, V3, V4, V5 and V6 belong to the transverse leads. They use only one positive electrode, which is placed on the chest.

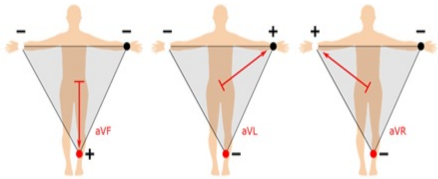

Figure 3.3: Illustration of augmented vector left (aVL), augmented vector right (aVR) and augmented vector foot (aVF), which belong to the augmented leads class [16] .

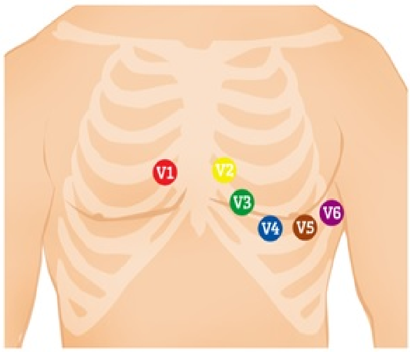

Figure 3.4: Illustration of the V1, V2, V3, V4, V5 and V6, which belong to the transverse leads [16] .

3.2.1 Einthoven's Triangle

The location of electrodes, which are placed for lead I, lead II and lead III, forms as triangle. These electrodes are placed at the right arm, left arm and left leg, so the Einthoven's triangle is occurred. This special triangle may be evaluated as a circuit, because logic of the Kirchhoff's law may be applied to the Einthoven's triangle, as mentioned in [16] [17]. In the Einthoven's triangle, the Kirchhoff's law is formulated for potential in lead I, lead II and lead III, as below.

$$LeadI + LeadIII = LeadII \tag{3.1}$$

3.3 Advantages of the ECG Signal

The most important advantages of the ECG, which makes it more reliable than other biometric treats against the frauds, is that the ECG signal is inimitable, because the ECG signal is occurred by the effect of autonomic nervous system and other sympathetic and parasympathetic factors, as clarified in [11]. Other advantages of the ECG signal is that generation of the ECG signals requires alived person and also, this requirement means that every alive human generate the ECG signal. And also, using the ECG signal as a biometric treat assures that owner of this treat has to be physically there. The assessment of the ECG according to the general properties, which is explained previously, as following.

- The universality is provided by help of that generation of the ECG signals require person to be alive and also, this requirement means that every persons have the ECG.
- The uniqueness is ensured, because the ECG signal may varies person to person by help of differences of length, weight, age and gender of people and also, heart position and structure.
- The permanence are also guaranteed and this characteristic was observed in researches as mentioned in [12].

3.4 The ECG Recognition System

Grouping of the ECG recognition system is like as grouping of the speaker recognition system, which is clarify previously, except sub-groups (text-independent and text-dependent).

And also, according to using methods, the recognition system based on the ECG signal may be divided in two main groups. These methods are named as fiducial methods and non-fiducial methods.

In the fiducial methods, features of the ECG signal are used. These features are classified into 3 fundamental sections, which are temporal feature, amplitude feature and morphological feature.

- Location of the ECG wave components are used to calculate various temporal intervals, which are assessed as the temporal feature.
- Amplitude of the ECG wave components are used to distinguish people from each other. The amplitude may be determined as the peaks of ECG wave components, which is commonly R peak of QRS complex.
- Morphological feature is related with the shape of the ECG wave. The morphological feature may be assessed by obtain average of the sampled values of intervals, such as QRS complex, with respect to multiple aligned the ECG wave.

The non-fiducial methods do not need any detection of fiducials for recognition. These methods performs with iterative structure of the ECG wave and these methods may be classified into autocorrelation based, phase space based and frequency based analyses.

3.5 History of The ECG Recognition System

In the last 15 years, various systems and applications based on ECG signal were improved. The ECG was examined by Biel and others to be used in human identification for the first time at 2001. In this research, a commercial ECG equipment was used to extract the 30 features, which are temporal and amplitude and then, this feature amount was decreased to 12 features by using the feature selection algorithm, which is based on analysis of correlation matrix and also, for classifying, multivariate analysis methods were performed. In this research, with using 360 features, which was obtained from 12 channels, the recognition rate accomplished to %100. Nevertheless, this method could not have an automatic recognition skill, because features were extracted by the ECG equipment [18]. Irvine and others discovered that the heart rate variability may be employed in human identification [19].

In 2002, Shen and others were proposed a method, which consists of the two steps, to recognize individuals from the 1 channel ECG signal. In the first step, distance between the two QRS complexes is compared by computing the auto-correlation coefficients with using the template matching method. In the second step, by help of the Decision-Based Neural Network-DBNN algorithm, recognition was performed for possible subjects, which was obtained by using the template matching method. In this research, seven features, which present the temporal and amplitude values were exploited. Consequently, the recognition rate of template matching method reached to %95 and the recognition rate of Decision-Based Neural Network-DBNN method reached to %80 recognition rate, and the recognition rate of fusion of these two methods reached to %100 [20].

In 2003, Israel and the others introduced a recognition system based on integrated of the ECG and face signals. This system can be accepted as an automatic recognition. However, the recognition rates of this system were low [21].

In 2005, Israel and others found out that 15 features, from the fiducial points, were exploited to describe the uniqueness of a human heartbeat and then, this amount of features was decreased to 12 by using Wilks Lambda method. In this research, the Linear Discriminant Analysis-LDA was exploited to classify the features and then, it reached to %81 heartbeat recognition rate and %100 human recognition rate [22] [12]. In the same year, Shen improved his previously research [20] and also expanded database, which consists of 168 healthy individuals ECG signals. In this research, the method consists of two steps. In the first step, template matching and mean-square error were analyzed comparatively and in the second step, the Decision-Based Neural Network-DBNN and distance classification method were analyzed comparatively. Consequently, the 17 features, which are temporal and amplitude, were obtained. As a result of comparison, the template matching and distinct classification has best recognition rate, which is %95.3 [12] [23].

In 2006, Shen ensure the %95 recognition rate via decreasing 17 feature, which were extracted in his previously research [23], to 7 feature for the same database [24] [12].

The Einthoven channels were used by Wübbeler and others to introduce a two-dimensional heart vector for biometric recognition system in 2007 and they obtained %98.1 recognition rate [25].

In 2008, Wang and others developed new method, which does not require any fiducial detection and also, based on estimation and comparison of the coefficients, which are obtained by applying the discrete cosine transform to the autocorrelation of heartbeat signals. They achieved to %94.47 and %97.8 recognition rate by help of using the two different databases [12]. At the same year, Chuang-Chien and others presented a method, which apply wavelet transform to the ECG signal to obtain coefficients for feature set. They used a database, which comprises ECG signals of the 35 healthy individiuals and 10 ECG signals of 10 arrhythmia individiuals. This method reached to %100 recognition rate for healthy individiuals and %81 recognition rate for arrhythmia individiuals [26]. Again in 2008, Irvine and others introduced a technique, which called as eigenPulse and based upon principle component analysis, for identification. They concluded that their technique had a better performance than fiducial feature according to experimental results of their research [27].

In 2009, unsupervised ECG, which means that the ECG signals are remodeled on phase space, was improved by Fang and Chan for identification. In the proposed method, the ECG signals may be 1 channel or 3 channel and recognition performs according to the similarities and differences between the phase space descriptions. This research reached to %93 recognition rate for 1 channel ECG and %99 recognition rate for 3 channel ECG [28]. In the same year, a recognition system based on ECG signals was proposed by Fatemian and Hatzinakos. In this proposed system, QRS segments of the ECG signals were exploited to recognize individuals. This proposed system reached to %99.61 recognition rate [29].

In 2010, Li and Narayanan improved new biometric recognition algorithm with exploiting the temporal and power spectrum information of the ECG signals. In this research, firstly, QRS determination process was applied to the ECG signals for extracting the temporal features for the owner of these ECG signals. And then, the heartbeat waveforms of the owner were obtained and the heartbeat waveforms were adjusted with using normalization method to be in the same time and amplitude scale. The Hermites polynomial expansion is exploited to normalize the heartbeat waveform. The hermite polynomial coefficients were computed for normalized heartbeat waveform. The support vector machine method was applied the Hermite polynomial coefficients, which represents the temporal information, to classify them. And also, the cepstral features were computed to model the frequency of the ECG signals and these features were modeled by Gaussian Mixture Model. This research reached to %98.26 recognition rate with using two feature sets [30]. During the 2010, Sufi and others introduced a biometric recognition system, which uses polynomial distance measurement to recognize the individuals [31]. At the same year, Loong and others introduced a biometric recognition system, which exploits the Linear Zpredictive coding to recognize the individuals from their ECG signals. They reached to %99 recognition rate [32]. In 2010, Ting and Salleh proposed a method, which based on Kalman filters for identification [33].

In 2011, Sufi and Khalil introduced a method to extract the ECG feature sets from the compressed ECG signals with using the data mining techniques [34]. In the same year, Safie and others generated a new feature extraction method, which is known as Pulse Active Ratio-PAR, and also, they extracted a new feature set via using this new method. In this proposed method, it has been determined that the length of the feature vector should be between 20 and 40 according to experimental work [15]. In 2011, Shen, Tompkins and Hu improved ECG signal based recognition system, which records ECG signals from the palm on the channel 1 [35].

In 2013, Gürkan and other exploited the AC/DCT features, MFCC features and QRS complex of ECG signals, which were obtained from lead 1, to constitute the feature sets. They improved a biometric recognition algorithm based on these feature sets. In this research, the recognition rate reached to %97.31 [36].

Between 2014 and 2016, Gökhan designed the measuring system, which is cheap, easy to use, portable and also, has low noise level, to measure the ECG signals on the lead 1. This system measures the ECG signals from the left and right hands thumbs as taken fingerprint. In this research, a biometric recognition algorithm was improved with using the ECG signals, which were recorded by this designed measuring system. Consequently, recognition rate of this research reached to %94.72. The research work was supported by Coordination Office for Scientific Research Project, ISIK University BAP project, Project Number: 14A203, Project name: Fingertip ECG Signal Based Biometric Recognition System [37].

Chapter 4

Feature Extraction and Classification Methods

4.1 Feature Exraction Method

4.1.1 Mel Frequency Cepstrum Coefficients

MFCCs are commonly used as feature in speaker recognition systems, because MFCC simulates the perception of human's ear. So, MFCC may not perceive frequencies over 1Khz [38] and also, MFCC may be influenced far less from chancing. Thus, it makes MFCC desirable. The MFCC method consists of 7 sections, which may be named as Pre-Emphasis, Frame Blocking, Windowing, Fast Fourier Transformation, Mel-Frequency Warping, Cepstrum and Liftering, respectively.

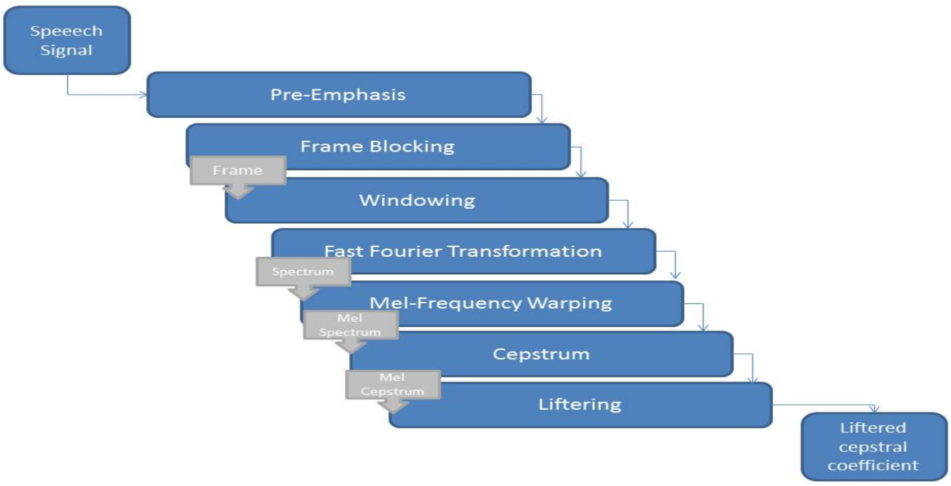

Figure 4.1: Block Diagram of the MFCC.

4.1.1.1 Pre-Emphasis

The first section is Pre-emphasis. "Pre-emphasis refers to a system process designed to increase, within a band of frequencies, the magnitude of some (usually higher) frequencies with respect to the magnitude of the others(usually lower) frequencies in order to improve the overall Signal-to-noise ratio." [39]. Thus, filter is applied to speech signal. Thanks to it, high frequencies are emphasized by increasing the energy in this section. It is applying with following equation:

$$y[n] = x[n] - ax[n] \tag{4.1}$$

where value of a is between 0.9 and 1.

4.1.1.2 Frame Blocking

The second section is Frame Blocking. The character of speech is not stable, while speech organs perform for the speech production. However, the character of speech may be stable in the short period of time. The speech should be divided into small parts, where the parameters of speech is stable. [38] Thus, the frame blocking provides the speech signal to be divided into frame, the stationary character of speech may be obtained by help of frame blocking.

During the frame blocking, overlapping is applied to all frame. Thanks to overlapping, the character of speech, where beginning and end of frame, may be protected to lose its significance.

Overlapping structure is ilustrated below.

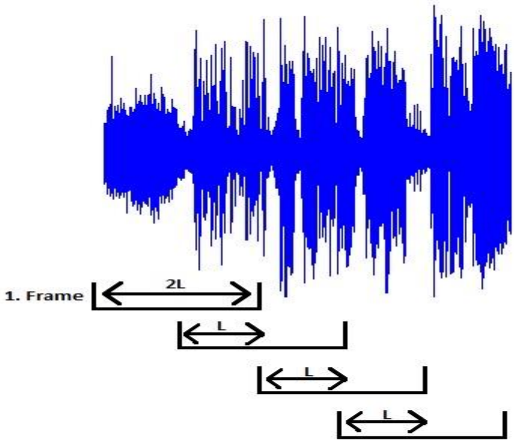

Figure 4.2: Frame blocking with 2L length frames and overlapping with L length shifting

4.1.1.3 Windowing

The third section is windowing. To reduce discontinuity in beginning and end of frame, windowing is applied to all frame. The purpose of windowing is to cut the part, which does not include any information, in beginning and end of frame. In other words, basically the spectral distortion is minimized by using windowing at both beginning and end of each frame [40].

The hamming window is one of the windowing method, which is used to reduce discontinuity. The formula of hamming window is as following.

$$w(n) = 0.54 - 0.46Cos\left(\frac{2\pi n}{N-1}\right) \qquad n = 0, ..., N-1 \tag{4.2}$$

In the formula, N represents the number of samples in each frame.

Windowing may be defined as the multiplication of a rectangle window and signal.

$$y(n) = x(n)w(n) \tag{4.3}$$

4.1.1.4 Fast Fourier Transformation

The Fast Fourier Transform is fourth section, which is used to convert frames from time domain to frequency domain. The FFT, a method for computing the Discrete Fourier Transform with reduced execution time [41]. In other words, FFT may be defined as a fast algorithm to perform for applying the Discrete Fourier Transform, which is formulated as following.

$$X_n = \sum_{k=0}^{N-1} x_k e^{\frac{-2\pi jkn}{N}} \qquad n = 0, ..., N-1 \tag{4.4}$$

4.1.1.5 Mel-Frequency Warping

The Mel-Frequency Warping is fifth section. In this section, mel-spectrum is procured by converting the frequency components, which are obtained from previous section, to the Mel scale. "The frequencies range in FFT spectrum is very wide and voice signal does not follow the linear scale."[38]. Thus, the mel scale is required. The structure of Mel scale may be defined as linear up to the 1 KHz and logarithmic after 1KHz. The Mel-Frequency Warping is computed by following formula.

$$Mel(f) = 2595 * log10(1 + f/700) \tag{4.5}$$

where f indicates frequency in Hz, and *Mel(f)* indicates the perceived frequency.

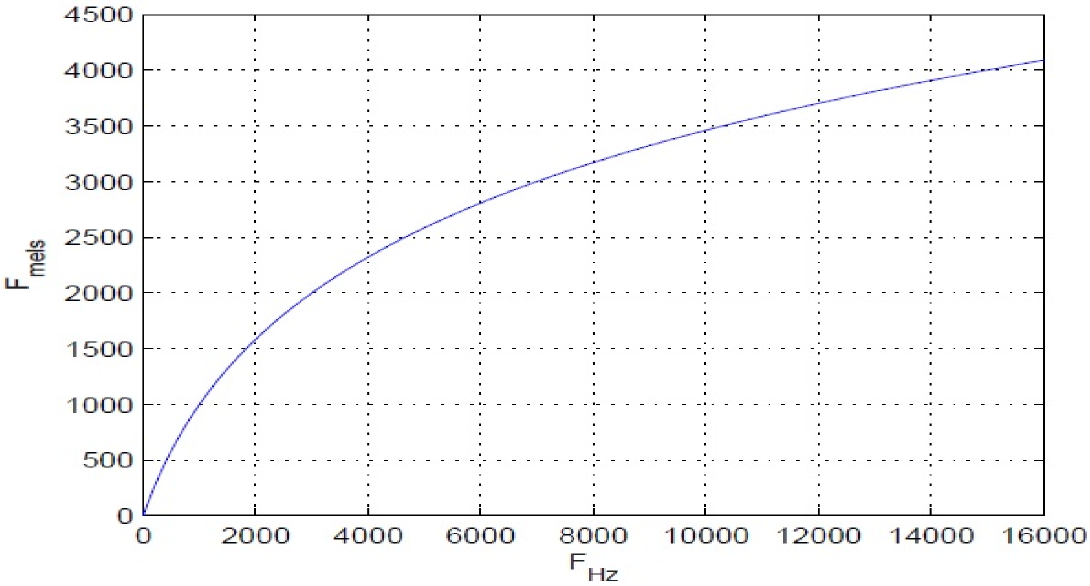

Figure 4.3: The Mel Scale

"One approach to simulating the subjective spectrum is to use a filter bank, one filter for each desired Mel frequency component. That filter bank has a triangular band pass frequency response, and the spacing as well as the bandwidth is determined by a constant Mel-frequency interval. The modified spectrum thus consists of the output power of these filters."[39]

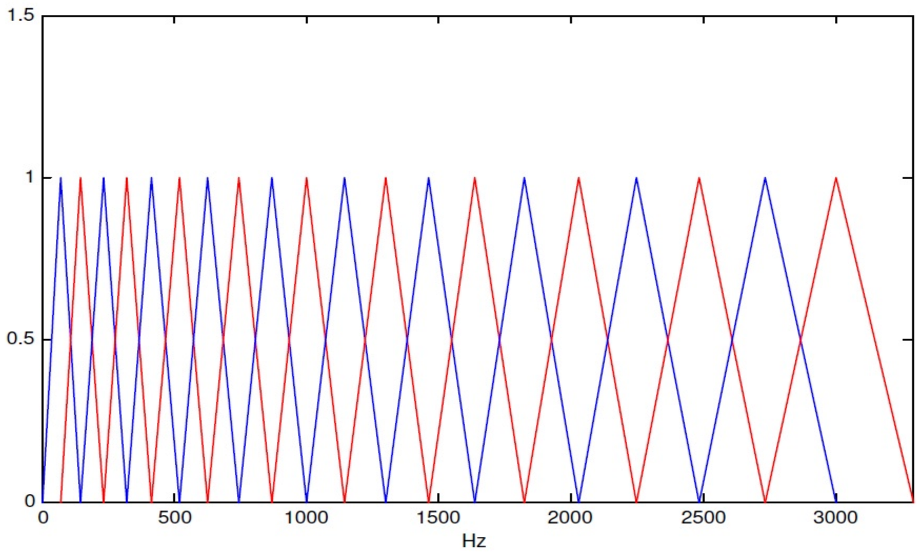

Figure 4.4: The Filter Bank

4.1.1.6 Cepstrum

The sixth section is called as Cepstrum. In this section, the Mel Frequency Cepstrum Coefficients are obtained by converting logarithm of previous section's result, which is called as Mel Spectrum, to time domain via using Discrete Cosine Transform. "The cepstral representation of the speech spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis." [39]. Thus, via using DCT, previous section's results are converted into time domain, thanks to that Mel Spectrum Coefficients consist of real numbers. After this operation, the MFCCs may be obtained as following formula;

$$C_n = \sum_{k=1}^{K} (\log S_K) \cos[n(k - \frac{1}{2})\frac{\pi}{K}] \qquad n = 1, 2, ..., K \tag{4.6}$$

where *Sk* represents to Mel Spectrum Coefficients.

And also, it should be known that the first *Co* is eliminated from DCT, because the *Co* defines the mean of input, which involves few information about the related person.

4.1.1.7 Liftering

The final section is Liftering. "If the cepstral coefficients are used as recognition features, which the performance of the recognition system can be improved significantly by liftering the cepstral coefficients." [42]. Thanks to apply the Liftering function to the cepstral coefficients, undesirable components are reduced or removed. Eventually, the recognition rate increases. One of type of Lifters is the Sinusoidal lifter, which is formulated as following expression.

$$w_i = 1 + \frac{D}{2} \sin \frac{\pi i}{D} \qquad i = 1, 2, ..., D \tag{4.7}$$

Where wi represents weight and i represents lifters.

The Liftered Cepstral Coefficient may be formulated as following expression.

$$\widehat{C}_i = wiC_i \tag{4.8}$$

4.1.2 The Discrete Cosine Transform of the Autocorrelation

The feature extraction methods and features are classified into 4 different classes for ECG. These classes are named as temporal and amplitude features, features based on transform methods, feature based on parametric methods and feature based on frequency information. The Discrete Cosine Transform is assessed as member of the feature based on transform methods. In this class, The ECG feature sets may be obtained by applying linear transform methods.

ECG signal are generally used by health workers to make assessments, medical diagnosis for their patients. But today, ECG signals may be used for also recognition system as biometric input. Nonetheless, determination of fiducial detection is quite hard for classifying of features of ECG for the recognition systems. Because, there is no methods, rules or information for exact localization of wave boundaries. Nevertheless, the AC/DCT does not need localization of wave boundaries to achieve the target of biometric recognition systems, as mentioned in [12].

The Discrete Cosine Transform of autocorrelation of ECG consists of 3 steps, which are framing, computing autocorrelation and applying DCT.

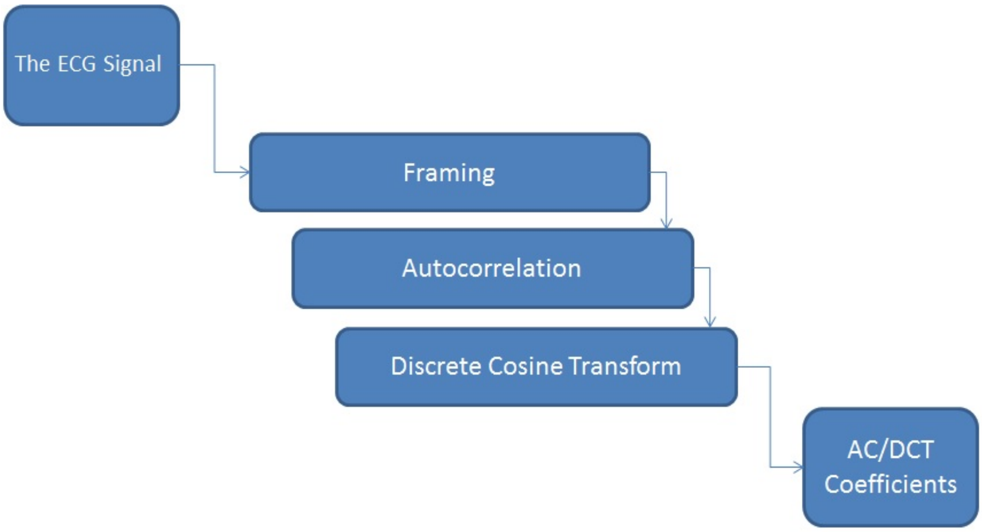

Figure 4.5: Block Diagram of the AC/DCT.

4.1.2.1 Framing

In the Framing step, the ECG signals are divided into frames. Thanks to framing, half stationarity of ECG signals may be captured. Because stationarity is not captured for ECG signal without the framing, as clarified in [14].

4.1.2.2 Autocorrelation

After the framing, the Autocorrelation Coefficients are calculated with following formula.

$$\widehat{r}_m = \sum_{i=0}^{N-|m|-1} x(i)x(i+m) \qquad m = 0, 1, 2, ..., L \tag{4.9}$$

Where x[i] represents frames of the ECG signal, x[i+m] represents time-shifted version of these frames.

After computation of the AC coefficients, these coefficients should be normalized by dividing with the R[0], which is maximum value, to eliminate the biasing factor, as mentioned [12].

The formulation of normalization is as following.

$$r_m = \frac{\widehat{r}_m}{\widehat{r}_0} \tag{4.10}$$

The AC is used for discovering of nonrandom patterns and also, the using AC provides to mix in a sequence of sums of products samples, hence exact localization of wave boundaries, which may be assessed as fiducial, are not requisite, as mentioned [12].

4.1.2.3 Discrete Cosine Transform

Final step of the AC/DCT is applying DCT to results of the previous step. This step provides to decrease the dimensionality of AC. The DCT is formulated as following.

$$c(u) = \alpha(u) \sum_{m=0}^{L} r(m) \cos\left[\frac{\pi(2m+1)u}{2N}\right] \tag{4.11}$$

Where $\alpha(u)$ is formulated as following.

$$\alpha(u) = \begin{cases} u = 0, & \text{for } \sqrt{\frac{1}{N}} \\ u \neq 0, & \text{for } \sqrt{\frac{2}{N}} \end{cases} \tag{4.12}$$

Because of the energy compaction of DCT, amount of the important DCT coefficients is decreased by that value of these coefficients will equals to near-zero, as clarified in [12]. By help of this operation, the dimensionality of AC is decreased.

4.2 Classification Method

4.2.1 Gaussian Mixture Model

The Gaussian distribution is prevalently preferred to use into huge amount of discipline of engineering and science, because that it has suitable computational attributes and also, it may simulate real-world data, as clarified in [43]

As mentioned in [44], the Gaussian Mixture Model is a parametric probability density function represented as a weighted sum of Gaussian component densities. In other words, the GMM may be determined as a classifier, which performs with a weighted combination of multi-variate Gaussians distributions to model the data.

The Gaussian Mixture Model is formulated as following.

$$p(x \mid \lambda) = \sum_{i=0}^{M} w_i g(x \mid \mu_i, \varepsilon_i) \tag{4.13}$$

Where wi is one of the parameter of GMM and called as the mixture weight and M represents number of components as GMM of order.

The formula may be interpreted as a weighted sum of M components Gaussian densities, so it shows that the enormous ability of representing a large class of sample distributions.

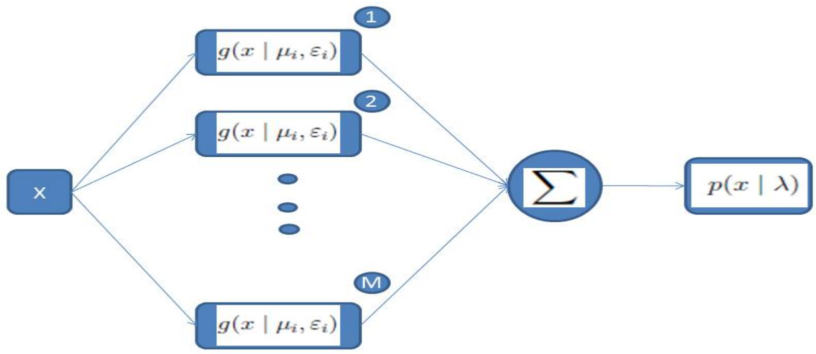

Figure 4.6: The GMM models the probability density function of a data set as a weighted sum of multivariate Gaussian probability density function.

The component Gaussian density is formulated as following.

$$g(x \mid \mu_i, \varepsilon_i) = \frac{1}{(2\pi)^{2D} \mid \varepsilon_i \mid^{1/2}} e^{-\frac{1}{2}(x-\mu_i)'\varepsilon_i^{-1}(x-\mu_i)} \tag{4.14}$$

Where μ represents mean vector,

ε represents covariance matrix and D is the dimension of data set.

By help of analyzing these formulas, the GMM has 3 parameters, which are the mixture weight, covariance matrix and mean vector.

The condition of the mixture weight is to accomplish the following formula.

$$\sum_{i=1}^{M} w_i = 1 \tag{4.15}$$

This condition requires to be accomplished to gather that it is a correct model.

These 3 parameters of GMM, which are means, variances and weights, are estimated from the data set to procure the model. This estimation operation is performed by using the maximum likelihood methods and expectation maximization algorithm.

These parameters set is shown as $\lambda = (w_i, \mu_i, \varepsilon_i)$

4.2.1.1 The Parameter Estimation

The parameter estimation aims to estimate the most possible prober parameters of GMM with using maximum likelihood method and the expectation maximization algorithm. The maximum likelihood algorithm aims to determine the parameters by maximizing the likelihood of GMM given the data set and then, the expectation maximization maximizes the likelihood function, iteratively. Because ML is non-linear function, as mentioned in [44], so the EM may be exploited to gather the ML. The EM has an iterative structure to perform and the structure consists of two steps, which are called as E-step for expectation and M-step for maximization, respectively.

The GMM likelihood may be formulated as following for a data vectors.

$$p(X \mid \lambda) = \prod_{t=1}^{T} p(x_t \mid \lambda) \qquad X = x_1, ..., x_T \tag{4.16}$$

In the Maximization Step (M-Step), the likelihood of GMM is increased monotonically by using formulas of parameters of GMM, as below.

$$\text{Mixture Weight} \qquad \overline{w_i} = \frac{1}{T} \sum_{t=1}^{T} Pr(i \mid x_t, \lambda) \tag{4.17}$$

$$\text{Mean} \qquad \overline{\mu_i} = \frac{\sum_{t=1}^{T} Pr(i \mid x_t, \lambda) x_t}{\sum_{t=1}^{T} Pr(i \mid x_t, \lambda)} \tag{4.18}$$

$$\text{Variance} \qquad \overline{\sigma_i^2} = \frac{\sum_{t=1}^{T} Pr(i \mid x_t, \lambda) x_t}{\sum_{t=1}^{T} Pr(i \mid x_t, \lambda)} - \overline{\mu_i^2} \tag{4.19}$$

In the Estimation Step (E-Step), a posteriori probability is computed by using following formula.

$$Pr(i \mid x_t, \lambda) = \frac{w_i g(x_t \mid \mu_i, \varepsilon_i)}{\sum_{k=1}^{M} w_k g(x_t \mid \mu_k \varepsilon_k)} \tag{4.20}$$

The EM starts to perform with initial parameters to estimate new parameters, which satisfy the condition as below.

$$p(X \mid \overline{\lambda}) \geq p(X \mid \lambda)$$

The iterations work by that new parameters become the initial parameters for the next iteration and it continues until some convergence threshold is provided.

Chapter 5

Speech Signal and ECG Signal based Recognition System

5.1 Feature Extraction Methods

In this project, two unimodal recognition systems, which are based on ECG signals and speech signals, are fused to constitute a multimodal recognition system. For ECG signals and speech signals, there are two databases, which were gathered from various sources, were used. There are two feature sets, which were extracted from ECG signals and speech signals by using feature extraction methods. The feature set for speech signals was extracted by using the MFCC method and the feature set for the ECG signals was extracted by using the AC/DCT method. All biometric recognition systems have 2 main processes, which are called as training and test. In the training process, the system is trained to perform successfully for recognition in the test process. For both of these two processes, feature extraction has to be applied to the biometrics.

5.1.1 Extraction of MFCC based Features from the Speech Signals

Applying the MFCC method, which were explained previously;

In the pre-emphasis phase, a first order Finite Impulse Response (FIR) is applied to the speech signals with following formula.

$$y[n] = x[n] - ax[n]$$

where value of a is 0.97 for this thesis. After this operation, the high frequency is increased to improve the overall SNR. In the frame blocking phase, speech signals are divided into frames to obtain stable parameters of speech. The sampling frequency of speech signals is 44.1 kHz. There are two databases were constituted; first database consists of 2 seconds speech signals and second database consists of 10 seconds speech signals. In this project, 10 ms frames were used to divide speech signals into frames and also, frame blocking was applied with 5 ms overlapping, which prevents loss of information at the beginning and end of frames. The overlapping rate is 50%. Thereafter, windowing, FFT, Mel-Frequency Warping, Cepstrum and Liftering phases were applied to obtain 12 MFCCs.

5.1.2 Extraction of AC/DCT based features from the ECG signals

Applying the AC/DCT method, which was explained previously;

The database for the ECG signals was provided from MIT-BIH. In this database, frequency rate of the ECG signals is 360 Hz and their duration is 30 minutes [45]. Before the framing, the Butterworth filter was applied to the ECG signals to eliminate noise and distortion. And then, the filtered signal was normalized with its mean and variance values. The ECG signals were divided into frames, whose consist of 1800 samples, so its duration is 5 seconds. After this operation, the autocorrelation function was applied to all frames to obtain the autocorrelation coefficients, whose amount is 50. Finally, DCT was applied to these coefficients to decrease dimensionality.

5.2 Block Diagram

First part of this project is performing unimodal recognition system for both the ECG signals and the speech signals to obtain recognition rates. These recognition rates are used for comparison of unimodal recognition system and multimodal recognition system.

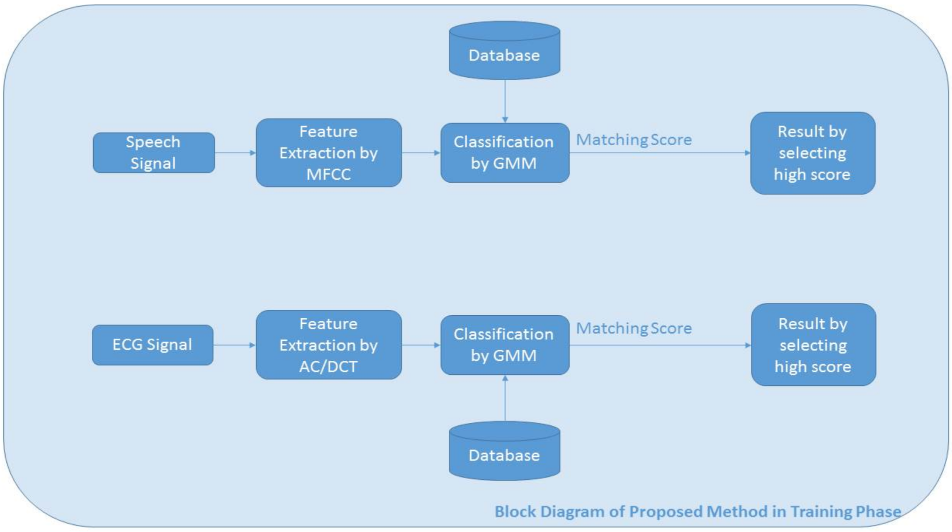

Figure 5.1: Block Diagram of Unimodal Speech Signal based Recognition System and Unimodal ECG Signal based Recognition System.

Second part of this project is fusion of these two unimodal recognition systems. The fusion was performed at matching level.

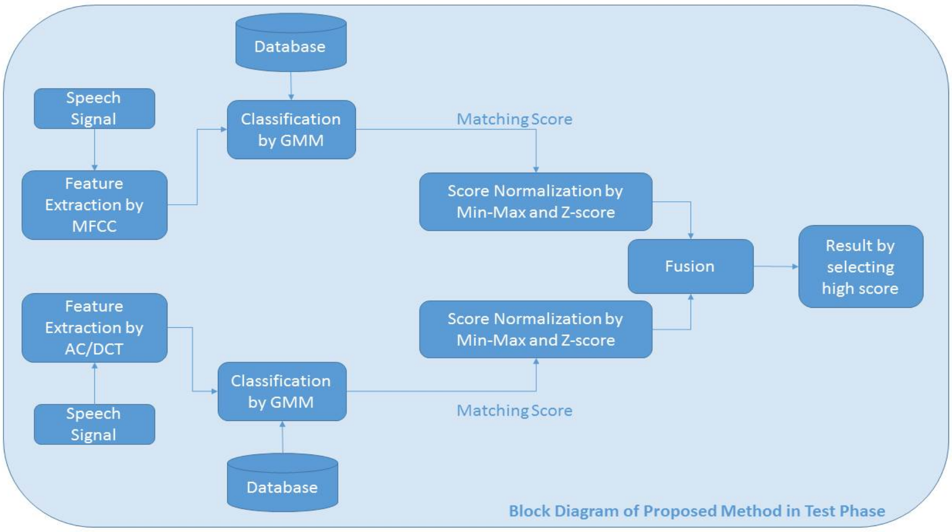

Figure 5.2: Block Diagram of Multimodal Speech Signal and ECG Signal based Recognition System.

5.3 Fusion & Matching

After the normalization, fusion can be applied to score matrices. The fusion is basic sum operation and its result, which is called as fusion matrix, is obtained from the sum of two score matrices. After gathering the fusion matrix, the matching is performed in the test process. The maximum value of per columns of matrix is the matching result. However, before the fusion, matching is performed to score matrices separately for obtaining the recognition rates of unimodal biometric recognition based on the ECG signal and speech signal.

Chapter 6

Experimental Work

6.1 Datasets

For this project, there are two datasets were used. The one of these datasets, which was provided from MIT-BIH, is for the ECG signals. This dataset consists of 48 persons, the 22 of 48 are females and the 26 of 48 are males. Duration of these signals is 30 minutes and these signals were divided into two parts to be used into training and test phaseses. Therefore, duration of the ECG signals are 15 minutes.

The other dataset was provided from the website given in [1]. This dataset was chosen according to the ECG dataset, hence it consists of 48 persons, the 22 of 48 are females and the 26 of 48 are males. In this project, there were two speech datasets, which were consist of the speech signals from the same persons. The difference of dataset is duration. The first dataset's duration is 2 seconds and other's duration is 10 seconds.

After training and testing phases, each speech signal and ECG signal were matched according to the gender to represent only one person. End of the matching, there are 48 persons and each of them have both ECG signal and speech signal. Thus, the dataset is obtained as artificial. Instead of constituted dataset, the artificial data set is used, because that there was not enough time, equipments and persons.

6.2 Normalization

Normalization prepares scores, which are into different nature, to be merged with adjusting them to be into same range. Normalization is nonignorable for multimodal recognition systems, which may be exploited to fuse scores in different scale. Normalization may be generally performed before the fusion in the multimodal recognition systems. Two normalization techniques, which are Z-score normalization and min-max normalization, were used in this project.

6.2.1 Min-max Normalization:

This technique adjusts all scores to be into interval [0,1]. Firstly, the maximum and minimum scores are selected from score set and then, following formula is applied to all scores.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{6.1}$$

6.2.2 Z-score normalization:

Firstly, the mean score and standard deviation are calculated from the score set and then, following formula is applied to all score.

$$x' = \frac{x - \text{mean}(x)}{\text{std}(x)} \tag{6.2}$$

Both of these normalization methods were applied to results of classification operation before the fusion. The recognition rates of them are same. These matching results are listed as below.

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 6	True
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 9	True
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 12	True
The test person no: 13	Match to the train person no: 13	True
The test person no: 14	Match to the train person no: 14	True
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 16	True
The test person no: 17	Match to the train person no: 17	True
The test person no: 18	Match to the train person no: 18	True
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 20	True
The test person no: 21	Match to the train person no: 21	True
The test person no: 22	Match to the train person no: 22	True
The test person no: 23	Match to the train person no: 42	False
The test person no: 24	Match to the train person no: 24	True

Table 6.1: Matching results for using the min-max normalization and the Z-score method to ECG signal and 10 seconds speech signal based recognition system with % 95.8. (Part 1)

The test person no: 25	Match to the train person no: 25	True
The test person no: 26	Match to the train person no: 26	True
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 29	True
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 37	True
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 40	True
The test person no: 41	Match to the train person no: 24	False
The test person no: 42	Match to the train person no: 42	True
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 45	True
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 47	True
The test person no: 48	Match to the train person no: 48	True

Table 6.2: Matching results for using the min-max normalization and the Z-score method to ECG signal and 10 seconds speech signal based recognition system with % 95.8. (Part 2)

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 6	True
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 16	False
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 16	False
The test person no: 13	Match to the train person no: 27	False
The test person no: 14	Match to the train person no: 7	False
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 16	True
The test person no: 17	Match to the train person no: 17	True
The test person no: 18	Match to the train person no: 18	True
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 27	False
The test person no: 21	Match to the train person no: 27	False
The test person no: 22	Match to the train person no: 34	False
The test person no: 23	Match to the train person no: 23	True
The test person no: 24	Match to the train person no: 24	True

Table 6.3: Matching results for using the min-max method to ECG signal and 2 seconds speech signal based recognition system with % 70.8. (Part 1)

The test person no: 25	Match to the train person no: 25	True
The test person no: 26	Match to the train person no: 26	True
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 11	False
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 8	False
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 27	False
The test person no: 41	Match to the train person no: 5	False
The test person no: 42	Match to the train person no: 36	False
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 45	True
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 27	False
The test person no: 48	Match to the train person no: 45	False

Table 6.4: Matching results for using the min-max method to ECG signal and 2 seconds speech signal based recognition system with % 70.8. (Part 2)

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 6	True
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 16	False
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 16	False
The test person no: 13	Match to the train person no: 27	False
The test person no: 14	Match to the train person no: 7	False
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 16	True
The test person no: 17	Match to the train person no: 17	True
The test person no: 18	Match to the train person no: 18	True
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 27	False
The test person no: 21	Match to the train person no: 27	False
The test person no: 22	Match to the train person no: 34	False
The test person no: 23	Match to the train person no: 23	True
The test person no: 24	Match to the train person no: 24	True

Table 6.5: Matching results for using the Z-score method to ECG signal and 2 seconds speech signal based recognition system with % 70.8. (Part 1)

The test person no: 25	Match to the train person no: 25	True
The test person no: 26	Match to the train person no: 26	True
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 11	False
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 8	False
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 27	False
The test person no: 41	Match to the train person no: 5	False
The test person no: 42	Match to the train person no: 23	False
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 45	True
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 27	False
The test person no: 48	Match to the train person no: 45	False

Table 6.6: Matching results for using the Z-score method to ECG signal and 2 seconds speech signal based recognition system with % 70.8. (Part 2)

Chapter 7

Results

The recognition rate for the ECG signal based recognition system was %87.5 and 42 persons were matched correctly. Matching results table is as following.

The recognition rate for 2 seconds speech signals is %58.3 and 28 persons are matched correctly. Matching results table is as following.

Normalization was applied before fusion of these two score matrices. The recognition rate after fusion was %70.8 and 34 persons were matched correctly. Matching results table is shown in previous chapter.

However, when the recognition rates are considered, it has been observed that the recognition rate, which obtained after the fusion, is lower than recognition rate of the ECG signal. Therefore, instead of 2 second speech signals, 10 second speech signals were used. In this case, the recognition rate of the speech signals was %97.9 and 47 persons were matched correctly. Matching results table is as following.

Then, normalization was applied again and two sets of data were fused. After the fusion, the rate of recognition reached %95.8 and 46 persons were matched correctly. Matching results table is shown in previous chapter.

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 27	False
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 9	True
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 12	True
The test person no: 13	Match to the train person no: 13	True
The test person no: 14	Match to the train person no: 14	True
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 16	True
The test person no: 17	Match to the train person no: 17	True
The test person no: 18	Match to the train person no: 18	True
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 27	False
The test person no: 21	Match to the train person no: 21	True
The test person no: 22	Match to the train person no: 22	True
The test person no: 23	Match to the train person no: 32	False
The test person no: 24	Match to the train person no: 24	True

Table 7.1: The recognition rate for the ECG signal based recognition system was %87.50. (Part 1)

The test person no: 25	Match to the train person no: 25	True
The test person no: 26	Match to the train person no: 26	True
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 29	True
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 37	True
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 40	True
The test person no: 41	Match to the train person no: 5	False
The test person no: 42	Match to the train person no: 32	False
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 31	False
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 47	True
The test person no: 48	Match to the train person no: 48	True

Table 7.2: The recognition rate for the ECG signal based recognition system was %87.50. (Part 2)

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 6	True
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 48	False
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 16	False
The test person no: 13	Match to the train person no: 27	False
The test person no: 14	Match to the train person no: 7	False
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 27	False
The test person no: 17	Match to the train person no: 23	False
The test person no: 18	Match to the train person no: 38	False
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 1	False
The test person no: 21	Match to the train person no: 38	False
The test person no: 22	Match to the train person no: 34	False
The test person no: 23	Match to the train person no: 23	True
The test person no: 24	Match to the train person no: 24	True

Table 7.3: The recognition rate for the 2 seconds speech signal based recognition system was %58.3. (Part 1)

The test person no: 25	Match to the train person no: 12	False
The test person no: 26	Match to the train person no: 36	False
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 11	False
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 23	False
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 39	False
The test person no: 41	Match to the train person no: 45	False
The test person no: 42	Match to the train person no: 23	False
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 35	False
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 27	True
The test person no: 48	Match to the train person no: 45	True

Table 7.4: The recognition rate for the 2 seconds speech signal based recognition system was %58.3. (Part 2)

The test person no: 1	Match to the train person no: 1	True
The test person no: 2	Match to the train person no: 2	True
The test person no: 3	Match to the train person no: 3	True
The test person no: 4	Match to the train person no: 4	True
The test person no: 5	Match to the train person no: 5	True
The test person no: 6	Match to the train person no: 6	True
The test person no: 7	Match to the train person no: 7	True
The test person no: 8	Match to the train person no: 8	True
The test person no: 9	Match to the train person no: 9	True
The test person no: 10	Match to the train person no: 10	True
The test person no: 11	Match to the train person no: 11	True
The test person no: 12	Match to the train person no: 12	True
The test person no: 13	Match to the train person no: 13	True
The test person no: 14	Match to the train person no: 14	True
The test person no: 15	Match to the train person no: 15	True
The test person no: 16	Match to the train person no: 16	True
The test person no: 17	Match to the train person no: 17	True
The test person no: 18	Match to the train person no: 18	True
The test person no: 19	Match to the train person no: 19	True
The test person no: 20	Match to the train person no: 20	True
The test person no: 21	Match to the train person no: 21	True
The test person no: 22	Match to the train person no: 22	True
The test person no: 23	Match to the train person no: 23	True
The test person no: 24	Match to the train person no: 24	True

Table 7.5: The recognition rate for the 10 seconds speech signal based recognition system was %97.9. (Part 1)

The test person no: 25	Match to the train person no: 25	True
The test person no: 26	Match to the train person no: 26	True
The test person no: 27	Match to the train person no: 27	True
The test person no: 28	Match to the train person no: 28	True
The test person no: 29	Match to the train person no: 29	True
The test person no: 30	Match to the train person no: 30	True
The test person no: 31	Match to the train person no: 31	True
The test person no: 32	Match to the train person no: 32	True
The test person no: 33	Match to the train person no: 33	True
The test person no: 34	Match to the train person no: 34	True
The test person no: 35	Match to the train person no: 35	True
The test person no: 36	Match to the train person no: 36	True
The test person no: 37	Match to the train person no: 37	True
The test person no: 38	Match to the train person no: 38	True
The test person no: 39	Match to the train person no: 39	True
The test person no: 40	Match to the train person no: 40	True
The test person no: 41	Match to the train person no: 24	False
The test person no: 42	Match to the train person no: 42	True
The test person no: 43	Match to the train person no: 43	True
The test person no: 44	Match to the train person no: 44	True
The test person no: 45	Match to the train person no: 45	True
The test person no: 46	Match to the train person no: 46	True
The test person no: 47	Match to the train person no: 47	True
The test person no: 48	Match to the train person no: 48	True

Table 7.6: The recognition rate for the 10 seconds speech signal based recognition system was %97.9. (Part 2)

Chapter 8

Conclusion & Discussion

In this project, the multimodal recognition system was constituted. The main objective of the project is to increase the recognition rate with using two different biometric characteristics. One of these biometric characteristics is speech, which is commonly used in the today's technology for recognition and other one is ECG, which is more robust than other biometrics against the fraud by help of its low possibility of being imitated and also, it requires that the persons have to be alive in order to be used by recognition systems.

There are also traditional recognition systems based on the inputs, which are known or possessed by person. The most known inputs of traditional recognition systems are passwords, which require that person has to remember them, and cards, which require that person has to carry them. For the traditional recognition system, the inputs, which are known by person, have the high possibilities of being forgotten, lost and other inputs, which are possessed by person, have the high possibilities of being stolen or guessed by third persons. Because of these high possibilities, the biometric recognition systems are more reliable and useful than traditional recognition systems. However, there is weakness point of the biometric recognition systems, which is that the biometrics do not be regenerated in the case of fraud and accidents. At this point, the importance of multimodal recognition systems increases. Thanks to using more than one biometrics, the reliability and security reach to high level for the multimodal biometric recognition systems.

In this project, since there was not enough time and possibility, an artificial database was created with obtaining these biometric signals from various sources. The artificial database consists of 48 ECG and speech signals, which belong to 22 females and 26 males. After the constitution of the artificial database, the feature extraction operation was applied to biometric signals. In the feature extraction operation, MFCC method, which is commonly used in speech recognition systems, was used to obtain the features of speech signals and AC/DCT method was used to obtain the features of ECG signals. After the feature extraction operation, GMM method was used to classify these features. Before the fusion phase, results of GMM method were used in the decision phase to obtain recognition rate for ECG signal and speech signal, individually. The recognition rate of ECG signal was 87.5% and 42 persons were matched correctly. The recognition rate of 2 seconds speech signals was 58.33% and 28 persons were matched correctly. After the obtaining these recognition rates, normalization was applied to results of GMM method and then, these two type of biometrics of results were fused for the decision phase. Matching results were obtained in the decision phase. The recognition rate after fusion was 70.83% and 34 persons were matched correctly. However, when the recognition rates are considered, it has been observed that the recognition rate, which obtained after the fusion, is lower than recognition rate of the ECG signals. Hence, 10 seconds speech signals were used instead of 2 seconds speech signals and normalization methods were applied to the speech signals after applying the MFCC and GMM methods. The recognition rate of the 10 seconds speech signals is 97.9% and 47 persons were matched correctly. And then, the speech signals and ECG signals were fused again to obtain recognition rate from the matching results in the decision phase. The recognition rate of fusion reached to 95.8% and 46 persons were matched correctly. Consequently, fusion recognition rate approximate to unimodal recognition rates, which were obtained from ECG signals and speech signals, individually. Another normalization method and biometrics may be used to improve the fusion rate.

References

[1] "Internet archive." Internet Archive. [Online]. Available: http://archive.org/

[2] D. Gibbon, R. Moore, and R. Winski, in *Spoken Language System Assessment*, Berlin, New York, July 1998, pp. 96–97.

[3] M. Adolph, "Biometrics and standards," in *ITU-T Technology Watch Reports*. ITU-T Technology, 2009. [Online]. Available: http://www.itu.int/ITU-T/techwatch

[4] J. N. Pato and L. I. Millwtt, in *Biometric Recognition: Challenges and Opportunities*. Washington, DC, USA: IEEE Computer Society.

[5] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuts and Systems for Video Technology, Special Issue on Image and Video-Based Biometrics*, vol. 14, no. 1, 2004.

[6] H. Ma and A. Ming, "The biometric grid: A solution to biometric technologies," *IEEE Distributed Systems Online*, vol. 8, p. 1, 2007.

[7] "Sound." Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Sound

[8] "Speech production." [Online]. Available: https://kom.aau.dk/group/04gr742/pdf/speech_production.pdf

[9] "Speech recognition." Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Speech_recognition

[10] G. Kaur, G. Singh, and V. Kumar, "A review on biometric recognition," *Bio-Science and Bio-Technology*, vol. 6, no. 4, pp. 69–75, 2014.

[11] Z. Zhao, L. Yang, D. Chen, and Y. Luo, "A human ecg identification system based on ensemble empirical mode decomposition," 2013.

[12] Y. Wang, F. Agrafioti, D. Hatzinakos, Konstantionos, and N.Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1–11, 2008.

[13] A. Fratini, M. Sansone, P. Bifulco, and M. Cesarelli, "Individual identification via electrocardiogram analysis," *Biomed Eng Online*, 2015.

[14] H. Gürkan, "Compression of ecg signals using variable-length classified vector sets and wavelet transforms," *EURASIP Journal on Advances in Signal Processing, Springer (ISI)*, no. 2012:119, pp. 1–17, 2012.

[15] S. Safie, J. Soraghan, and L. Petropoulakis, "Electrocardiogram (ecg) biometric authentication using pulse active ratio (par)," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1315–1322, 2011.

[16] "12-lead ecg placement guide with illustrations." Cables and Sensors LLC [US]. [Online]. Available: http://www.cablesandsensors.com/pages/12-lead-ecg-placement-guide-with-illustrations

[17] "The ecg leads and lead theory." Clinical ECG Interpretation. [Online]. Available: https://ecgwaves.com/the-ecg-leads-and-lead-theory/

[18] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "A new approach in human identification," *IEEE Trans. on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.

[19] J. M. Irvine, B. K. Wiederhold, and L. W. Gavshon, "Heart rate variability: A new biometric for human identification," *Int. Conf. on Artificial Intelligence*, pp. 1106–1111, 2001.

[20] T. Shen, W. Tompkins, and Y. Hu, "One-lead ecg for identity verification," *The 2nd Joint Engineering in Medicine and Biology, 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society (EMB-S/BMES 02)*, p. 6263, 2002.

[21] S. Israel, W. Scruggs, W. Worek, and J. Irvine, "Fusing face and ecg for personal identification," *The 32nd Applied Imagery Pattern Recognition Workshop (AIPR03)*, p. 226231, 2003.

[22] S. A. Israel, J. M. Irvine, A. Chang, M. D. Wiederhold, and B. K. Wiederhold, "Ecg to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[23] T. Shen, "Biometric identity verification based on electrocardiogram (ecg)," *Ph.D. dissertation*, 2005.

[24] T. W. Shen, "Quartile discriminant measurement (qdm) method for ecg biometric feature selection," *International Symposium of Biomedical Engineering*, no. 13094, 2006.

[25] G. Wubbeler, M. Stavridis, D. Kreiseler, R. D. Bousseljot, and C.Elster, "Verification of humans using electrocardiogram," *Elsevier Pattern Recognition Letters*, vol. 28, pp. 1172–1175, 2007.

[26] C. Chuang-Chien, C. C.-M. C, and H. Chih-Yu, "A novel personal identity verification approach using a discrete wavelet transform of the ecg signal," *Int. Conf. on Multimedia and Ubiquitous Eng*, pp. 201–206, 2008.

[27] Irvine, J.M., Israel, S. A., Scruggs, W.T., and W. J. Worek, "Eigenpulse: Robust human identification from cardiovascular function," *Pattern Recognition*, vol. 41, no. 11, pp. 3427–3435, 2008.

[28] C. Fang and H. L. Chan, "Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space," *Elsevier Pattern Recognition*, vol. 42, no. 9, pp. 1824–1831, 2009.

[29] S. Fatemian and D. Hatzinakos, "A new ecg feature extractor for biometric recognition," *The 16th International Conference on Digital Signal Processing*, pp. 1–6, 2009.

[30] M. Li and S. Narayanan, "Robust ecg biometrics by fusing temporal and cepstral information," *Journal of Eng. and Comp. Innovations*, pp. 1326–1329, 2010.

[31] F. Sufi, I. Khalil, and I. Habib, "Polynomial distance measurement for ecg based biometric authentication," *Security and Communication Networks*, vol. 3, no. 4, pp. 303–319, 2010.

[32] J. Loong, K. Subar, R. Besar, and M. Abdullah, "A new approach to ecg biometric systems: A comparitive study between lpc and wpd systems," *World Academy of Science Engineering and Technology*, vol. 68, pp. 759–764, 2010.

[33] C. M. Ting and S. H. Salleh, "Ecg based personal identification using extended kalman filter," *10. Int. Conf. on Information Sciences Signal Processing and their Applications*, pp. 774–777, 2010.

[34] F. Sufi and I. Khalil, "Faster person identification using compressed ecg in time critical wireless telecardiology applications," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 282–293, 2011.

[35] T. W. Shen, W. J. Tompkins, and Y. H. Hu, "Implementation of a one-lead ecg human identification system on a normal population," *International Conference on Pattern Recognition*, vol. 2, no. 1, pp. 12–21, 2011.

[36] H. Gurkan, U. Guz, and B. Yarman, "A novel biometric authentication approach using electrocardiogram signals," *35th Annual International IEEE EMBS Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4259–4262, 2013.

[37] G. Gokhan, "Fingertip ecg signal based biometric recognition system," *Master of Science Thesis, Graduate School of Science and Engineering, (Advisor: Do. Dr. Hakan Gurkan, Co-Advisor:. Dr. Umit Guz)*, 2016.

[38] L. Muda, M. Begam, and I. Elamvazuthi, "Voice recognition algorithms using mel frequency cepstral coefficient and dynamic time warping techniques," *JOURNAL OF COMPUTING*, vol. 2, no. 3, pp. 138–143, 2010.

[39] A. V. Bhalla, S. Khaparkar, and M. R. Bhalla, "Performance improvement of speaker recognition system," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 3, pp. 138–143, 2012.

[40] P. P. Singh and P. Rani, "An approach to extract feature using mfcc," *IOSR Journal of Engineering*, vol. 4, no. 8, pp. 21–25, 2014.

[41] "Discrete fourier transform." Mathwork. [Online]. Available: http://mathworks.com/help/signal/ug/discrete-fourier-transform.html

[42] K. K. Paliwal, "Decorrelated and liftered filter bank energies for robust speech recognition."

[43] D.Yu and L. Deng, in *Automatic Speech Recognition.* Springer-Verlag London, 2015.

[44] D. Reynolds, "Gaussian mixture models," *MIT Lincoln Laboratory.*

[45] A. L. Goldberger, L. A. N. Amaral, and L. Glass, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

Curriculum Vitae

Ceren Dursun was born on 14 March 1991, in İstanbul. She received her BS degree in Electrical and Electronics Engineering from Işık University in 2014 with 3^{nd} ranking out of all graduates. She completed her M.S. degree in 2017 in Electronics Engineering from Işık University. She is currently working as a consultant at PwC, starting from 2016 to present.