Modelling and Analysis of Transient Evoked Otoacoustic Emissions for Human Biometric Recognition

Johnson I Agbinya and Shaza Abuelgasim

[1]*School of Information Technology and Engineering*
Melbourne Institute of Technology
288 La Trobe Street, Melbourne 3000, Australia
Email: jagbinya@mit.edu.au
[2]*Faculty of Computer Science and Information Technology*
Sudan University of Science & Technology, Khartoum, Sudan
Email: shazasaber2015@gmail.com

Abstract: This paper reports on application of data analytics techniques for determining the status of human ear, variation in ear conditions and the specifics of whose ear they are. Principal component analysis of transient evoked otoacoustic emissions from human ears were derived based on recording of responses to external audio excitations undertaken fifteen months apart. Results indicate the method established in the paper is suitable for person recognition and for identifying when deterioration of the hearing performance of an ear has taken place. Ear transformation matrix is introduced. The transformation matrix represents a scaling of the eigenvectors using a Hermitian matrix or scaling matrix. They are however known to be the eigenvalues obtained from the PCA analysis. While the eigenvalues could be seen to represent audio loudness scaling, the eigenvectors represent further ear deterioration. Eigenvalues are maintained when the audio performance of ear to external excitation has not changed. Variation of the Hermitian is also variation of the ear condition. The lengths of the eigenvectors are considered as estimates of the change in ear loudness performance as it can be seen as equivalent to the power content of the eigenvectors.

1 Introduction

There has emerged growing interest in the application of machine learning in determining ear conditions [1-4]. A comparative analysis using various classification methods was undertaken in [2] using support vector machine (SVM), neural network multilayer perceptron (NNMP), random forest (RF), and adaptive boosting (AB). These established machine learning algorithms were used to predict noise-induced hearing-impairment among workers in industrial environments. Predictions between 78.6 and 80.1% accuracies were recorded with the SVM outperforming other methods. Unfortunately, the opportunity provided by the multiple classification method to evaluate if and how SVM compare with ensemble modelling was missed by the authors. Crawson et al sought to show that depression can be predicted using hearing impairment [3] and arrived at a conclusion that "machine learning algorithms can accurately predict PHQ-9 depression scale scores" using Patient Health Questionnaire-9 [PHQ-9]. The goal in the study undertaken by and Pamela Souza [4] was to test the machine learning algorithms predicting audiometric configurations when given *limited* information as in when using uncalibrated equipment in a person's home [4]. This approach although much more suited to current pandemic settings as in covid-19 situations provides for self-tests and thus if successful could lead to a risk-free self-assessment of hearing loss. As more and more modern data processing approaches are being adopted, the state of the ear remains the ultimate interest in the estimation methods. The state of an object can be estimated using methods such as principal component analysis to establish ear eigenvalues and eigenvectors. Such methods rely on the scaling properties and the evolution of the resulting Hermitian matrices which help to establish ear state from time-to-time or its evolution. Other emerging data analytics methods can therefore provide further insight into the study of ear conditions. That is the subject of this paper.

2 Structure of the Human Ear

The human ear is well described anatomically by biologists and in terms of acoustic performance by physicists and engineers. Discussions in this section are informed by the physics and electronic models of the human ear. The human ear is modelled in sections, as a one-dimensional lossy transmission line circuit (ear canal, see Figure 1). This is terminated with a distributed load impedance formed by the middle ear and inner ear including the cochlea [5].

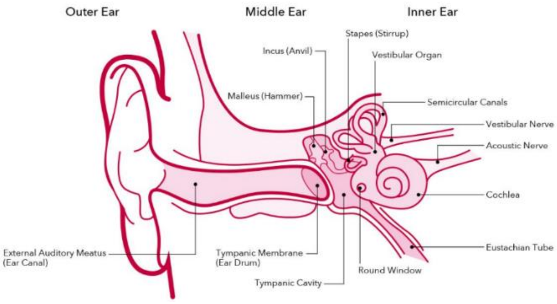

Figure 1: Biological Structure of The Human Ear [5]

The ear canal behaves like a pipe organ typically of length between around 2.1cm and 2.6cm. therefore the fundamental frequency excited by the ear canal in response to an external stimulus can be estimated based on an expression which takes into account its length and the frequency modes it excites as:

$$f_n = \frac{n\upsilon}{4\ell} \qquad (1)$$

The fundamental frequency $\left(f_1 = 3.6kHz \right)$ occurs when n=1, the other modes are obtained for n being integers greater than one and υ is the speed of sound. The length of the canal lies typically in the range $2.1 \le \ell \le 2.6cm$. The ear drum is traditionally represented with an oscillator excited by the sound pressure from outside the ear and made to vibrate due to the combined actions of the hammer, anvil and stirrup.

3 Helmholtz Resonance Theory of Hearing

Resulting from understanding of the ear canal, the audio structure of the ear was conceived in the von Helmholtz' theory of hearing [8] and later explained by several authors [8, 9]. A typical structure of the frequency response of the cochlea originated by the theory of hearing is given in the encyclopaedia Britannica and reproduced in Figure 2 [7]. The sound pressure travelling through the canal reaches the ear drum which responds to it through the combined activities of the ear hammer, anvil and stirrup (stapes). The stirrup rocks back and forth and transmits the sound pressure into the cochlea. Since the cochlea is filled with fluid, the air vibrations from the ear drum are transmitted through the fluid vibrations in the cochlea.

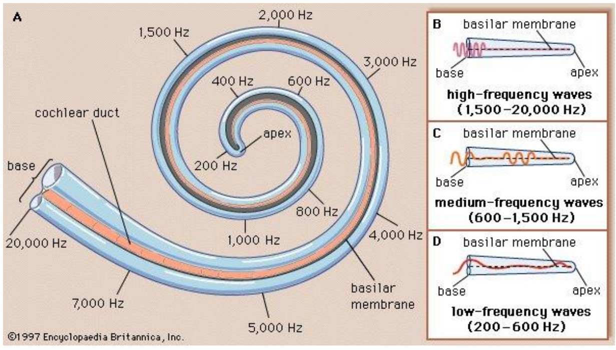

Figure 2: Acoustic structure of the human cochlea [7]

Hermann von Helmholtz's initial theory of the ear as a multi-resonant biological circuit inspired other researchers including leading to clinical and experimental investigations leading to the "place theory". The place theory explains that the sound waves at different frequencies activate "regions of the basilar membrane and organ of Corti" [7]. Subsequent experiments by Georg von Bekesy a Hungarian American led to the modern understanding of how the ear analyses acoustic signals leading to his Nobel Prize [9, 6].
According to Georg von Békésy [9] the ear distinguishes pitch due to continuous changes along the length of the cochlea. These changes occur in the form of increases in mass and width of the basilar membrane and decreases in stiffness from near the oval window down to the apex of the cochlea. Therefore, sections of the membrane become affected by specific frequency vibrations in proportion of their length (see equation (1)) and stiffness. At the apex of the cochlea, low frequency sounds are excited and towards the basal end of the cochlea high-frequency sounds are excited. High frequency waves travel through shorter distances and low frequency waves travel longer distances to be excited. The number of hair cells in the cochlea stimulated is thus a function of how vigorously the vibration is in those sections. The hair cells send nerve impulses to the auditory nerve and the brain which interprets the place on the basilar membrane and hence the tone and its pitch. Hence sorting of different frequencies is done in the cochlea. It also analyses the frequencies and the inherent energy in the sound vibrations. The frequency response of the cochlea is therefore an improvement on the Helmholtz theory of hearing leading to what is now understood as the acoustic spectrum of the cochlea (Figure 2) [9].

4 Resonance Inductive Model of the Cochlea

As shown by the author in [10] resonance frequency modes can be excited using both loose and strong coupling regimes in inductive systems to cover not only the audio frequency range but also higher frequencies. This section limits the consideration to the audio frequency range up to 20kHz. We therefore propose an inductive cochlea system as an array of resonant inductive atoms each one consisting of RLC circuits which resonate at different frequencies. The theory of how the human ear works was provided by Bekesy [9].

The ear drum and ossicles is thus modelled as a bank of resonant filters. While Figure 2 is provided for discourse and understanding of how the ear could support electrical waves, it is representative only of how the applied electrical wave travels into and out of the ear; we do not go further to show simulations as that is not the objective of this paper. The objective of this paper is to demonstrate how the signal received from this model could be used to determine the state of an ear and recognition of persons with the responses of their ears to external excitation.

The rest of the paper is organised as follows. Section 5 provides an ear theory based on principal components. Section 6 validates this theory through experiments by recording the responses of the ears to external audio excitations over a period. Details of the experiments are given in this section. Analysis and results are presented in section 7 with further proposal for how the state of an ear changes through time as a transformation model. Conclusions are draw in section 8.

5 Proposed Theory

The human ear has been the subject of numerous studies throughout history, none more so than recent interests in the use of radio frequency responses for determining identity of human beings. This paper provides insight into the status of an ear in relation to its health and application in biometric identity based on radio signals. The interest of the paper is in the determination of the ear status with reference to the y health of the ear over time as evidence of ear deterioration. This links directly to the effects of long-term use of earphones, long term exposure to high decibel noise as in factories, battle fields and listening to loud music. The paper compares methods based on spectral analysis of the ear response and principal component analysis scheme. Consider two N-dimensional vectors described by the following expressions:

$$\left.\begin{aligned} \bar{u}_1 &= \left(u_1, u_2, \cdots, u_N\right)_{t_1} \\ \bar{\upsilon}_1 &= \left(\upsilon_1, \upsilon_2, \cdots, \upsilon_N\right)_{t_2} \end{aligned}\right\} \quad (1)$$

Consider in the first place these vectors are formed from real numbers and from two experiments conducted at two different times t_1 and t_2 separated by some duration of time $T = \left(t_2 - t_1\right)$. Provided the components of the vectors are not all identically zero, the normalised vectors can be written as

$$u' = \frac{u}{|u|} \; and \; \upsilon' = \frac{\upsilon}{|\upsilon|} \quad (2)$$

Equation (2) establishes the magnitudes of the two vectors as the summations

$$|u| = \left(\sum_{j=1}^{N} u_j\right)^{\frac{1}{2}} \; and \; |\upsilon| = \left(\sum_{j=1}^{N} \upsilon_j\right)^{\frac{1}{2}} \quad (3)$$

These expressions (1) to (3) lead to the fundmental equations used in this work to establish similarity and the measure of disimilarity. The dot product of the vectors is essentially given by the covariance expression

$$\Gamma(u,\upsilon) = \left(u' \bullet \upsilon'\right) = \left(\frac{u}{|u|} \bullet \frac{\upsilon}{|\upsilon|}\right) = \frac{\left(u \bullet \upsilon\right)}{|u||\upsilon|} \quad (4)$$

We formally write this expression

$$\Gamma(u,\upsilon) = \frac{\left(\sum_{j=1}^{N} u_j \upsilon_j\right)}{\left(\sum_{j=1}^{N} u_j\right)^{\frac{1}{2}} \left(\sum_{j=1}^{N} \upsilon_j\right)^{\frac{1}{2}}} \quad (5)$$

The geometrical interpretation of the dot product between two vectors is illustrated in Figure 1 and comes from vector theory. Mathematically this is

$$\Gamma(u,\upsilon) = \cos\theta = \frac{\left(u \bullet \upsilon\right)}{|u||\upsilon|} \quad (6)$$

where θ is the angle between the two vectors $\bar{u}$ and $\bar{\upsilon}$. Therefore, when the two vectors are identical the angle between them is zero and the cosine of the angle between them is 1, or perfect correlation between them. In terms of what is happening to the ear, we draw the conclusion that if the ear performance has not deteriorated, the returned vectors from the measurement system within the two-recording times separated by a duration T must be the same. This conclusion is subject to two assumptions, that noise is not recorded during the two recordings and the

recording system is relatively stable in the reproduction of the state of the ear.
The variation of the cosine of the angle is considered in this paper as a measure of deterioration. This is for as long as the value remains positive. The measure of dissimilarity can also be estimated from the sign of the cosine angle. If the sign becomes negative, we infer that the two vectors come from the ears of two individuals.

6 Details of Experiments

A set of data was recorded at the King Khalid Specialist Hospital at Hail-Saudi Arabia three times over a period of about two years. The aim is to have a time gap between the three recordings to enable study of ear performance deterioration, stability and processing for person identification and recognition.
By taking a simple and intended random sample, 144 signals were recorded in 3 different sessions during a year and 4 records per session for a total of 12 subjects, and the following table shows test levels at different time periods as table (1).

Variable	Number	Notes
No of samples (subjects)	12	samples
No of sessions	3	variable
NO of recordings per session	4	Variable repeat rate
Total record	144	Sample test

Table 1: Number of Recordings for 12 Persons

The following steps were taken during recording. The person's ear was examined by an expert to assess if the person's hearing performance is deemed normal and wax-free (clean ear). Normally TEOAE signals are affected by the state of the tympanic membrane and middle ear. Adverse effects on the TEOAE results from the states of the tympanic membrane and middle ear. For example, an infected ear will lead to poor TEOAE. During the recording process a number of cases were rejected for these reasons.
The effects of the recording environment were assessed. Records were taken in environments or rooms with people conversing with background noise level below the 40dB suitable threshold. Subjects were in general sitting or lying down and thus have limited mobility to influence the acquired signals.
By permission from the King Khalid Specialist Hospital at Hail-Saudi Arabia, their portable Otoacoustic Emission device, the MADSEN Capella was used. The portable device provided easy USB-based connection to the USB port of a laptop or computer. Through the connection it is powered and also collects data from the human ear. Internally the MADSEN Capella contains a digital signal processor and allows real-time signal processing and internal audio amplifier. The device fits to the human ear through an audio transmitting and receiver probe. The ear stimulus sound level was kept below 80dB and allows normal audio frequency response up to 4kHz. For the TEOAE stimulus either a tone burst at unique frequencies, or a wider band click may be used.
To understand ear performance across female and male subjects, young or older persons, a mixture of subjects were invited for the tests. Over three (3) sessions and 4 tests, data was collected from twelve (12) subjects. Demographically, 33% of the subjects were male and 67% female. From historical perspectives, losses in ear performance is known to be much more prominent among the older generations. Therefore, to accommodate for varying ages, the number of subjects range from 8% within the age bracket of 20 to 30 years, 42% of the subjects were in the age bracket 31 to 40 years and the remaining 50% of the subjects were between the ages of 40 to 50 years. Through this selection process it is assumed that the physiological and psychological influences on hearing and ear performance may have been captured. The ear performance is known to be influenced by long term exposure to loud noise sessions (for example loud music and long-term use of earphones), form of employments where noise is prevalent. For these reasons, the number of subjects in this study include 55% teachers where noise is assumed minimal, 17% office staff and 25% truck drivers (assumed to be a demography where loud noise is to be expected in the work environment). In general, we have also assumed that the internal acoustic shape of an ear is maintained as we age. This unique feature may be useful in the use of ear as a biometric recognition system.

7 Results and Discussions

Using Matlab, the acquired TEOAE signals were analysed. Principal component analysis was undertaken to extract subject-based eigenvalues and eigenvectors. The recordings include six sessions for 12 persons at different times for the right and left ears. The first session was recorded in May 2018; the second session was recorded nine months later in March 2019 and the third session was recorded three months after the second session in June 2019.

Principal component analysis was undertaken on the transient evoked signals from the left and right ears of a person. The eigenvalues and eigenvectors were computed for both ears based on the independently recorded signals over three periods of time. The first data set was recorded in May 2018, the second in March 2019 and the third recording in June 2019. This represents a separation of twenty-five months between the first and third recording. The objectives were to find out how the two ears deteriorated over time, whether the deteriorations were independent and whether the signals provide indications recognising persons. Three pairs of ear signals were used to compute three sets of eigenvalues and eigenvectors. Table 1 shows the comparison of the angles between the sets of eigenvectors.

For the eleven persons shown in Table 1, the angles between pairs of angles were identical for the left and right ears (row-wise values). The angles thus show that for each person the angle between the pairs of signals were distinct and can be used for recognizing each person. There were changes in the angles between the ear eigenvector pairs taken from recordings many months apart. They thus also show deterioration of the ear audio performances over time.

Table 2: Person Recognition Angles Using Eigenvectors ***(F=female, M=Male)***

No	Person	Ear	Angle (v0, v1) (May-18,Mar-19)	Angle (v0, v2) (May-19,Jun-2019)	Angle (v1, v2) (Mar-19,Jun-2019)
1	F1	Left	2.8809	4.1197	1.2388
		Right	2.8809	4.1197	1.2388
2	F2	Left	4.0348	4.0348	0
		Right	4.0348	4.0348	0
3	F3	Left	162.211	16.6254	178.8364
		Right	162.211	16.6254	178.8364
4	F4	Left	0	180	180 (0.3391)
		Right	4.656	174.8558	179.5118 (0.3391)
5	F5	Left	173.4429	174.7247	1.2819
		Right	173.4429	174.7247	1.2819
6	F6	Left	19.0271	128.8482	109.8211
		Right	19.0271	128.8482	109.8211
7	M1	Left	0.3039	177.0753	177.3792
		Right	0.3039	177.0753	177.3792
8	M2	Left	34.6229	1.6191	36.242
		Right	34.6229	1.6191	36.242
9	F7	Left	172.3992	2.1658	170.2334
		Right	172.3992	2.1658	170.2334
10	F8	Left	29.7781	26.4967	3.2815
		Right	29.7781	26.4967	3.2815
11	F9	Left	5.7897	177.6795	171.8898
		Right	5.7897	177.6795	171.8898
12	F10	Left	164.9722	0.838	164.1342
		Right	164.9722	0.838	164.1342

8 Conclusions

This paper has applied data analytics algorithms for assessing biometric features of human ear response to external excitation. Principal component analysis was undertaken on transient evoked signals from the left and right ears of a person. The eigenvalues and eigenvectors were computed for both ears based on independently recorded signals over three periods of time. Results show that a person can be identified through the eigenvectors for the ears. For a person, the eigenvectors and the angles between the eigenvectors are maintained even when the ear responses were measured over a year apart. The ear eigenvectors differ for individuals. In cases where there is deterioration of the state of an ear (or both), the eigenvectors and angles between them also differ. Further development to use the eigenvectors and angles between them for biometric recognition of persons is ongoing.

9 References

[1] Yi-Wen Liu, Sheng-Lun Kao, Hau-Tieng Wu, Tzu-Chi Liu, Te-Yung Fang, Pa-Chun Wan, "Transient-evoked otoacoustic emission signals predicting outcomes of acute sensorineural hearing loss in patients with Ménière's disease", Acta-Laryngologica, Volume 140, 2020, pp. 230 – 235.
[2] Yanxia Zhao, Jingsong Li, Meibian Zhang, Yao Lu,1 Hongwei Xie, Yu Tian, and Wei Qiu, "Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial Noise: A Pilot Study", Ear and Hearing, Vol. 40, 2019, pp. 690–699.
[3] Crowson, Matthew G.; Franck, Kevin H.; Rosella, Laura C; Chan, Timothy C. Y., "Predicting Depression From Hearing Loss Using Machine Learning", Ear and Hearing: Volume 42, Issue 4, July/August 2021, pp. 982-989.
[4] Gregory M. Ellis and Pamela E. Souza, "Using Machine Learning and the National Health and Nutrition Examination Survey to Classify Individuals With Hearing Loss", Frontiers in Digital Health, https://www.frontiersin.org, August 2021, Volume 3, Article 723533, pp. 1-9.
[5] G. von B´ek´esy, Experiments in Hearing, McGraw-Hill, New York, 1960.
[6] https://www.britannica.com/science/ear/Transmission-of-sound-waves-through-the-outer-and-middle-ear; accessed April 26, 2021
[7] HARTRIDGE, H. The Resonance Theory of Hearing. *Nature* **110,** 9–10 (1922). https://doi.org/10.1038/110009b0
[8] Georg von Bekesy, https://www.britannica.com/biography/Georg-von-Bekesy
[9] Robert H. Withnell and Lauren E. Gowdy, " An Analysis of the Acoustic Input Impedance of the Ear", Journal of the Association for Research in Otalaryngology (JARO),Vol 14: 611–622 (2013) DOI: 10.1007/s10162-013-0407-y, August 2013, pp. 611 – 622.