

TRANSIENT OTOACOUSTIC EMISSIONS FOR BIOMETRIC RECOGNITION

Jiexin Gao, Foteini Agrafioti, Siyuan Wang, Dimitrios Hatzinakos

The Edward S. Rogers Sr. Department of Electrical and Computer Engineering,
University of Toronto,
10 King's College Road, Toronto, ON, Canada, M5S 3G4

ABSTRACT

This paper investigates the potential use of Transient Otoacoustic Emissions (TEOAE) for biometric recognition. Multiresolution decomposition of TEOAE is done by a modified Bivariate Empirical Mode Decomposition (BEMD) combined with an auditory model. Matching scores are computed by combining ranked correlations across different levels. Recognition rate with recording from left ear is 96.30% and can be improved to 98.15% by utilizing a matching score fusion with information from right ear.

Index Terms— Empirical Mode Decomposition, Transient Otoacoustic Emission, Biometrics

1. INTRODUCTION

Security concerns involved in human identification applications are continuously increasing. Traditional methods of identity authentication require something that the subject can remember (e.g. password) or possess (e.g. ID cards). Due to the advances in technology, the effectiveness of such identification methods is now questioned. The need for universal, cost efficient and difficult to fraud techniques is prominent.

This work focus on the investigation of the potential use of the Transient Evoked Otoacoustic Emissions (TEOAE) for human authentication. TEOAE are low level sound waves generated by an active process in the cochlea that can be stimulated by a click sound (white noise pulse) and can be collected by a sensitive microphone in the ear. Since its discovery in 1978 by Kemp [1], TEOAE has been widely applied to areas such as early diagnosis of hearing loss and newborn screening. Its uniqueness within each person and long term stability make it feasible as biometric modality and can be applied to areas such as new born verification and potentially any environment that a microphone can be used, such as human identification on mobile devices.

The paper is organized as follows. In Section 2 a brief review of previous works are given, followed by a introduction to TEOAE in Section 3. Signal collection device and procedure is presented in Section 4. In Section 5 detailed methodol-

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

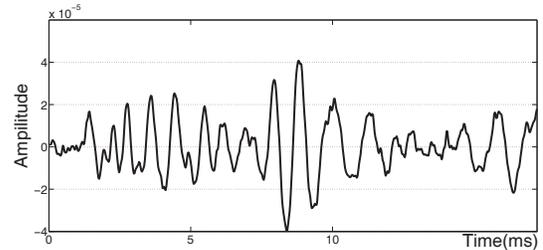


Fig. 1. Example of TEOAE

ogy is covered including the modified BEMD, recognition algorithm and fusion scheme. Simulation results are presented in Section 6 with some discussion on possible influencing factors.

2. RELATED WORKS

A feasibility study of using TEOAE as a biometric modality was conducted by Swabey [2]. The dataset that was investigated consists of one adult short-term dataset with 23 subjects (recorded within same session, which is of less value for biometric evaluation purpose), one neonate dataset with 760 subjects (no report on time interval between two recording sessions) and one adult long-term dataset with 6 subjects (time interval between two sessions was 6 months). Maximum likelihood estimation was employed to approximate the probability density function of inter-class and intra-class distance. The reported Equal Error Rate (EER) was 1.24%, 2.29% and 2.35%, all with 90% confidence, respectively.

3. TRANSIENT OTOACOUSTIC EMISSIONS

Transient Otoacoustic Emissions are low level sound waves produced by an active process in cochlea. The signal is highly nonlinear and nonstationary. Example of TEOAE recorded from Vivosonic Integrity system, with low frequency trend removed, is shown in Figure 1.

Table 1. Experiment Protocol

Stimulus Parameters	
STI-Mode	Nonlinear
Click Interval	21.12ms
Click Duration	80ms
Click Level	80dB peSPL
Test Control	
Record Window	2.8 – 20ms
Low Pass Cut-off	6000 Hz
High Pass Cut-off	750Hz
Artifact Rejection Threshold	55dB SPL

4. SIGNAL COLLECTION

Signal collection was conducted in BioSec Laboratory at University of Toronto, approved by University of Toronto Protocol reference number 23018. Vivosonic Integrity system was used with protocol details as shown in Table 1. To ensure the quality of recording but not to constrain the environment too much, earmuff was used for noise canceling but the experiment room was a regular office where there were people talking and entering/leaving the office. The participants were only given the instruction to sit on a chair and relax. 54 subjects were successfully recorded in both sessions, with the time between sessions at least one week to validate long-term stability. Most of the subjects are between the age 20 and 30.

Dataset consists of one recording of length 17.2ms per ear per session for each subject.

5. METHODOLOGY

In this section we present the methodology used to process and classify TEOAE recordings. Denote TEOAE recorded during first session(enrollment) as $\{X_{Li}\}, \{X_{Ri}\}$ and TEOAE recorded during second session(recognition) as $\{Y_{Li}\}, \{Y_{Ri}\}$, with subject IDs $i = 1, 2 \dots 54$ and $\{L, R\}$ for left or right ear.

5.1. Modified BEMD with Auditory Model

The frequency selectivity of cochlea and oscillatory nature of TEOAE signal make it suitable to apply Empirical Mode Decomposition(EMD) which decomposes a signal into multi-level local oscillation components. EMD is a method proposed by Huang et al [3] for processing nonlinear and nonstationary data. It is adaptive and efficient but suffers from the problem of variable number of levels among decomposition of different signals. The only way to address this uniqueness problem is to apply Bivariate EMD (BEMD) but there is no guarantee that decomposition of signal pair (f_1, f_2) will results in the same number of intrinsic mode functions (IMF) as the decomposition of (f_1, f_3) , which makes comparison

at multiple levels difficult. To address the problem of uniqueness and to get a meaningful decomposition of OAE signal, an auditory model proposed by Zheng [4] is incorporated into the decomposition procedure as shown in Figure 2. Synthesized sinusoids of different frequency that represent the characteristic response of one region of the cochlea are used to guide decomposition of each level. For $(M + 1)$ -level decomposition (M IMFs and 1 residue), the reference frequency used at level k is calculated by:

$$f = \frac{f_o}{q^{2Mk}}$$

where base frequency $f_o = 15165.4\text{Hz}$ and $q = 1.0352952$. At each level residue from previous level is decomposed together with the synthesized reference sinusoid by a 1-level BEMD, which only extracts the highest frequency at the corresponding level. After completing each level of decomposition, the extracted IMF is removed from the signal and the procedure continues on the residue.

In the algorithm we implemented for TEOAE the decomposition stops after level 4 since in our recognition stage only the first few IMFs are needed, although in order to get a detailed and physically meaningful representation of the TEOAE signal, a 9-level decomposition is necessary. Also in particular for TEOAE signals, since high frequency components of TEOAE exhibit shorter latency and duration, in order to remove noise from the recording, IMF1 and IMF2 are multiplied with a mask:

$$W(t) = \begin{cases} 1 & 0 \leq t < 3.9\text{ms} \\ 1 + \cos(\frac{t-3.9}{2.6}\pi) & 3.9\text{ms} \leq t < 6.5\text{ms} \\ 0 & 6.5\text{ms} \leq t \leq 17.2\text{ms} \end{cases}$$

All OAE recordings are processed by using the above procedure. Denote the decomposed signals after applying mask as $\{x_{Lik}\}, \{x_{Rik}\}$ for enrollment session and $\{y_{Lik}\}, \{y_{Rik}\}$ for recognition session with subject ID i and $k = 1, 2 \dots 8$ denote IMF index. An example of IMF1 – 4 from this procedure is depicted in Figure 3.

5.2. Recognition

Recognition can be done with recording from either the left or right ear. For the simplicity of discussion, we assume the use of recording from left ear in this subsection.

For the recording from an unknown subject n with recording Y_{Ln} in recognition session, we want to find the best match identity in enrolled recordings. This is done by the following steps:

- **Correlation matrix** Correlations between IMFs are calculated with the corresponding subject ID for each entry as follows:

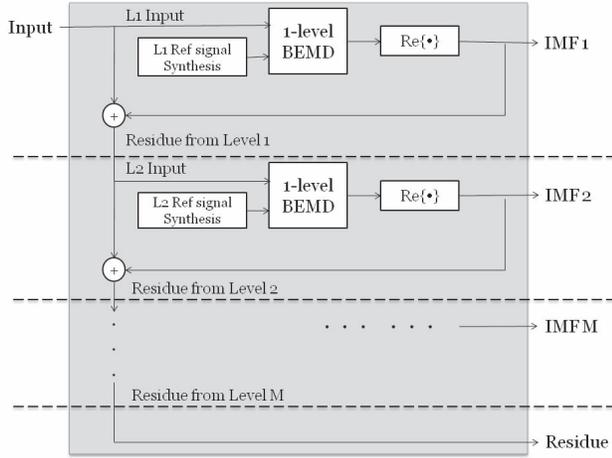


Fig. 2. Modified BEMD with auditory model

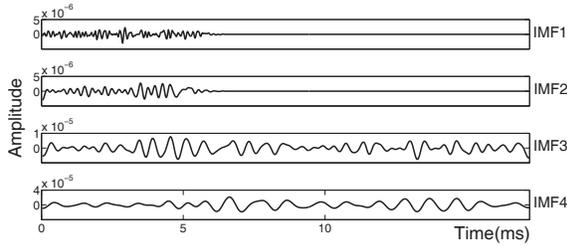


Fig. 3. IMF1 – 4 from decomposing a raw TEOAE recording using the procedure discussed in Section 5.1

$$\mathbf{C}(\mathbf{i}, \mathbf{k}) = \text{corr}(x_{L_{ik}}, y_{L_{nk}})$$

$$\mathbf{I}(\mathbf{i}, \mathbf{k}) = i$$

with $i = 1, 2 \dots 54$, $k = 1, 2, 3, 4$, and $\text{corr}(\mathbf{a}, \mathbf{b})$ denotes the correlation between two vectors \mathbf{a} and \mathbf{b} .

• **Ranked, normalized and weighted correlation matrix**

After computed the correlation matrix $\mathbf{C}_{54 \times 4}$ (subscript denotes size of the matrix), we sort each column in descending order and get $\mathbf{C}'_{54 \times 4}$. Normalize each entry in $\mathbf{C}_{54 \times 4}$ with respect to first row and keep the first 3 rows (3 highest ranked matches for each level) yield $\tilde{\mathbf{C}}_{3 \times 4}$ with $\tilde{\mathbf{C}}_{(1,k)} = 1$ and $\tilde{\mathbf{C}}_{(i,k)} = \frac{\mathbf{C}'_{(i,k)}}{\mathbf{C}'_{(1,k)}}$ for $i = 1, 2, 3$, $k = 1, 2, 3, 4$. $\mathbf{I}_{54 \times 4}$ is also re-ordered correspondingly as $\tilde{\mathbf{I}}_{3 \times 4}$.

With an empirical weight imposed, score matrix is computed as:

$$\mathbf{S}_{3 \times 4} = \begin{bmatrix} 1 & 0.8 & 0.8 & 0.6 \\ \tilde{\mathbf{C}}_{(2,1)} & 0.8\tilde{\mathbf{C}}_{(2,2)} & 0.8\tilde{\mathbf{C}}_{(2,3)} & 0.6\tilde{\mathbf{C}}_{(2,4)} \\ \tilde{\mathbf{C}}_{(3,1)} & 0.8\tilde{\mathbf{C}}_{(3,2)} & 0.8\tilde{\mathbf{C}}_{(3,3)} & 0.6\tilde{\mathbf{C}}_{(3,4)} \end{bmatrix}$$

• **Matching score**

Denote the collection of all unique IDs in $\tilde{\mathbf{I}}_{3 \times 4}$ as $\{I_u\}$ with $1 \leq u \leq N$ and $N \leq 12$. For every $I_u \in \{I_u\}$, final scores

$$S_u = \sum_{\tilde{\mathbf{i}}(m,n)=I_u} \mathbf{S}(\mathbf{m}, \mathbf{n})$$

• **Decision**

Sort $\{S_u\}$ in descending order as $\{\tilde{S}_u\}$ with $\{I_u\}$ re-ordered as $\{\tilde{I}_u\}$. 3 best matched identities are \tilde{I}_1 , \tilde{I}_2 and \tilde{I}_3 with corresponding scores \tilde{S}_1 , \tilde{S}_2 and \tilde{S}_3 . The subject is identified as \tilde{I}_1 .

5.3. Fusion of left and right ear

In application scenario where high accuracy is required, a score level fusion from both ears can be employed to improve system performance. Suppose we have the three best matches from left ear enrolled recordings with their identities I_{L1} , I_{L2} , I_{L3} and scores S_{L1} , S_{L2} , S_{L3} . Those from right ear are denoted as I_{R1} , I_{R2} , I_{R3} and S_{R1} , S_{R2} , S_{R3} .

If results from both sides agree with each other, that is $I_{L1} = I_{R1}$, final identified subject ID is I_{L1} .

If $I_{L1} \neq I_{R1}$ the matched identity is calculate as:

• **Concatenate subject IDs and scores**

$$I_c = [I_{L1} \ I_{L2} \ I_{L3} \ I_{R1} \ I_{R2} \ I_{R3}]$$

$$S_c = [S_{L1} \ S_{L2} \ S_{L3} \ S_{R1} \ S_{R2} \ S_{R3}]$$

• **Fused matching score**

Denote the collection of all unique IDs in I_c as $\{I_u\}$. Final scores $\{S_u\}$ are computed as follows:

$$S_u = \sum_{I_c(m)=I_u} S_c(m)$$

• **Decision**

Sort $\{S_u\}$ in descending order as $\{\tilde{S}_u\}$ with $\{I_u\}$ re-ordered as $\{\tilde{I}_u\}$. The subject is identified as \tilde{I}_1 .

6. RESULT

Recognition performance is summarized in Table 2 for three different scenarios: using left ear recording only, using right ear recording only and the fusion of two ears. Right ear performance is slightly lower than left ear. One possible cause may be the additive Spontaneous Otoacoustic Emission (SOAE), which might not be as unique within each individual as TEOAE and has been proved to exist together with TEOAE and exhibit greater intensity in right ear than in left ear [5]. With the fusion of information from two ears a recognition rate of 98.15% is achieved which is quite promising since there is no statistical analysis required in the algorithm.

Table 2. Performance

Scenario	Correctly Recognized	Performance
Left	52 out of 54	96.30%
Right	49 out of 54	90.74%
Fusion	53 out of 54	98.15%

7. CONCLUSION

In this paper a framework for biometric recognition using Transient Otoacoustic Emissions was presented. TEOAE signals from 54 subjects for long-term stability validation purpose was collected. By using a modified BEMD with auditory model and a recognition algorithm without statistical analysis, a recognition rate of 98.15% can be achieved with fusion of information from both ears. In the future, the group would like to work on comparing the proposed method with different signal decomposition approaches such as wavelet transform, as well as evaluating the performance on larger datasets and under special circumstances such as subjects with hearing difficulties or with excessive noise exposure.

8. ACKNOWLEDGEMENT

The authors would like to thank Vivosonic Inc.(Toronto)[6] for providing the Vivosonic Integrity System. Moreover, the authors would like to thank professor Francis M. Bui from University of Saskatchewan, for his insightful suggestions on investigating the signal.

9. REFERENCES

- [1] David T Kemp, "Otoacoustic emissions, their origin in cochlear function, and use," *British Medical Bulletin*, vol. 63, no. 1, pp. 223–241, 2002.
- [2] Matthew A. Swabey, Paul Chambers, Mark E. Lutman, Neil M. White, John E. Chad, Andrew D. Brown, and Stephen P. Beeby, "The biometric potential of transient otoacoustic emissions," *Int. J. Biometrics*, vol. 1, pp. 349–364, March 2009.
- [3] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [4] Ling Zheng, Yuan-Ting Zhang, Fu-Sheng Yang, and Da-Tian Ye, "Synthesis and decomposition of transient-evoked otoacoustic emissions based on an active auditory

model," *Biomedical Engineering, IEEE Transactions on*, vol. 46, no. 9, pp. 1098 –1106, sept. 1999.

- [5] Yael Raz, "Otoacoustic emissions: Clinical applications, 3rd edition, martin s. robinette, theodore j. glattke, eds., new york: Thieme medical publishers, 2007," *The Laryngoscope*, vol. 117, no. 9, pp. 1700–1700, 2007.
- [6] "<http://www.vivosonic.com/>," .