EARPRINT: TRANSIENT EVOKEDOTOACOUSTIC EMISSION FOR B IOMETRICS

by

Yuxi Liu

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

© Copyright 2014 by Yuxi Liu
Abstract

Earprint: Transient Evoked Otoacoustic Emission for Biometrics

Yuxi Liu
Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto
2014

Biometrics is attracting increasing attention in privacy and security concerned issues. However, advanced forgery techniques are threatening the reliability of conventional biometric modalities. This has been motivating our investigation of a novel modality Transient Evoked Otoacoustic Emission (TEOAE), which is an acoustic response generated from cochlea. As a physiological outcome from auditory system, TEOAE is immune to replay and falsification attacks. In this work, we derive the time-frequency representation of the signal using wavelet analysis, revealing individual uniqueness. Discriminant analysis is subsequently utilized to reduce intra-subject variability. We also introduce a complete framework of the biometric system in both verification and identification modes, as well as a solution to secure sensitive template data. Comparative experiments on a biometric setting dataset show the merits of the proposed methods. Performance is further improved with fusion of information from both ears, with an average identification rate 99.44% and equal error rate 0.02% achieved.
Acknowledgements

First and foremost, I would like to sincerely thank Professor Dimitrios Hatzinakos, whose guidance and support made the completion of this work possible.

I also gratefully thank former and current members of Biometric Security Laboratory, especially Jiexin Gao and Sahar Javaher Haghighi for their inspiring interaction and encouragement.
## Contents

1 Introduction .......................................................... 1  
   1.1 Security and Privacy Motivations ................................. 1  
   1.2 Research Goals and Contributions ............................... 4  
   1.3 Thesis Organization ............................................. 7  

2 Transient Evoked Otoacoustic Emission .......................... 9  
   2.1 Brief Physiology of the Ear and Hearing ....................... 9  
   2.2 Otoacoustic Emission ........................................... 12  
      2.2.1 Spontaneous Otoacoustic emission ......................... 13  
      2.2.2 Distortion product otoacoustic emission .................. 13  
      2.2.3 Transient evoked otoacoustic emission ................... 14  
   2.3 TEOAE for Biometrics: Literature Review ..................... 18  

3 TEOAE Analysis for Biometrics .................................... 20  
   3.1 Time-frequency Representation .................................. 20  
   3.2 LDA-based Dimensionality Reduction ............................ 24  
   3.3 Biometric Matching ............................................. 29  
      3.3.1 Verification .................................................. 29  
      3.3.2 Identification .............................................. 30  
   3.4 Framework of TEOAE Biometric System ......................... 33  
   3.5 Application Scenario ........................................... 35
4 TEOAE for Biometrics Performance Evaluation

4.1 Dataset and Experimental Setting ............................................ 38
4.2 Single Ear Results ................................................................. 42
  4.2.1 Verification performance .................................................... 42
  4.2.2 Identification performance ................................................ 44
4.3 Fusion of Both Ears ............................................................... 45
  4.3.1 Fusion in Verification ....................................................... 45
  4.3.2 Fusion in Identification .................................................... 47

5 Template Protection with BioHashing ........................................ 49

5.1 Motivation for TEOAE Template Protection ............................... 49
5.2 Related Works ........................................................................ 50
5.3 BioHashing for TEOAE Biometrics .......................................... 51
  5.3.1 Overview of BioHashing .................................................... 51
  5.3.2 Detailed Analysis of Random Projection ............................... 52
  5.3.3 Non-invertible Quantization .............................................. 53
  5.3.4 Framework of Biometric System with BioHashing ................. 55
5.4 Performance Evaluation .......................................................... 57
  5.4.1 Experimental Results ....................................................... 57
  5.4.2 Security Analysis ............................................................ 58

6 Conclusions and Future Improvements ....................................... 61

Bibliography ................................................................. 64
List of Tables

1.1 Comparison of major biometric modalities as well as TEOAE. ........ 5

4.1 TEOAE recording protocol ............................................. 39

4.2 Comparison of various dataset properties. ............................ 41

4.3 Performance in verification of the proposed method under 5 empirical CWT scales, and a similar one without subsequent LDA, as well as that of PDF estimate proposed in [1, 2] and BEMD in [3]. .................. 44

4.4 Performance in identification of the proposed method under 5 empirical CWT scales, and a similar one without subsequent LDA, as well as that of BEMD proposed in [3]. .......................... 45

4.5 Verification performance of the proposed system based on different fusion strategies. ............................................. 46

4.6 Identification performance of the proposed system based on different fusion strategies. ............................................. 48

5.1 Performance of proposed method with BioHashing and previous one without BioHashing respectively based on various CWT scales. ............ 59
List of Figures

1.1 Illustration of TEOAE’s clinical application in hearing assessment: (a) testing probe placement, (b) measurement system with (c) Vivosonic Integrity V500 System [4], (d) a measurement system includes a laptop computer, "DP2000" (Starkey, USA) Otoacoustic Emission PCMCIA card, cable interface, foam probe ear tips, and noise protection headset [5]. . . . 3

2.1 Simplified diagram of human auditory system (edited from Fig. 1A in [6]). 10

2.2 Cross-sectional view of the cochlea (edited from Fig. 1 in [7]). . . . . . . 11

2.3 Using earphone with built-in microphone for OAE measurement. . . . . 12

2.4 Measurement of SOAE. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

2.5 Measurement of DPOAE. . . . . . . . . . . . . . . . . . . . . . . . . . . 14

2.6 Measurement of TEOAE. . . . . . . . . . . . . . . . . . . . . . . . . . . 15

2.7 Example of (a) the click stimulus, (b) a corresponding TEOAE recording. 16

2.8 Waveform and spectrogram of TEOAE data collected from the same sub-
ject (a) before one year, (b) after one year. [8] . . . . . . . . . . . . . . 17

3.1 The Mexican hat wavelet, in time domain: (a) mother wavelet, (c) scaled wavelet with scale a = 2, (d) scaled wavelet with scale a = 0.5; and in frequency domain: (b) mother wavelet, (d) scaled wavelet with scale a = 2, (f) scaled wavelet with scale a = 0.5. . . . . . . . . . . . . . . . . . . . . . . 22
3.2 Three different frequency components extracted from (a) original signal at scale (b) \(a = 4\), (c) \(a = 8\), (d) \(a = 16\) respectively. 

3.3 Frequency components extracted from signal of (a) subject A in the 1st and the 2nd session, (b) subject B in the 1st and the 2nd session, (c) subject A in the 1st and subject B in the 2nd session.

3.4 Block diagram of verification biometrics.

3.5 Block diagram of identification biometrics.

3.6 Framework of the proposed biometric system.

3.7 Architecture of the TEOAE biometric system for identity authentication in mobile commerce.

3.8 Architecture of the TEOAE biometric system for personalization in multimedia device.

4.1 TEOAE recordings from one collection (a) the first 10 recordings (first 200 ms), (b) the last 10 recordings (last 200 ms), when \(WWR > 90\%\).

4.2 Verification performance of left ear signal with CWT scale = 8.

5.1 Procedure of BioHashing.

5.2 Histogram of pairwise distances between genuine pairs (red) and those between imposter pairs (blue) in the following feature spaces: (a) Original feature, (b) Projection feature and (c) Projection feature after quantization.

5.3 Distribution of \(v\), approximately \(N(0,1)\).

5.4 Framework of the proposed biometric verification system with BioHashing.

5.5 Verification performance of left ear signal with CWT scale = 7.

5.6 Histogram of pairwise distances of the same users' hash templates across 100 systems.
List of Acronyms

**BEMD**  Bivariate Empirical Mode Decomposition. 18

**CWT**  Continuous Wavelet Transform. 20

**DPOAE**  Distortion Product Otoacoustic Emission. 13

**EER**  Equal Error Rate. 18

**FAR**  False Acceptance Rate. 42

**FRR**  False Rejection Rate. 42

**IHC**  Inner Hair Cell. 10

**LDA**  Linear Discriminant Analysis. 27

**MLE**  Maximum Likelihood Estimation. 18

**OAE**  Otoacoustic Emission. 12

**OHC**  Outer Hair Cell. 10

**PDF**  Probability Density Function. 18

**SOAE**  Spontaneous Otoacoustic Emission. 13
**SPL**  Sound Pressure Level.  14

**TEOAEN**  Transient Evoked Otoacoustic Emission.  2

**WWR**  Whole Wave Reproducibility.  39

x
Chapter 1

Introduction

1.1 Security and Privacy Motivations

Accurate and automatic identity recognition is becoming continuously essential in privacy and security concerned situations, such as access control and financial transactions. Traditional strategies for individual identification include entity-based (e.g. access card) and knowledge-based (e.g. PIN) methods. However, the reliability of such methods has been questioned concerning the risk of identity fraud or theft. Biometrics provides an efficient and reliable solution through physiological or biological characteristics of a person, for instance, iris, face, fingerprint and voice etc.

Nowadays biometrics systems are rapidly supplanting traditional identity management systems (e.g., passwords, access card), by utilizing the uniqueness of physiological, biological or behavioral traits of an individual. For example, e-passport embeds one’s face and/or iris characteristics to verify the identity; fingerprint recognition is widely used in access control, laptop lock and mobile device security.

Along with the pervasion of biometrics technologies, however, sophisticated forgery and spoofing techniques are threatening the reliability of conventional biometric modalities. For instance, fingerprints can be easily captured from a glass surface, phone screen
or anywhere one touched, and an artificial clone can be created by cheap and readily available plastic mold and gelatin [9]. Moreover, signature and voice recognition system become vulnerable against sophisticated falsification attacks. Such vulnerability has been motivating the recent investigation of a novel yet promising biometric modality Transient Evoked Otoacoustic Emission (TEOAE) [2, 3, 10], which is naturally immune to replay and falsification attacks as a physiological signal from human auditory system.

TEOAE refers to an approximate 20 ms acoustic response generated by an active process occurring inside the cochlea after a low level transient click stimulus [11]. TEOAE is present in most normal ears (99+%) [12] and widely used in clinical applications as a diagnostic tool for hearing assessment [13] (see the illustration in Figure 1.1). Importantly, a number of biological and clinical studies [12,14,15] have reported that TEOAE recordings show significant differences among individuals, due to different inner ear structure and many other genetic factors. Such uniqueness enables its role in identity recognition as “fingerprints” or “signature”.

Using TEOAE for biometrics has several advantages. Primarily, the inherent immune to falsification and replay attacks: it is more robust against falsified credential compared to traditional biometric modalities, as fabricating one’s auditory system is almost impossible. Also it is extremely difficult to steal one’s TEOAE which is the outcome of physiological activities of the auditory perception in response to a specific acoustic stimulation. This likewise enhances biometric reliability and effectiveness with implicit liveness detection.

Further, framework of TEOAE biometrics can be readily integrated into multimedia devices with headset or earphone (e.g. smart phone, media player) for user authentication, identification (individualization) and personal data security. TEOAE measurement system is simple, only an earphone or probe made up of a speaker and microphone is used to induce and record the acoustic response. And the cable interface and memory card (Figure 1.1 (d) for example), can be embedded into a multimedia device.
Figure 1.1: Illustration of TEOAE’s clinical application in hearing assessment: (a) testing probe placement, (b) measurement system with (c) Vivosonic Integrity V500 System [4], (d) a measurement system includes a laptop computer, ”DP2000” (Starkey, USA) Otoacoustic Emission PCMCIA card, cable interface, foam probe ear tips, and noise protection headset [5].
Last but not least, it can be employed to some special scenarios where face and fingerprint recognition fail to work, such as neonate identification. A comparison of major biometric modalities as well as TEOAE is provided in Table 1.1, based on the following five criteria [16]: performance, the attainable recognition accuracy, considering potential limitation including personal, operational and environmental factors that affect the accuracy; acceptability, the extent to which people are willing to accept this biometric technology; cost, the expense of employing this technology; and robustness, against circumvention and forgery techniques. Each modality has its perks and disadvantages and there is no such one that can be applicable to all scenarios. TEOAE can be a reliable even favorable alternative in some particular applications, for instance mobile commerce and personalization in multimedia device, which will be discussed in an upcoming section.

1.2 Research Goals and Contributions

This thesis comprehensively studies a relatively new biometric modality: the transient evoked otoacoustic emission. The primary research objective is to unveil the biometric characteristics of TEOAE and investigate a novel approach for reliable and robust identity recognition based on TEOAE. A set of contributions that have been made in research work are outlined as follows:

- We design an algorithm deriving both time and frequency characteristics of the original TEOAE signal. Previously proposed methods addressing TEOAE for biometrics mainly explored the signal in solely time domain. In this work, time-frequency analysis based on continuous wavelet transform is described, which preserves both time and frequency features of TEOAE signal. It also reveals the intra-subject stability and the inter-subject disparity that are obscure in original time series signals.

- We consider the time-variant nature of TEOAE as a physiological signal, and ap-
Table 1.1: Comparison of major biometric modalities as well as TEOAE.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Performance</th>
<th>Acceptability</th>
<th>Cost</th>
<th>Robustness</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>N/A</td>
</tr>
<tr>
<td>Face</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Lighting, makeup, age</td>
</tr>
<tr>
<td>Fingerprint</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Dirt, injury, inebriety</td>
</tr>
<tr>
<td>Gait</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
<td>Weight, lighting, age</td>
</tr>
<tr>
<td>Iris</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>Lighting, injury, age</td>
</tr>
<tr>
<td>Palmprint</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Inconsistency, fatigue</td>
</tr>
<tr>
<td>Signature</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Noise, injury, disease, injury</td>
</tr>
<tr>
<td>Voice</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>Noise, hearing impairment</td>
</tr>
<tr>
<td>TEOAE*</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

* Evaluated according to our work, as one of the pioneer studies on TEOAE for biometrics.
ply a discriminant subspace learning algorithm to further reduce the intra-subject variability prior to matching procedure. Unlike previous works which examined the performance in datasets of TEOAE recorded in the same session, we conduct experiments on a dataset in biometric setting, i.e., time interval between different recording sessions is at least one week. Besides, a supervised subspace learning method is performed in this work to select more effective features in a class-optimal way.

- We introduce a complete framework of the TEOAE biometric system in both verification and identification modes, considering practical application. Efficiency of the proposed system is also tested in both modes.

- With promising performance based on data from single ear, we further improve the accuracy and system robustness by integrating information of both ears. As TEOAE emitted from the left and the right ear are not identical, engaging binaural information reduces the probability of false authentication and that of miss classification.

- We additionally present a data cryptography method to secure TEOAE template data. Due to TEOAE’s natural immune to replay and falsification attacks, corresponding biometric system where sensitive template data are stored becomes the main intrusion target. To tackle this issue, feature sets are projected onto random subspaces and subjected to binary quantization, which ensures non-linkability of hash templates across various systems and prevents recovery of original biometric data from the distorted templates.

- Last but not least, we examined the proposed methods on our biometric setting TEOAE database [17], which to our best knowledge is one of very few used for biometric purpose. Data collection was conducted at the Biometrics Security Laboratory, University of Toronto, under protocol reference #23018.
1.3 Thesis Organization

The remainder of this thesis is structured as follows:

Chapter 2 describes the general background of otoacoustic emission following a brief physiology of the ear and hearing. It emphasizes on fundamental information about TEOAE and the feasibilities and perks of using TEOAE for biometrics. Besides, a literature review of related works on this issue is provided in the same chapter.

Chapter 3 presents detailed methodology of biometrics based on TEOAE along with a complete system framework considering practical application. It is followed by Chapter 4 where experimental results and analysis on both verification and identification modes are reported.

Chapter 5 introduces a Biohashing approach to secure TEOAE template data. Further, its effectiveness and robustness is justified through experiments and security analysis on attack scenarios.

Finally, Chapter 6 concludes the work presented in this thesis and outlines the directions for future developments.

Contents of this Thesis are partially presented or submitted in the following referred journal and conference publications:


• Y. Liu and D. Hatzinakos, Human Acoustic Fingerprints: A Novel Biometric Modality for Mobile Security, in Proceedings of the IEEE International Confer-

Chapter 2

Transient Evoked Otoacoustic Emission

2.1 Brief Physiology of the Ear and Hearing

Hearing is one of the most vital human senses which enables communication and exploration of the world around us through sound. To fulfill this function, the ear, or specifically the human auditory system, converts sound vibration into nerve impulses for subsequent perception and interpretation as sound by the brain. The human ear serves as a biological transducer detecting the frequencies and amplitude of sound waves through a complicated process, which allows us to discern the pitch, loudness and timbre of sounds [20].

The human ear, which mainly consists of outer ear, middle ear and inner ear (see Figure 2.1), is a fully developed part of our bodies at birth. The outer ear collects incoming sounds and channels them along a tube called the auditory canal towards the middle ear. At the end of the canal, also the beginning of middle ear, tympanic membrane (commonly known as eardrum) vibrates once the incoming acoustic pressure waves strike it. Movements of the tympanic membrane also set three bones that it connects to, the
malleus, incus and stapes, into motion at the same frequency of the sound wave [21]. From stapes, vibrations are transported through the fluid of the inner ear and arrive at cochlea.

![Simplified diagram of human auditory system](image)

Figure 2.1: Simplified diagram of human auditory system (edited from Fig. 1A in [6]).

Mechanical energy (i.e., vibration) from sounds is translated into electrical signals in the inner ear. Specifically, inside the cochlea (see the cross-sectional view in Figure 2.2), basilar membrane as the organ of hearing responds to different frequency in a location-specific way, where higher frequencies are responded near the basal end and lower frequencies are responded near the apical end. Resting on the basilar membrane, there are two types of cells of particular importance to auditory system, the inner hair cells (IHCs) and the outer hair cells (OHCs). Activation of the IHCs causes firing of auditory nerves, which transduce the sound vibrations into electrical pulses that are then relayed to the auditory cortex and interpreted as “sound” by the brain. The IHCs act as conveyers of sound level (intensity) information to the brain, while the OHCs are responsible for enhancing sound frequency selectivity of human auditory system, as reviewed in the next section.
Figure 2.2: Cross-sectional view of the cochlea (edited from Fig. 1 in [7]).
2.2 Otoacoustic Emission

In mammals, in order to enhance hearing sensitivity and frequency selectivity (i.e., discrimination of hearing), some OHCs resonate with a larger amplitude of vibration when the frequency of the sound wave matches the natural frequency of those OHCs. The stimulus-specific vibration acts as energy sources for the forward amplification of sound and increases mechanical energy within the cochlea [12].

The energy generated by the stimulus-related oscillations of OHCs, meanwhile, echoes back to middle ear from the cochlea. It then vibrates the tympanic membrane, which produces acoustic response propagating backwards into the ear canal. The residual energy arriving at the ear canal is generally lower than 15 dB. Such nearly inaudible response is the otoacoustic emission (OAE), which was first discovered by David Kemp [22] in 1977.

Since OAE travels back to ear canal, it can be easily measured with an earphone with a built-in sensitive microphone or a small disposable probe inserted into the outer part of the ear canal (Figure 2.3). The earphone or the miniature speaker within the probe broadcasts the sound stimulus/stimuli; simultaneously, the assembled miniature microphone records the OAE triggered by the aforementioned active process in the cochlea.

Figure 2.3: Using earphone with built-in microphone for OAE measurement.
There are, in fact, different types of OAE that are determined by the sound stimulus used to generate the response. They are summarized into three categories:

### 2.2.1 Spontaneous Otoacoustic emission

Spontaneous Otoacoustic Emission (SOAE) is low-level sound continuously emitted without any presentation of an external acoustic stimulus. A microphone is solely needed to measure SOAE (Figure 2.4). However, SOAEs occur in only about 60% of ears with normal hearing [12], and thus it is not prevalent in clinical use.

![Figure 2.4: Measurement of SOAE.](image)

### 2.2.2 Distortion product otoacoustic emission

Distortion Product Otoacoustic Emission (DPOAE) is an evoked OAE as a result of intermodulation distortion in response to two simultaneously presented pure-tone stimuli of different frequencies [13]. The prevalence of DPOAE reaches 99+% of normal hearing ears. However, DPOAE measurement system has high hardware requirements, including a probe with two speakers generating stimulus of different frequencies and a microphone associated with an analog-to-digital converter (ADC) to distinguish the response from the sinusoidal waveform stimuli (Figure 2.5).
2.2.3 Transient evoked otoacoustic emission

Transient evoked otoacoustic emission (TEOAE) is an acoustic response stimulated by a short low level click sound, which is a flat-band signal in frequency domain. The shorter the stimulus in duration, the broader its frequency range. Therefore in practice, the stimulus is approximately 100 $\mu$s in duration [12] and has a frequency range from 500 to 6000 Hz. The peak intensity of the stimulus is typically set around 80 dB sound pressure level (SPL). TEOAE emerges shortly (nearly 2 to 3 ms) after the click stimulus diminishes and lasts approximately 20 ms. Overall, TEOAE measurement system is much simpler than that of DPOAE, only an earphone or probe made up of one single speaker and microphone is used to induce and record the acoustic response (Figure 2.6). As the stimulus vanishes before the TEOAE arises, no extra converter is required to distinguish the TEOAE from the collected signal.

An example of the stimulus and TEOAE recording from the database collected in Biometrics Security Laboratory at the University of Toronto [17] is displayed in Figure 2.7 (properties of the database will be described in detail in Section 4.1).

In human cochlea, a tone-burst stimulus can excite the basilar membrane and evoke a frequency-specific response; when a broadband click stimulus is used, the evoked emission becomes a cumulative response consisting of several dominant frequency components with different latency [23]. This makes TEOAE highly non-stationary with a complex...
waveform compare to other OAE signals.

Satisfaction of the following essential requirements for biometric traits [16] qualifies TEOAE to be a feasible biometric modality:

- **Universality**: ideally each member of the population should have the biometric trait. The prevalence of TEOAE in normal ears is close to, slightly less than 100% [12].

- **Collectability**: the biometric trait should be easy to acquire without discomfort to users, and the biometric data should be measured quantitatively. As aforementioned, TEOAE, as a simple and non-invasive tool for hearing assessment, is prevalently used in universal infant hearing screening programs. Placing the probe for TEOAE measurement is similar to wearing an earphone. Besides, data collected are simply one dimensional acoustic signal, which can be effortlessly quantized.

- **Uniqueness**: the distinctive ability of the biometric trait should be sufficient so as to separate different individuals. A number of biological and clinical studies [12,14,15] have reported that TEOAE recordings show significant differences among individuals, due to different inner ear structure and many other genetic factors. Interpersonal differences of TEOAE signal suggest an individual morphology of inner ears especially OHCs.

- **Stability**: the characteristics or features of the biometric trait should remain moder-
Figure 2.7: Example of (a) the click stimulus, (b) a corresponding TEOAE recording.
ately constant (regarding a particular matching criterion) over a reasonable period of time. The stability of TEOAE is proved by the high reproducibility both on short-term and on long-term progression, which is discovered in clinical tests and biological studies [24,25].

Take the study [8] by Konopka and Grzanka for example, where 152 healthy subjects participated in ear examinations twice in a one-year interval, high similarity is observed in the spectrogram of data collected from the same subject before and after one year. As shown in two spectrograms in Figure 2.8 (a) and (b), a star-shaped object centres at frequency around 3 kHz and a few crossing lines correspond to the dominant frequency components. Regardless of the changes in TEOAE waveform morphology, which is probably associated with the proportions of particular components, configuration of TEOAE in time and frequency coordinates remains stable over a long time period [8].

![Figure 2.8](image_url)

Figure 2.8: Waveform and spectrogram of TEOAE data collected from the same subject (a) before one year, (b) after one year. [8]


2.3 TEOAE for Biometrics: Literature Review

The idea of using TEOAE for biometrics is relatively innovative despite of its prevalence in clinical application. To our best knowledge, TEOAE is one of very few biometric traits triggered by stimuli. Following are the existing prior studies:

In 2004, a primary study of TEOAE for biometrics was introduced by Swabey et al. [26] by analyzing the Euclidean distance of the frequency components in the power spectra of the signal. After that, a quantitative study was conducted in [1] to approximate the probability density function (PDF) of the intra- and inter-class distance distribution-s by maximum likelihood estimation (MLE) to the time series data. In [27], cepstrum transformation was applied for feature extraction and pairwise distance was calculated in both temporal and cepstral domain. This has provided a potential route to better extract TEOAE biometric features by applying time-frequency analysis. For the reason that TEOAE is a highly non-stationary cumulative signal with different frequency components, analyzing TEOAE in only time domain or frequency domain is suboptimal, where features of the other domain are neglected.

Recently, Grabham et al. [2] summarized the above works [1, 26], examined biometric verification performance in 4 datasets of TEOAE recorded in the same session respectively (either in one sitting or with the ear probe removed and refitted shortly), which was of little biometric value without considering the high intra-subject variability in separate recording sessions. The reported equal error rate (EER) of the first 3 datasets was 2.30% (measured), 2.29% (predicted) and 4.58% (predicted) respectively.

TEOAE Dataset in biometric setting was first investigated in [3], where Bivariate Empirical Mode Decomposition (BEMD) was applied to decompose a signal into multi-level local oscillation components with selected frequencies. Correlations across different levels were then calculated to derive the matching scores. Empirical weights for score computation were chosen based on limited testing data (only one TEOAE from 54 subjects) without statistical analysis, which failed to guarantee generalizability and applicability
Motivated by these studies, in this thesis we propose a wavelet based approach to derive the time-frequency representation of the signal, which reserves features of both time and frequency domains. Besides, the proposed work has two major considerations: first, as slight variance in spectrograms of TEOAE from the same subject is observed (see Figure 2.8 (b)), we apply discriminant analysis to reduce the intra-subject variability prior to matching procedure; second, in our TEOAE dataset, time interval between different recording sessions was at least one week so as to verify the long-term stability for biometric evaluation purpose. Last but not least, we introduce a complete framework of TEOAE biometric system and a solution to secure sensitive template data stored in the system, which further strengthens the comprehensiveness of this work on TEOAE for biometrics.
Chapter 3

TEOAE Analysis for Biometrics

In this work, we derive the time-frequency representation of the original signal by adopting wavelet analysis to reveal individual uniqueness and long-term reproducibility. Considering the time-variant nature of such physiological signal, we apply discriminant analysis to further reduce the intra-subject variability prior to matching procedure. Lastly, we introduce a complete framework of the biometric system in both verification and identification modes. Specifically, similarity score is measured in Pearson correlation in the first mode; best matching identity is attained by multiclass classifier softmax regression in the latter one.

3.1 Time-frequency Representation

TEOAE is a cumulative response due to the fact that the broadband stimuli can excite the entire basilar membrane and different resulting frequency components arrive at the ear canal at different times. Hence, investigating TEOAE signal in either time domain or frequency domain is suboptimal. Analysis through a time-frequency approach is therefore advisable based on its close relation with cochlear mechanisms [28].

Among time-frequency analysis tools, Continuous Wavelet Transform (CWT) [29] is the more ideal candidate than Discrete Wavelet Transform for time series signal by
Chapter 3. TEOAE Analysis for Biometrics

virtue of its more fine-grained resolution without skipping samples. Therefore in this work, CWT is applied to separate the mixed frequency components and derive the time-frequency representation of TEOAE. Regarding mother wavelet $\psi(t)$, which is a narrow band-pass function, the wavelet transform to a signal $x(t)$ is given by:

$$WT_x(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*(\frac{t-b}{a})dt$$  \hspace{1cm} (3.1)

where $\psi(\frac{t-b}{a})$ is the corresponding window function with scale factor $a$ and time translation factor $b$. Let $X(\omega)$ and $\Psi(\omega)$ be the Fourier transform of $x(t)$ and $\psi(t)$ respectively. The equivalent frequency-domain version of wavelet transform is:

$$WT_x(a, b) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega)\Psi^*(a\omega)e^{+j\omega b}d\omega$$  \hspace{1cm} (3.2)

The adjustable window function is introduced so that the mixed frequency components in TEOAE can be separated. More precisely, if the narrow band-pass mother wavelet $\Psi(\omega)$ is centered around frequency $f_0$, according to Fourier transform theorems, $\Psi(a\omega)$ is the same function centered at $f_c = f_0/a$ with constant quality factor. Take Mexican hat wavelet for example, its mother wavelet in time and frequency domain are presented in Figure 3.1 (a) and (b) respectively. The central frequency $f_c$ is proportional to the scale $a$, as displayed in Figure 3.1 (d) and (f), where the central frequency is half of that of the mother wavelet with $a = 2$, and twice of that of the mother wavelet with $a = 0.5$ respectively.

Therefore, different frequency components can be extracted from the TEOAE original signal $x(t)$ by the narrow band-pass wavelet function with different scales. As Figure 3.2 displays, three frequency components are extracted from the original TEOAE signal at scale $= 4, 8, 16$ respectively with Daubechies 5 as mother wavelet. A higher frequency fraction corresponds to a smaller scale and vice versa.

Further, variation of scale alters not only the central frequency $f_c$ of the band-pass
Figure 3.1: The Mexican hat wavelet, in time domain: (a) mother wavelet, (c) scaled wavelet with scale $a = 2$, (d) scaled wavelet with scale $a = 0.5$; and in frequency domain: (b) mother wavelet, (d) scaled wavelet with scale $a = 2$, (f) scaled wavelet with scale $a = 0.5$. 
Figure 3.2: Three different frequency components extracted from (a) original signal at scale (b) $a = 4$, (c) $a = 8$, (d) $a = 16$ respectively.
function, but also the window length in time. Specifically, the time-duration is inversely proportional to $f_c$ (see Figure 3.1 (c)(d) and (e)(f)). By virtue of this particular structure of the wavelet filters, wavelet transform can provide highly accurate extraction of the time-frequency features from a mixed signal with high frequency components of brief duration and low frequency ones of long duration, TEOAE is such a case [28, 30]. As depicted in Figure 3.2, the discriminated frequency components show a clear difference in time duration, i.e., a longer latency for a lower frequency component (higher scale). CWT preserves good resolution in both time and frequency domain and thus is a suited approach for feature extraction of TEOAE.

A particular frequency component extracted is then used to represent the original signal if it can well differentiate (“identify”) itself from those of other sources. Take two subjects in the dataset [17] for example, frequency components at scale = 8 are extracted from both subjects per different recording session respectively. Extracted components from the same subject (A or B) exhibit high similarity between different sessions (Figure 3.3 (a) and (b)) although their original signals look disparate. However, extracted components from different subjects (A and B) highly diverge despite the close waveform of their original signals (Figure 3.3 (c)). Time-frequency representation reveals the intra-subject stability or reproducibility and the inter-subject disparity that are obscure in original signals.

### 3.2 LDA-based Dimensionality Reduction

The resulting TEOAE time-frequency representation vectors can be directly used for matching. However, unlike most conventional biometric modalities, e.g., fingerprints and iris, TEOAE as a physiological signal is time-variant over a long time period with observed changes in the waveform morphology, and so is the corresponding time-frequency feature derived from it (see Figure 3.3 (a) for example). Due to this dynamic nature,
Chapter 3. TEOAE Analysis for Biometrics

(a)

(b)
Figure 3.3: Frequency components extracted from signal of (a) subject A in the 1st and the 2nd session, (b) subject B in the 1st and the 2nd session, (c) subject A in the 1st and subject B in the 2nd session.
it is important to further diminish the intra-subject variability in order to reduce false rejection. Besides, the dimensionality of the time-frequency representation vector is as same as that of the original signal (see Figure 3.3). Dimensionality reduction is therefore prudent to avoid possible redundance and irrelevance of the features, and additionally to save storage space for template data in a biometric system. For these reasons, a supervised subspace learning method linear discriminant analysis (LDA) [31, 32] is performed in this work to select more effective features in a class-optimal way prior to matching procedure.

The objective of LDA is to preserve the maximum amount of class separability information with the minimum-dimensional data after linear projection. Details of the implementation are herein discussed, beginning with the following definitions:

- Suppose we have a training set consisting of $K$ classes, i.e., $K$ subjects registered in the biometric system.
- Let $N_k$ be the number of time-frequency representation vectors from class $k$, where $k = 1...K$.
- $z_{ki}$ is defined as a representation vector (observation, sample) from class $k$, where $i = 1...N_k$.
- The training feature set comprises $Z_1, Z_2, ..., Z_K$, where $Z_k = \{z_{k1}, z_{k2}, ..., z_{kN_k}\}$ is the group of observations from class $k$.
- Let $\mu_k$ be the class mean vector for class $k$, i.e., $\mu_k = \frac{1}{N_k} \sum_{i=1}^{N_k} z_{ki}$.
- Further, $\mu$ is the overall mean vector.

The feature extraction procedure of LDA is a linear transformation which maps the high-dimensional data onto a $m$-dimensional ($m = K - 1$) space where between-class distance is maximized and within-class distance is minimized, so that observations are
best separated into classes. Two types of distance are respectively measured in between-class and within-class scatter matrices $S_b$ and $S_w$, which are computed as below:

$$S_b = \sum_{k=1}^{K} N_k (\mu_k - \mu)(\mu_k - \mu)^T \quad (3.3)$$

$$S_w = \sum_{k=1}^{K} \sum_{i=1}^{N_k} (z_{ki} - \mu_k)(z_{ki} - \mu_k)^T \quad (3.4)$$

Accordingly, the projection weight $W$ should maximize the Fisher criterion function $J(W)$:

$$W = \arg \max_W J(W)$$

$$= \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|} \quad (3.5)$$

which guarantees largest segregation between observations from different classes and smallest variance within those from the same classes.

It is noted that the objective function $J(W)$ is invariant to scaling of $W$, hence maximizing $J(W)$ is equivalent to the following constrained optimization problem:

$$\max \quad W^T S_b W$$

$$\text{s.t.} \quad W^T S_w W = 1 \quad (3.6)$$

The corresponding Lagrange multiplier can be expressed as:

$$L(W, \lambda) = W^T S_b W - \lambda(W^T S_w W - 1) \quad (3.7)$$

According to Karush-Kuhn-Tucker (KKT) conditions (i.e., $\frac{L(W, \lambda)}{\partial W} = 0$), we obtain

$$S_b W = \lambda S_w W$$

$$\Rightarrow \quad S_w^{-1} S_b W = \lambda W \quad (3.8)$$
which is a generalized eigenvalue problem and the optimum \( W^* \) can be obtained by computing the \( m \) most significant eigenvectors which correspond to the \( m \) largest eigenvalues of \( S_w^{-1}S_b \). The projected feature vector \( \tilde{z}_{ki} \) \((\tilde{z}_{ki} \in \mathbb{R}^m)\) for a training sample \( z_{ki} \) is consequently computed by:

\[
\tilde{z}_{ki} = W^*Tz_{ki}
\] (3.9)

and is stored in the system gallery as a reference template. \( W^* \) is also saved in the system and will be reapplied to a new input feature vector \( z' \) so that it is projected onto the optimal and comparable subspace, i.e., \( \tilde{z}' = W^*Tz' \).

### 3.3 Biometric Matching

After LDA-based feature extraction from time-frequency representations, the next procedure is biometric matching which compares the probe against the templates stored in system gallery. Depending on a specific application, there are two modes of matching to contemplate, verification and identification. The matching module of a biometric system either validates a user’s claimed identity (in verification mode) or establishes an unknown user’s identity (in identification mode).

#### 3.3.1 Verification

Verification is basically a one-to-one matching, comparing the input sample presented by a user against the template corresponding to the claimed identity which is stored in the system. The decision of match or non-match is finally made based on the similarity score and a pre-defined threshold. Briefly, biometric verification aims to answer the question “Is this user the one he/she claims to be?” (as depicted in Figure 3.4).

In this work, Pearson’s distance is chosen as a metric, which is defined from the Pearson correlation coefficient as a common similarity measure in biometrics \([33, 34]\). The similarity score \( S_P \) between a probe vector \( z' \) and the reference template \( z_k \) (here
for convenience, we also write $z$ to denote the optimal features after LDA, instead of $\tilde{z}$ from the claimed identity $k$ is calculated by:

$$S_P(z', z_k) = 1 - \frac{\text{cov}(z', z_k)}{\sqrt{\text{cov}(z', z') \text{cov}(z_k, z_k)}}$$  \hspace{1cm} (3.10)$$

where $\text{cov}(a, b)$ is the covariance matrix of $a, b$. An ACCEPT decision is made only if $S_P(z', z_k) \leq t_P$, where $t_P$ is the corresponding decision threshold.

### 3.3.2 Identification

In identification mode, the system checks the probe presented by an unknown individual against the templates in the dataset from all enrollees in attempt to establish his/her identity. Biometric identification is basically a one-to-many comparison process, which is equivalent to multiclass classification. It seeks to answer the question “What is the identity of this person?” (as depicted in Figure 3.5).

An efficient and accurate multiclass classifier softmax regression, also known as multinomial logistic regression [35, 36], is employed in this work, which analyses the configuration of feature vectors from different classes (subjects) and provides probabilistic prediction results. It generalizes binary logistic regression by allowing the model to learn
and predict more than two categorical values. Similarly, it uses maximum likelihood estimation (MLE) to evaluate the probabilities of different possible outcomes.

Given a training set \( \{(z^{(1)}, y^{(1)}), (z^{(2)}, y^{(2)}), \ldots, (z^{(M)}, y^{(M)})\} \), where \( z^{(i)} \) is an input feature vector consisting of the optimal feature vector after dimensionality reduction and an appended dimension for the intercept term, i.e., \( z^{(i)} \in \mathbb{R}^{m+1} \); \( y^{(i)} \in \{1, 2, \ldots, K\} \) is the corresponding class label; and \( M \) is the number of training samples. Let \( \theta = \{\theta_1, \theta_2, \ldots, \theta_K\} \) denote the parameters of the regression model to be estimated, where \( \theta_k \in \mathbb{R}^{m+1} \) represents the coefficients of class \( k \) for each input feature vector. The hypothesis that \( z^{(i)} \) is classified as class \( j \) is written as:

\[
P(y^{(i)} = j | z^{(i)}, \theta) = \frac{\exp(\theta_j^T z^{(i)})}{\sum_{k=1}^{K} \exp(\theta_k^T z^{(i)})}
\]

(3.11)

Also note that \( 1/ \sum_{k=1}^{K} \exp(\theta_k^T z_i) \) is the normalization factor so that

\[
\sum_{j=1}^{K} P(y^{(i)} = j | z^{(i)}, \theta) = 1
\]

(3.12)

for any \( i \).

Model parameters are then learned from the training data by MLE. The log-likelihood
The well-trained model can be used to classify an unlabeled input $z'$. The hypothesis function $h_\theta(z')$ will return a $K$-dimension vector with the estimated probabilities of each class. The prior acts as a regularizer on $\theta$ and MLE in this way can be viewed as a penalized version or maximum a posteriori (MAP) estimation. The estimate in (3.14) is typically accomplished by an iterative optimization method such as gradient descent or Newton's method (e.g., L-BFGS [37] as the one we employ).
class:

$$h_{\hat{\theta}}(z') = \begin{bmatrix} P(y' = 1|z', \hat{\theta}) \\ P(y' = 2|z', \hat{\theta}) \\ \vdots \\ P(y' = K|z', \hat{\theta}) \end{bmatrix}$$

$$= \frac{1}{\sum_{k=1}^{K} \exp(\hat{\theta}_k^T z')} \begin{bmatrix} \exp(\hat{\theta}_1^T z') \\ \exp(\hat{\theta}_2^T z') \\ \vdots \\ \exp(\hat{\theta}_K^T z') \end{bmatrix}$$

(3.16)

The largest element indicates the most likely class that the input vector is labeled.

### 3.4 Framework of TEOAE Biometric System

Following all aforementioned modules, a complete framework of biometric system is logically proposed in this section. A general setup for the biometric system involves an enrollment stage and a verification and/or identification stage. A block diagram of the proposed system is depicted in Figure 3.6 and detailed explanation is provided as follows.

During the enrollment session, TEOAE is acquired from each enrollee and subjected to CWT-based time-frequency analysis at a particular scale. Resulting representation vectors are then used to drive the LDA-based optimal subspace learning algorithm. The fully trained transformation rule with weight $W^*$, and resulting feature sets are stored in the system database (see the dash line in Figure 3.6). Notably for identification system, feature templates are further utilized to force the learning of softmax regression model. Optimized classifier with parameter $\hat{\theta}$ is included in the system database used for
Chapter 3. TEOAE Analysis for Biometrics

Figure 3.6: Framework of the proposed biometric system.
future prediction. In a word, the enrollment operation establishes the template pool and creates a paradigm of individual TEOAE morphologies, acting as a training procedure from machine learning perspective.

In the verification mode, an individual contributes a TEOAE recording and claims an identity $k$. Probe feature vector will be generated after CWT and linear transformation according to the pre-trained rule. It is then matched against the respective template of subject $k$ retrieved from the gallery. And decision of either accept or reject the claim is finally made.

As for the identification operation, an individual from the pre-enrolled subject pool gets recorded. The biometric system will similarly conduct feature extraction and selection to a new reading, and compute the hypothesis function using the pre-determined classifier with the probe vector as an input. The user’s identity will be consequently established after a one-to-many search.

### 3.5 Application Scenario

Nowadays, mobile devices have outgrown their initial use for communication and have added many powerful functionalities, such as data storage, web access and remote commerce, etc. The increasing need for greater mobile security has been opening up new areas in biometrics. Hence, the need for greater security for mobile devices and more robust protection against unauthorized use has been rapidly increasing.

As aforementioned, TEOAE is naturally immune to falsification or replay attacks as an outcome of physiological activity of the auditory perception. Besides, its compatibility with mobile devices and robustness make it an advantageous modality for mobile security. As an example, Figure 3.7 exhibits how the proposed system is applied to mobile commerce (e.g., remote banking, phone purchase). To fulfill registration for the secure system, a client’s phone is responsible for TEOAE data acquisition and transmission to
the remote server in the administrator’s side, where data processing and learning algorithms are afterwards implemented. When a user (who is either the authorized client or an illegitimate one) is signing in, the server receives data sent from user’s phone and performs matching after computation. It finally responds to both the administrator and the user whether the authentication is accepted or declined, whether follow-up transaction is allowed or refused. It is expected that security questions whose answers can be easily discovered (e.g., “What is your mother’s maiden name?”, “What is your date of birth?”, “where did you go to college?”) will no longer be asked.

Figure 3.7: Architecture of the TEOAE biometric system for identity authentication in mobile commerce.

Further, local application (e.g., user verification and identification, private data access) of the proposed biometric system share the same architecture, except that datasets are stored and computation is conducted locally or on a client-server with which the device needs to communicate. For instance, a multimedia device (e.g., ipod, MP4) can be used by multiusers. Each user has his/her own profile, including history, subscription channel and favourite play list. Utilizing TEOAE biometrics, automatic personalization can be realized. A user registers with the system when wearing the earphone at the first time. In later use, the user will be recognized by the system and automatically presented
his/her preferred list without manually selecting. The conceptual framework is displayed in Figure 3.8.

Figure 3.8: Architecture of the TEOAE biometric system for personalization in multimedia device.
Chapter 4

TEOAE for Biometrics Performance Evaluation

This chapter presents the experiments and corresponding results of the proposed method for TEOAE based biometric system. Following an introduction of the TEOAE dataset [17] and experimental setting, comparative experiments are first conducted in single ear scenario, using signals collected from either left or right ear. Results show the merit of the proposed method by comparing with previous approaches. Experiments next focus on binaural scenario in order to further improve the system performance by fusing results from both ears. Moreover, performance of proposed method is herein evaluated by cross-validation. It meanwhile evaluates how accurately our proposed system will perform.

4.1 Dataset and Experimental Setting

In our experiments, we make use of the biometric setting dataset with 54 subjects collected at the Biometrics Security Laboratory [17], at the University of Toronto (protocol reference #23018). Vivosonic Integrity System was used for data collection with recording protocol [3] listed in Table 4.1. To alleviate constraints on the environment, the collection was conducted in a regular office with people conversing and ambulating, instead of a
Table 4.1: TEOAE recording protocol

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Stimulus</th>
<th>STI-Mode</th>
<th>Nonlinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click Interval</td>
<td>STI-Mode</td>
<td>Click Interval</td>
<td>21.12 ms</td>
</tr>
<tr>
<td>Click Duration</td>
<td>Parameters</td>
<td>Click Duration</td>
<td>80 µs</td>
</tr>
<tr>
<td>Sound Level</td>
<td></td>
<td>Sound Level</td>
<td>80 dB peSPL</td>
</tr>
<tr>
<td>Record Window</td>
<td>Test</td>
<td>Record Window</td>
<td>20 ms</td>
</tr>
<tr>
<td>Low Pass Cut-off</td>
<td>Control</td>
<td>Low Pass Cut-off</td>
<td>6000 Hz</td>
</tr>
<tr>
<td>High Pass Cut-off</td>
<td></td>
<td>High Pass Cut-off</td>
<td>750 Hz</td>
</tr>
<tr>
<td>Artifact Rejection</td>
<td></td>
<td>Artifact Rejection</td>
<td>55 dB SPL</td>
</tr>
<tr>
<td>Threshold</td>
<td></td>
<td>Threshold</td>
<td></td>
</tr>
</tbody>
</table>

sound proof clinical setting.

For each subject, TEOAE signals were collected from the left and right ear respectively in each of two sessions. Time interval between different sessions was at least one week so as to verify the long-term stability for biometric evaluation purpose, which previous datasets in [2] nonetheless fail to achieve with ear probe removed and reinserted in single sitting (see Table 4.2 for the comparison of various datasets).

In each collection, number of recordings or length of collection time depends on how fast the response steadies, which varies by person, body condition and body movement, etc. Collection is stopped when the whole wave reproducibility (WWR) measure displayed on the Integrity system is high enough (≥ 90%), which indicates high steadiness of response. For instance, in one collection, TEOAE recording is gradually stabilizing, as shown by the first 10 recordings (i.e., first 200 ms) plotted in Figure 4.1 (a) and the last 10 recordings (i.e., last 200 ms) plotted in Figure 4.1 (b) when the steadiness indicator WWR > 90%.

Current dataset consists of in average 122 TEOAE recordings for the left ear and 113 recordings for the right ear per person per session. Last 10 samples/segments of the first and second session per ear per subject, which are in general the 10 steadiest ones, are used to constitute the experimental training and testing set respectively. Daubechies 5 is again chosen as the mother wavelet of CWT.
Figure 4.1: TEOAE recordings from one collection (a) the first 10 recordings (first 200 ms), (b) the last 10 recordings (last 200 ms), when $WWR > 90\%$.
Table 4.2: Comparison of various dataset properties.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>23</td>
<td>760</td>
<td>561</td>
<td>33</td>
<td>54</td>
</tr>
<tr>
<td>Stimulus sound level</td>
<td>80dB peSPL</td>
<td>70dB peSPL</td>
<td>70dB peSPL</td>
<td>80dB peSPL</td>
<td>80dB peSPL</td>
</tr>
<tr>
<td>No. of sessions per subject</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>4L + 4R</td>
<td>2L + 2R</td>
</tr>
<tr>
<td>Interval between sessions</td>
<td>Single sitting</td>
<td>Single sitting</td>
<td>Single sitting</td>
<td>Single sitting</td>
<td><strong>At least one week</strong></td>
</tr>
<tr>
<td>Environment</td>
<td>quiet room</td>
<td>clinic</td>
<td>clinic</td>
<td>regular office</td>
<td>regular office</td>
</tr>
<tr>
<td>Age range of subjects</td>
<td>Adults 20-60 years</td>
<td>Neonates</td>
<td>Neonates</td>
<td>Adults 20-40 years</td>
<td>Adults 20-30 years</td>
</tr>
<tr>
<td>Dataset primary purpose</td>
<td>Hearing test</td>
<td>Hearing test</td>
<td>Hearing test</td>
<td>Biometrics</td>
<td>Biometrics</td>
</tr>
</tbody>
</table>
4.2 Single Ear Results

4.2.1 Verification performance

For a biometric verification system, performance is measured in false acceptance rate (FAR), false rejection rate (FRR) and equal error rate (EER):

- FAR refers to the percentage that the biometric system incorrectly authenticate imposters, which can be computed as:
  \[
  \text{FAR} = \frac{\text{Number of incorrectly authenticated subjects}}{\text{Total number of imposters}} \quad (4.1)
  \]

- False rejection arises when the system fails to authenticate a genuine subject. FRR is computed using Eq. 4.2.
  \[
  \text{FRR} = \frac{\text{Number of falsely rejected subjects}}{\text{Total number of legitimate subjects}} \quad (4.2)
  \]

- EER is defined as the rate at which FAR and FRR are equal. EER is a common metric used to compare the accuracy of different verification systems, and in general, the lower EER and the better performance of the biometric system.

For example, performance of the proposed method with CWT scale = 8 on the left ear signal is exhibited in Figure 4.2, which shows a trade-off between FAR and FRR over a broad range of distance thresholds \( t_P \). Lowering \( t_P \) will improve system security, but meanwhile increase false rejection. On the other hand, raising \( t_P \) at the expense of increase of false acceptance will make the system more noise-tolerant or variation-tolerant. Also, an EER 1.87% is obtained when \( t_P = 0.440 \).

Similarly, Table 5.1 displays the EERs under 5 empirical CWT scales without and with subsequent LDA respectively. Results of using previous method proposed in [1, 2] by approximating PDF upon Euclidean distance of time series data, as well as those of
BEMD [3] at 5 most refined levels of intrinsic mode function (IMF) are also appended. Our proposed method obviously outperforms previous ones, with an EER 1.85% achieved at CWT scale = 7 for the left ear and 0.39% at CWT scale = 6 for the right. This again proves that analyzing TEOAE signal in only time domain is suboptimal, since its frequency features are neglected. Time-frequency representation based on CWT preserves both time and frequency features of the signal and flexibly extracts a particular component that can best identify itself (e.g., component extracted at scale = 7 for the left ear and 6 for the right). Besides, LDA effectively reduces the intra-class variability of such time-dependent physiological signal, thus essentially lowers false rejection and consequently the equal error. For instance, the performance measured on the left ear data at CWT scale = 7 is largely improved, with EER dropping from 9.27% (without LDA) to 1.85%.
Table 4.3: Performance in verification of the proposed method under 5 empirical CWT scales, and a similar one without subsequent LDA, as well as that of PDF estimate proposed in [1, 2] and BEMD in [3].

<table>
<thead>
<tr>
<th>Method</th>
<th>Property</th>
<th>Left Ear</th>
<th>Right Ear</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF Estimate</td>
<td>–</td>
<td>38.28 %</td>
<td>36.18 %</td>
</tr>
<tr>
<td>BEMD</td>
<td>IMF1</td>
<td>16.77 %</td>
<td>16.63 %</td>
</tr>
<tr>
<td></td>
<td>IMF2</td>
<td>9.72 %</td>
<td>9.26 %</td>
</tr>
<tr>
<td></td>
<td>IMF3</td>
<td>9.33 %</td>
<td>9.26 %</td>
</tr>
<tr>
<td></td>
<td>IMF4</td>
<td>22.19 %</td>
<td>12.96 %</td>
</tr>
<tr>
<td></td>
<td>IMF5</td>
<td>35.22 %</td>
<td>35.53 %</td>
</tr>
<tr>
<td>CWT</td>
<td>Scale = 6</td>
<td>9.73 %</td>
<td>3.69 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 7</td>
<td>9.27 %</td>
<td><strong>1.85 %</strong></td>
</tr>
<tr>
<td></td>
<td>Scale = 8</td>
<td>9.26 %</td>
<td>1.87 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 9</td>
<td>10.39 %</td>
<td>3.82 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 10</td>
<td>10.19 %</td>
<td>5.57 %</td>
</tr>
</tbody>
</table>

### 4.2.2 Identification performance

In an identification system, the performance metric is identification rate, which measures how accurate it correctly establishes an individual.

Under the same experimental scheme as verification, the efficiency of the proposed identification approach (CWT+LDA+SR) is compared against that of BEMD followed by matching with the largest pairwise correlation coefficient [3].

The proposed method also exceeds previous method and achieves the best identification rate **94.26%** at scale = 8 for the left ear and **96.30%** at scale = 6 for the right ear. BEMD, on the other hand, successively traces out local oscillation components based on extrema of the original signal. It is highly distorted by noise and not able to efficiently capture a significant trend for differentiation purpose. Moreover, the latent space learned by LDA is optimal for classification, which improves the identification accuracy.
Table 4.4: Performance in identification of the proposed method under 5 empirical CWT scales, and a similar one without subsequent LDA, as well as that of BEMD proposed in [3].

<table>
<thead>
<tr>
<th>Method</th>
<th>Property</th>
<th>Left Ear</th>
<th>Right Ear</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEMD</td>
<td>IMF1</td>
<td>68.52 %</td>
<td>66.67 %</td>
</tr>
<tr>
<td></td>
<td>IMF2</td>
<td>77.78 %</td>
<td>72.22 %</td>
</tr>
<tr>
<td></td>
<td>IMF3</td>
<td>75.93 %</td>
<td>88.89 %</td>
</tr>
<tr>
<td></td>
<td>IMF4</td>
<td>61.11 %</td>
<td>68.52 %</td>
</tr>
<tr>
<td></td>
<td>IMF5</td>
<td>29.63 %</td>
<td>25.92 %</td>
</tr>
<tr>
<td>CWT</td>
<td>Scale = 6</td>
<td>Without LDA</td>
<td>87.03 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 7</td>
<td>Without LDA</td>
<td>90.00 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 8</td>
<td>Without LDA</td>
<td>92.59 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 9</td>
<td>Without LDA</td>
<td>87.03 %</td>
</tr>
<tr>
<td></td>
<td>Scale = 10</td>
<td>Without LDA</td>
<td>83.33 %</td>
</tr>
</tbody>
</table>

4.3 Fusion of Both Ears

With promising results from single ear, the goal in fusion scenario is to improve the accuracy and system robustness by integrating information of both ears. Possible fusion strategies for both two modes are discussed as follows:

4.3.1 Fusion in Verification

Fusion in verification system is implemented based on the distance score computed upon components at CWT scale = 7 for the left ear and score computed at CWT scale = 6 for the right ear. Possible strategies include:

- **SumScore** - sum up the scores of both ears, i.e., \( S_p^{fusion} = S_p^l + S_p^r \).

- **MulScore** - multiply the scores of both ears, i.e., \( S_p^{fusion} = S_p^l \times S_p^r \).

- **MatcherWeighting(MW)Score** [38] - weighted sum of scores of both ears. Weights \( w^l \) and \( w^r \) are associated with respective EERs \( eer^l \) and \( eer^r \) achieved in single ear.
scenario, in the following relation:

\[
\begin{align*}
    w^l &= \frac{1}{eer^l (1/eer^l + 1/eer^r)} \\
    w^r &= \frac{1}{eer^r (1/eer^l + 1/eer^r)}
\end{align*}
\]  

(4.3)

In order to estimate how efficiently a fusion strategy will function, cross-validation is performed, which meanwhile gauges the generalizability of the whole methodology. For each iteration, we shuffle data from the first and second session and split them in two sets \(d_0\) and \(d_1\) which are in equal size. We then train on \(d_0\) and test on \(d_1\) so that training set will not be fixed as data from only one session. Results including EERs and FRRs satisfying zero FAR are finally averaged over 1000 iterations to produce a performance evaluation, which is shown in Table 4.5. Baseline results, i.e., performance in single ear scenario under the same cross-validation scheme, are also included.

Table 4.5: Verification performance of the proposed system based on different fusion strategies.

<table>
<thead>
<tr>
<th>Fusion strategy</th>
<th>EER (mean ± std)</th>
<th>FRR @ 0% FAR (mean ± std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ear only</td>
<td>2.67 % ± 0.93 %</td>
<td>11.96 % ± 4.06 %</td>
</tr>
<tr>
<td>Right ear only</td>
<td>0.48 % ± 0.39 %</td>
<td>11.84 % ± 2.31 %</td>
</tr>
<tr>
<td>SumScore</td>
<td>0.26 % ± 0.25 %</td>
<td>3.70 % ± 0 %</td>
</tr>
<tr>
<td>MulScore</td>
<td>0.12 % ± 0.01 %</td>
<td>2.47 % ± 0.55 %</td>
</tr>
<tr>
<td>MWScore</td>
<td><strong>0.02 % ± 0.02 %</strong></td>
<td><strong>1.54 % ± 1.39 %</strong></td>
</tr>
</tbody>
</table>

Binaural fusion evidently improves system performance: with strategy \(MWScore\) the lowest EER 0.02% is achieved, and the lowest FRR 1.55% guaranteeing zero false acceptance is attained. Strategy \(MulScore\) yields a second best result with EER 0.12% and FRR 2.47% at the safest level. And its computational cost is much lower than that of strategy \(MWScore\). Selection between these depends on requirement of specific biometric application.
4.3.2 Fusion in Identification

Fusion in identification mode is operated on the best representation extracted at scale = 8 for the left ear and scale = 7 for the right. Fusion is based on score which is the value of elements in hypothesis vector obtained by Eq. (3.11), and rank which indicates the descending order of all elements. Possible fusion strategies are as follows:

- **MaxScore** - order by the maximum of the scores, e.g., if the maximum probability of left ear $h_{\text{max}}^l$ is larger than that of right ear $h_{\text{max}}^r$, fusion result will be consistent with left ear result.

- **SumScore** - add the scores of both ears, $h_{\text{fusion}} = h^l + h^r$.

- **MulScore** - multiply the scores of both ears, i.e., $h_{\text{fusion}} = h^l \times h^r$.

- **SumWtScore** - weighted sum of scores of both ears. Greater weight is assigned to the right ear since it outperforms the left. In this examination, $h_{\text{fusion}} = 1 \times h^l + 1.5 \times h^r$.

- **SumRank** - sum up the ranks of both ears.

- **MulRank** - multiply the ranks of both ears.

- **SumWtRank** - similar to SumWtScore, $R_{\text{fusion}} = 1 \times R^l + 1.5 \times R^r$.

Similarly, Table 4.6 details fusion results averaged over 1000 cross-validation iterations. With strategy MulScore best average accuracy 99.44% is achieved, as well as the second least variation.

Last but not least, it is worth noting that right ear regularly outperforms left ear, which is mainly due to the fact that TEOAE is slightly prominent in the right versus left ear as stated in previous biological research [39,40].
Table 4.6: Identification performance of the proposed system based on different fusion strategies.

<table>
<thead>
<tr>
<th>Fusion strategy</th>
<th>Accuracy (mean ± std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ear only</td>
<td>93.07 % ± 1.33 %</td>
</tr>
<tr>
<td>Right ear only</td>
<td>95.45 % ± 0.70 %</td>
</tr>
<tr>
<td>MaxScore</td>
<td>97.51 % ± 0.87 %</td>
</tr>
<tr>
<td>SumScore</td>
<td>97.62 % ± 0.82 %</td>
</tr>
<tr>
<td>MulScore</td>
<td>99.44 % ± 0.62 %</td>
</tr>
<tr>
<td>SumWtScore</td>
<td>98.23 % ± 1.75 %</td>
</tr>
<tr>
<td>SumRank</td>
<td>97.25 % ± 0.95 %</td>
</tr>
<tr>
<td>MulRank</td>
<td>98.96 % ± 0.85 %</td>
</tr>
<tr>
<td>SumWtRank</td>
<td>98.14 % ± 0.04 %</td>
</tr>
</tbody>
</table>
Chapter 5

Template Protection with BioHashing

5.1 Motivation for TEOAE Template Protection

Although TEOAE biometrics offer various advantages over conventional modalities, e.g., inherent robustness to attacks and natural liveness detection, there exist some problems yet to be addressed for the widespread deployment of such a new technique. Foremost, as TEOAE biometric data reflect personal traits, the sensitive user template should be stored in a highly protecting format in the biometric system. Otherwise, original TEOAE data will be recovered once the vulnerable template is compromised. Secondly, due to the long-term stability of TEOAE, its biometric characteristics are mainly immutable, i.e., user’s unprotected templates stored in different systems are interchangeable. Templates intercepted from one system can be re-injected to another system, which will result in permanent identity theft. Different passwords or access cards are used in different applications, similarly, no same user’s templates can be stored across various systems. To address these problems, protection of biometric data using cryptography (a.k.a. BioHashing [41,42]) for TEOAE biometrics is first proposed here, which fulfills the following
robustness criteria:

- Irreversibility: templates are transformed to a distorted version to prevent recovery of the original data.

- Diversity: templates stored in one database cannot be linked to other databases and meanwhile guarantees decent recognition performance.

5.2 Related Works

To the best of our knowledge, there has not been any prior work that addresses BioHashing for TEOAE. However, a number of previous works on BioHashing for face, fingerprints or other modalities are of significant reference value. Bolle et al. [43] proposed the concrete idea of cancelable biometrics, where a repeatable and non-invertible transformation is applied on the biometric templates. Subsequent comparison is directly conducted on the distorted and irreversible version without any decryption. Following this line, Savvides et al. [44] demonstrated correlation energy filters for face recognition; Ratha et al. [45] proposed fingerprint data transformation based on by Cartesian, polar, and functional projections. Although this strategy enhances template security by subspace projection, it enlarges intra-class variation and thus partly deteriorates the recognition performance. Alternatively, random projection proposed in [46, 47] transforms face image templates onto a user-specific random orthogonal matrix and meanwhile preserves the similarity between the original data. As the individual random orthogonal matrix is associated with an external key (token, USB) or passwords, whose storing raises another security concern and being compromised will jeopardize the performance [48].

In this work, we adopt the mechanism of random projection by projecting efficient feature sets extracted by previous algorithm [10] onto random subspaces and discretizing resulting vectors into hash code. By virtue of the inherent robustness to circumvention and compromise offered by TEOAE biometrics, the pseudo-random matrix simply
serves as a carrier of one-way projection. Thus, individual random matrices are directly stored in the system database, preventing another security threat. Quantitative analysis and experimental results validate that the proposed method fulfills robustness criteria, irreversibility and diversity, and meanwhile guarantees decent recognition performance.

5.3 BioHashing for TEOAE Biometrics

This section presents the BioHashing methodology for TEOAE based on random projection. An overview of the proposed method is first provided, followed by detailed specification and framework of the robust biometric system.

5.3.1 Overview of BioHashing

An extracted TEOAE feature $z \in \mathbb{R}^m$ after LDA before being stored in the system gallery is projected onto a random subspace for protection according to the following procedure:

1. Generate a pseudo-random matrix \( \{ r_i \in \mathbb{R}^m \mid i = 1, 2, ..., n \} \), where \( n \) is the length of the expected binary string \( b \) after hashing, and \( n \leq m \).

2. Apply Gram-Schmidt algorithm to transform the previous pseudo-random matrix \( r \) into an orthonormal matrix \( \{ R_i \in \mathbb{R}^m \mid i = 1, 2, ..., n \} \).

3. Realize random projection of \( z \) with \( R \) by computing
   \[
   \{ v_i = \langle z, R_i \rangle \in \mathbb{R} \mid i = 1, 2, ..., n \},
   \]
   where \( \langle \cdot, \cdot \rangle \) denotes the inner product operation; or alternatively:
   \[
   v = R^T z. \tag{5.1}
   \]

4. Binarize to obtain the \( n \) bits BioHash \( b = \{ b_i \mid i = 1, 2, ..., n \} \) from:
   \[
   b_i = \begin{cases} 
   0 & \text{if } v_i \leq \tau \\
   1 & \text{if } v_i > \tau 
   \end{cases} \tag{5.2}
   \]
where threshold $\tau$ is selected in a way that half of the projections $v_i$ are greater than $\tau$ and the other half are less than $\tau$.

The aforementioned procedure of BioHashing is also depicted in Figure 5.1.

Finally, similarity score $D$ is straightforwardly measured upon BioHash vectors in normalized Hamming distance as a metric. For instance, similarity score of $b^{(a)}$ and $b^{(b)}$ (both are of the same length $n$) is computed by:

$$D(b^{(a)}, b^{(b)}) = \frac{1}{n} \sum_{i} (b_i^{(a)} \oplus b_i^{(b)})$$  \hspace{1cm} (5.3)

where $\oplus$ denotes the XOR operation.

### 5.3.2 Detailed Analysis of Random Projection

Significantly, for effective biometrics, the BioHashing formulation should not weaken the recognition performance. Thus, pairwise similarity in the input feature space should be
preserved in the resulting projected space.

Given two TEOAE feature vectors $\mathbf{z}^{(a)}$ and $\mathbf{z}^{(b)}$, distance between respective random projected vectors $\mathbf{v}^{(a)}$ and $\mathbf{v}^{(b)}$ can be computed as follows:

\[
\|\mathbf{v}^{(a)} - \mathbf{v}^{(b)}\|^2 = \|\mathbf{R}^T \mathbf{z}^{(a)} - \mathbf{R}^T \mathbf{z}^{(b)}\|^2 \\
= \|\mathbf{R}^T \mathbf{z}^{(a)}\|^2 + \|\mathbf{R}^T \mathbf{z}^{(b)}\|^2 - 2(\mathbf{R}^T \mathbf{z}^{(a)})^T (\mathbf{R}^T \mathbf{z}^{(b)}) \\
= \|\mathbf{z}^{(a)}\|^2 + \|\mathbf{z}^{(b)}\|^2 - 2(\mathbf{z}^{(a)})^T \mathbf{R} \mathbf{R}^T \mathbf{z}^{(b)} \\
\approx \|\mathbf{z}^{(a)}\|^2 + \|\mathbf{z}^{(b)}\|^2 - 2(\mathbf{z}^{(a)})^T \mathbf{z}^{(b)} \\
= \|\mathbf{z}^{(a)} - \mathbf{z}^{(b)}\|^2
\]

In virtue of the orthonormal properties of $\mathbf{R}$, i.e., $\|\mathbf{R}\| = 1$ and $\mathbf{R} \mathbf{R}^T \approx \mathbf{I}$, random projection effectively preserves the pairwise distance of the original features. Particularly, $\mathbf{R}^T \mathbf{R} = \mathbf{I}$ when $n$ is sufficiently large. This indicates that the degree of preservation of pairwise distances increases with the dimension $n$ of the random projection subspace, and reaches the upper limit with the possible maximum one, i.e., $n = m$ [41].

Again, we take the dataset [17] as an example: after features are extracted by using CWT/LDA algorithm, TEOAEs from the same subjects (genuine pairs) are mapped close to each other with small pairwise distances, while those from different subjects (imposter pairs) are mapped far apart with comparably large distances, as shown in the distribution in Figure 5.2 (a). During BioHashing operation, random projection successfully maintains such pairwise similarity as in the original feature space. This can be indicated by the histogram of pairwise distances in Figure 5.2 (b).

### 5.3.3 Non-invertible Quantization

After random projection, each resulting vector is quantized into a binary string via a predefined threshold $\tau$. Such discretization makes it computationally infeasible to recover
Chapter 5. Template Protection with BioHashing

Figure 5.2: Histogram of pairwise distances between genuine pairs (red) and those between imposter pairs (blue) in the following feature spaces: (a) Original feature, (b) Projection feature and (c) Projection feature after quantization.

the projected feature vector from the binary string. Even if \( \tau \) is given to the adversary, possible solutions are still infinite.

As earlier stated, \( \tau \) is empirically selected based on a large dataset, so that half of the projected outputs are above it and the rest half below it. Suppose pseudo-random matrix \( r \) is generated according to standard normal distribution, its orthonormal matrix \( R \) is also normally distributed, which leads to an approximately normal distribution of the projected vectors \( v \) (see Figure 5.3). With \( \tau = 0 \), information loss due to quantization is minimized. Further, as Hamming distance is a measure of bitwise disagreement which relates to the considerable difference of the projected vectors, pairwise distances are consequently preserved to the largest degree (Figure 5.2 (c)).

5.3.4 Framework of Biometric System with BioHashing

Built upon the preceding TEOAE biometric system, a patched version with BioHashing is presented in Figure 5.4.

In the enrollment session, after TEOAE is acquired from each enrollee and subjected
to CWT/ LDA based feature extraction, individual pseudo-random matrix is generated by the system. BioHashing is then executed on the extracted feature sets of each user with respective pseudo-random matrix. Resulting binary templates \( \{b^{(1)}, b^{(2)}, \ldots, b^{(K)}\} \), along with the individual pseudo-random matrices and established threshold \( \tau \), are stored in system database.

In verification mode where an individual gets recorded and claims an identity \( k \), a hash vector \( b' \) is created after feature extraction and BioHashing operation with identity \( k \)’s pseudo matrix. Afterwards, system will retrieve the corresponding template \( b^{(k)} \) from database, which is then matched against \( b' \) via Eq. 5.3. An ACCEPT decision is made only if \( D(b^{(k)}, b') \leq D_{th} \), where \( D_{th} \) is the system threshold.

In [46,47], pseudo-random matrix is associated with user-specific external token (e.g., USB) or ancillary input (e.g., password), and essentially becomes an extra confidential key compensating the insufficient of conventional biometric modalities. However, it turns out to be another security threat since the external keys are easily lost or stolen as well. In this work, individual pseudo-random matrices are directly stored in the system database instead. By virtue of the inherent robustness to circumvention and compromise offered by TEOAE biometrics, the random matrix simply serves as a carrier projecting feature sets onto a secured subspace. When an identity claim is made, corresponding “carrier”

Figure 5.3: Distribution of \( v \), approximately \( N(0, 1) \).
Figure 5.4: Framework of the proposed biometric verification system with BioHashing.

is recalled from the system to project the probe vector onto the same subspace of the corresponding template set.

5.4 Performance Evaluation

In this section, experiments are conducted on the same dataset [17]. Same as previous experiments, last 10 recordings of the first and second session per subject are used to constitute the training and testing set respectively. Binary string (hash) for each subject is set to be of the same length as the feature vector after CWT/LDA operation, i.e., $n = m$.

5.4.1 Experimental Results

In Figure 5.5, the performance of the proposed hashing method using left ear signal under CWT scale = 7 is exhibited, where an EER 1.59% is obtained and FRR = 11.85% satisfies
a requirement of no false acceptance. Similarly, Table 5.1 displays respective performance of using signal from each ear under different CWT scales. Results are averaged over 1000 iterations as the pseudo matrices are randomly generated according to standard normal distribution. Performance of using previous method without BioHashing is also appended.

![Graph showing verification performance of left ear signal with CWT scale = 7.](image)

**Figure 5.5:** Verification performance of left ear signal with CWT scale = 7.

Despite the distortion of feature sets for privacy cryptography, the proposed method successfully guarantee biometric performance with comparable EERs achieved, i.e., 1.78% at CWT scale = 7 for the left ear and 0.51% at CWT scale = 7 for the right. Besides, proposed method outperforms previous one by operating a lower FRR that guarantees zero FAR, such as 6.58% (with BioHashing) versus 14.81% (without BioHashing) at CWT scale = 8 for the right ear.

### 5.4.2 Security Analysis

As stated above, unlike conventional modalities that are accessible or easily captured in public, TEOAE is an outcome of complex physiological activities of human auditory perception in response to specific click sound stimulation. Thanks to this nature, TEOAE biometrics provides an inherent robustness to replay and falsification attacks.
Table 5.1: Performance of proposed method with BioHashing and previous one without BioHashing respectively based on various CWT scales.

<table>
<thead>
<tr>
<th>CWT scale</th>
<th>Left Ear</th>
<th></th>
<th></th>
<th>Right Ear</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without BioHashing</td>
<td>With BioHashing</td>
<td>Without BioHashing</td>
<td>With BioHashing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.69 %</td>
<td>1.79 %</td>
<td>36.67 %</td>
<td>11.46 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.85 %</td>
<td>1.78 %</td>
<td>15.56 %</td>
<td>12.48 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.87 %</td>
<td>2.09 %</td>
<td>12.41 %</td>
<td>11.61 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.39 %</td>
<td>0.59 %</td>
<td>13.89 %</td>
<td>6.77 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.67 %</td>
<td>0.51 %</td>
<td>20.00 %</td>
<td>8.79 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.85 %</td>
<td>0.58 %</td>
<td>14.81 %</td>
<td>6.58 %</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The biometric system where sensitive template data are stored consequently becomes the main intrusion target. Security analysis on the robustness of the proposed system with BioHashing against attacks is given as follows:

(a) Imposter attack: an imposter claims to be a legitimate user with his own TEOAE signal. The experimental results shown above, i.e., EER = 1.78% at decision threshold $D_{th} = 0.344$ for the left ear signal and EER = 0.51% at decision threshold $D_{th} = 0.308$ for the right, indicate that fraudulent authentication hardly succeeds.

(b) Brute-force attack: the adversary does not have any knowledge of the hash template and tries out all possible combinations of 0 and 1. The computational complexity is $2^n$, where $n$ is the length of hash template. Moreover, as $n$ can be equal or any value close to but not exceeding the dimension $m$ of LDA-based feature extractor, it is almost impossible to be illegitimately authenticated with exhaustive search in a tolerable time.

(c) Recovery of original biometric data from compromised hash template: even if the
attacker has intercepted the hash template, it is still computationally infeasible to recover the projected feature vector from the binary string. Irreversibility is achieved by binary quantization as a one-way transformation.

(d) Re-injection across various systems using compromised hash template: now the attacker with intercepted hash template gives up recovering original biometric data, instead, he assumes similar templates are stored in various databases and attempts to implant this template in the other system(s). Nevertheless, as the individual pseudo-random matrix is randomly generated when a user registers for a system, original biometric data from the same subject will be projected onto different and independent subspaces among systems. We analyse such non-linkability of hash templates of the same users across systems by pairwise distance. Figure 5.6 displays the distribution of pairwise Hamming distances among 100 TEOAE biometric systems, where only 0.19% pairs are of distance smaller than decision threshold 0.344 and can be fraudulently accepted in mutual systems. Diversity/ non-linkability across various systems is well fulfilled in virtue of the randomly generated carrier matrix.

![Figure 5.6: Histogram of pairwise distances of the same users’ hash templates across 100 systems.](image-url)
Chapter 6

Conclusions and Future Improvements

Nowadays, advanced forgery and spoofing techniques are threatening the reliability of conventional biometric modalities. To address this issue, in this thesis, we propose and evaluate an identity recognition system based on Transient Evoked Otoacoustic Emission signals. The feasibility and effectiveness of TEOAE for biometrics are demonstrated, along with the corresponding perks, including inherent robustness against falsification and replay attacks, high reliability with implicit liveness detection and compatibility with mobile devices.

TEOAE, as a simple and non-invasive tool for hearing assessment, is prevalently used in universal infant hearing screening programs; interpersonal differences of TEOAE signal suggest an individual morphology of inner ears especially OHCs. Together with its collectability, universality and uniqueness, the stability proved by the high reproducibility both on short-term and on long-term progression, which is discovered in clinical tests and biological studies [24,25], essentially makes it a doable biometric modality.

The effectiveness of TEOAE for biometrics is fulfilled by the presented methodology: wavelet analysis to derive time-frequency representation of the signal and LDA-based
optimal subspace learning algorithm to reduce intra-subject variability.

As stimulated by a broadband click sound, TEOAE is a cumulative response consisting of several dominant frequency components with different latency. Investigating TEOAE in either time domain or frequency domain is suboptimal. In this work, time-frequency analysis based on continuous wavelet transform is described, which preserves both time and frequency features of TEOAE signal. It also reveals the intra-subject stability and the inter-subject disparity that are obscure in original time series signals.

Unlike most conventional biometric modalities, e.g., fingerprints and iris, TEOAE as a physiological signal is time-variant depending, and so is the corresponding time-frequency feature derived from it. Such a dynamic nature is taken into consideration, and LDA optimal subspace learning algorithm is employed to further reduce the intra-subject variability and select more effective features prior to matching procedure.

After LDA-based feature extraction from time-frequency representations, the proposed biometric matching operation compares the probe against the templates stored in system gallery. Depending on a specific application, there are two modes of matching to contemplate, verification and identification. In verification mode, Pearson’s distance is chosen as a metric to validates a user’s claimed identity. In identification mode, an unknown user’s identity is established by a multiclass classifier softmax regression model. A complete framework of the TEOAE biometric system in both modes is described afterwards, considering practical application, along with its utilization into mobile commerce as an example.

It is demonstrated that the proposed approach ensures state-of-the-art performance, with EER 1.85% achieved on the left ear signal and 0.39% the right ear in verification mode, and accuracy 94.26% attained on the left ear signal and 96.3% the right ear in identification mode. As TEOAE emitted from the left and the right ear are not identical, fusion of information from both ears further improves the accuracy of the biometric system and reduces the probability of false authentication.
Although TEOAE is naturally immune to replay and falsification attacks, corresponding biometric system where sensitive template data are stored will become the main intrusion target. A data cryptography method to secure TEOAE template data is additionally presented. Quantitative analysis and experimental results justify that the proposed method fulfills robustness criteria, irreversibility and diversity, and meanwhile guarantees decent recognition performance, i.e., EERs 1.78% and 0.51% are achieved for the left and right ear signal respectively. Further, robustness of the proposed acoustic “signature” system is verified by security analysis on four attack scenarios.

There are still two possible directions that could be extended in future research:

- General CWT scale configuration. In current work, CWT scales 6 to 10 are selected via empirical experiment on a 54-subjects database. In real applications, it is significant to choose a common CWT scale that is applicable to a larger population.

- Potential effects on the accuracy of biometrics by changing the stimulus parameters, including the click duration (or frequency band) and intensity.

- Factors that could affect accuracy of TEOAE for biometrics. External interferences include environmental noise and position of the earphone. Our data collection was conducted in a regular office, however in real world application, how much influence environmental noise will make should be examined. Besides, fitting of the earphone placed in the ears is vital, which affects not only the stimuli input to the ears but also the collected TEOAE response. Internal interferences are mainly body condition, such as health and conciseness. Although biological study [49] reveals that sleep mechanisms have no effect on TEOAE amplitude, it is worth exploring the case in TEOAE biometric characteristics.
Bibliography


