THear: Development of a Mobile Multimodal Audiometry Application on a Cross-Platform Framework

Wai-Kim Leung[1,2,3]*, Jia Jia*[1,2,3]*, Yuhao Wu*[1,2,3]*, Jiayu Long*[2,3,4] *and Lianhong Cai*[1,2,3]

[1]Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[2]Key Laboratory of Pervasive Computing, Ministry of Education
[3]Tsinghua National Laboratory for Information Science and Technology (TNList)
[4]Department of Information Art & Design, Tsinghua University, Beijing 100084, China

liangwj14@mails.tsinghua.edu.cn, jjia@mail.tsinghua.edu.cn
yh-wu14@mails.tsinghua.edu.cn, longjy14@mails.tsinghua.edu.cn and clh-dcs@mail.tsinghua.edu.cn

Abstract

This paper presents our ongoing research in the field of audiometry application development. There are many researches on the development of audiometry application but they are all platform specific. In this paper, we developed a new multimodal framework that enabled cross-platform development by using open standards such as HTML5, CSS5 and JavaScript. We will describe the architecture of the new multimodal framework which supports the conduction of the automatic audiometry. Based on the framework, we further propose three multimodal speech audiometry approaches for supporting automatic speech audiometry on mobile devices. By using the framework combined with the three multimodal speech audiometry approaches, we developed an application called THear (TsingHua ear) to support the conduction of pure-tone audiometry and speech audiometry on mobile devices. A preliminary performance evaluation on THear will also be reported.

Index Terms: audiometry, pure-tone audiometry, speech audiometry, multimodal framework, mobile application, cross-platform, e-health

1. Introduction

Speech chain consists of speech perception and speech production which plays a very important role in communication in daily life. There are several researches [1], [2], [3] indicated that the speech production could be affected by poor perception. In recent decades, Walkmans, MP3 players and mobile devices are getting more prevalent. It is very easy to observe that more people, especially teenagers, keep using earphones to listen to the music. Unfortunately, a research [4] found out that the inappropriate duration of exposure to high volume level could damage our hearing and the damage was not recoverable. Hearing impairment, or known as hearing loss, brings communication difficulty and inconvenience to daily life. Therefore, we should pay attention to protecting our hearing.

In audiology, Pure-Tone Audiometry (PTA) is the most traditional way to evaluate the hearing level. The aim of PTA is to identify the hearing threshold of an individual by using audiometer to play the pure-tone signal of different frequencies (usually from 250Hz to 8000Hz) [5] with different sound intensities (in dBHL). The results can then be plotted on an audiogram. The normal hearing thresholds should be less than 25 dBHL across all frequencies [6]. On the other hand, Speech Audiometry (SA) is also actively used for measuring ability in speech understanding.[7] SA determines the hearing level of an individual by playing a list of words at a certain sound intensity. The audiologist will ask the subject to repeat the words to determine the Speech Reception Threshold (SRT). SRT is defined as the lowest level in dB at which 50% of the words were repeated correctly by the subject. Some hearing diseases such as Presbyacusis can be diagnosed by PTA on audiogram (significant increase in hearing threshold at high frequency). For patients with Auditory Neuropathy(AN), their hearing thresholds on audiograms are usually flat like a normal hearing person but SA can help to distinguish them. Therefore, PTA and SA are the main audiometries to measure an subject's hearing level.

Traditional PTA and SA are conducted in hospital since it requires professional equipment such as audiometer. Fortunately, the development of mobile devices grows rapidly in recent years. A modern mobile device usually comes with many sensors including camera, microphone, speaker, accelerometers, gyroscopes, etc which can be used for healthcare[8][9][10][10] and e-learning[11]. The conduction of PTA on mobile devices are widely researched. Ayag et al. [12] proved the standard deviation of sound level output across different Android devices would be varied from 1.28 dB to 3.99 dB. Similar researches on iOS devices were also conducted [13], [14] and [15]. Larrosa et al. [13] proved the difference between mobile audiometry app on iPhone/ iPad and the standard audiogram thresholds was -0.21 $\pm$ 6.38 dB HL[13]. However, all of these researches are device specific which is not cross-platform compatible. SA is mainly conducted in hospitals so there are not many researches on the conduction of automatic speech audiometry on mobile devices or computers. Dutch speech-in-noise screening test [16] and French speech intelligibility in noise [17] were developed in 2002 and 2010 respectively. Two speech audiometries were conducted by telephone. The correlations between the SRTn test by telephone and pure-tone average were 0.732[16] and 0.77[17]. Over 20,000 calls were registered in four months after the launch of the two automatic speech-in-noise screening tests . This shows that there is a huge demand for the automatic speech audiometry in the society. An interactive multimodal speech audiomety is important since subjects are required to respond to their answers by clicking, typing or speaking to support the conduction of speech audiometry on mobile devices. There are not many researches in this field. In this paper, we will propose three multimodal speech audiometries on mobile device.

There are three main mobile operation systems (OS) nowa-

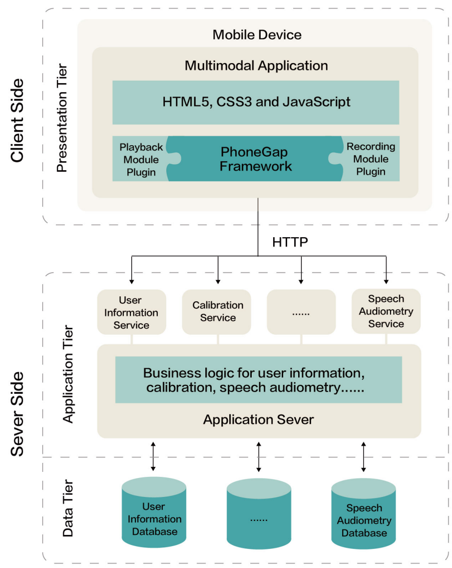

Figure 1: *The multimodal cross-platform framework for audiometry on mobile devices.*

days which are Apple's iOS, Google's Android and Microsoft's Windows Phone. However, their applications are platform independent. As a result, the cost of application development is expensive in order to support all three operation systems. Liu et al. leveraged the cross-platform advantage of HTML-based approach to develop a cross-platform computer-aided pronunciation training system for mobile devices[11]. By taking the advantage of cross-platform framework, a mobile audiometry framework can be developed to support the majority of mobile devices available in the market at minimal cost. The above approach is helpful to minimize the cost to promote the importance of protecting hearing to all people.

In this paper, we focus on three main contributions. 1) We will introduce our new multimodal cross-platform audiometry framework which supports different mobile devices including Audroid Phone and Apple iPhone. 2) Based on the framework, we further propose three multimodal speech audiometry approaches to support the conduction of audiometry automatically. 3) Finally, we developed our mobile, multimodal and cross-platform audiometry application for mobile devices. An evaluation had been conducted to prove the accuracy of our demo application.

2. Our framework and approach

2.1. Framework Architecture

The proposed mobile audiometry framework is based on a 3-tier client-server architecture shown in Figure 1.

The client side acts as the presentation tier and provides the multimodal user interface for user interaction. It resides on mobile devices such as Apple iPhone/iPad, Android Phone/Tablet, etc. The multimodal mobile application is installed at the client side. PhoneGap framework is used at the client side since it is based on the open-standard HTML5, CSS3 and JavaScript which provide cross-platform capability. The Playback Module and Recording Module are implemented as PhoneGap Plugins. The plugins include two parts: (1) the JavaScript APIs which are used to access the native code and (2) the native code which is called by the Javascript APIs.

The server side consists of the data tier and the application tier. The data tier includes databases which are used for storing user information, user audiometry history, calibration parameters for Playback Module and the speech audiometry word list with its corresponding audio files. The application tier contains an application server. The application server is connected to the databases for storing and retrieving information. It is also responsible for making logical decisions and performing calculations. At the same time, the application server exposes the services requested by the client side as Web APIs which allow the client to access through HTTP. For example, after the application server had received the speech audiometry answers from the client side, it would retrieve the model answers from the database and compute the speech audiometry results. The results would be encapsulated and responded to the client side.

The web services are RESTful that client access is allowed by Uniform Resources Identifier (URI) and perform “Create”, “Read”, “Update” and “Delete” access. The messages between the client and the server are in Javascript Object Notation (JSON) format which is supported by many programming languages.

2.2. Two Modules to Support Multimodal Audiometry

There are two important modules to support audiometry. The first module is called Playback Module. The main responsibility of this module is to play the audio clips or pure-tone signals at various sound intensities accurately by loading the model's specific calibration parameters from the server. Code 1 shows the pseudocode of this module. The second module is called Audio Recording Module. Audio Recording Module records the sound to our system. The recorded sound file can be used for analysing the background noise level in order to determine the accuracy of the audiometry results. This module can also be used for recording the user's answers during the speech audiometry. The audio will then be sent to the server and processed by the automatic speech recognizer or the audiologist remotely.

Code 1 Pseudocode of Playback Module

```
procedure PLAYAUDIOATDB
    filename ← audio Filename
    sourceDb ← orginal audio dB
    targetDb ← target audio dB
    calParams ← Calibration Params of devices
top:
    factor = calculateAmplificationFactor(sourceDb, targetDb, calParams) ← Calculate amplification factor
    samples = readAudioSamples(filename)
    newSamples = new array()
main:
    for each sample in samples do
        newSample = sample * factor
        newSamples.push(newSample)
final:
    playAudioSamples(newSamples)
```

2.2.1. *Implementation of Pure-Tone Audiometry*

These are standard procedures for measuring hearing threshold [5]. Pure-tone signals with various sound intensities are played accordingly and the subject needs to respond “hear” or “not hear”. The pure-tone signals should be played for 1 to 2 seconds[5]. The sound intensity will be reduced by 10 dB if the subject heard the signal, otherwise the sound intensity will be increased by 5 dB[5]. This step is repeated until the threshold is determined. Our pure-tone audiometry section provides two button for user response. After all the hearing thresholds across different frequencies had been determined, the thresholds would be plotted on an audiogram as the results.

2.2.2. Proposal of Three Approaches for Speech Audiometry

Traditional speech audiometry is conducted in hospitals. Audiologist uses calibrated player to play the speech signals at various sound intensities. The subject needs to repeat the words they heard and the audiologist needs to manually mark the answers on the list for word recognition rate calculation. We propose three multimodal approaches to overcome this manual process. (1)The first way is to allow the subject to choose the answer from multiple choices. (2)The second way is to allow the subject to type the answers in the textbox. The input text would then be sent to the server and compared with the model answers. (3)The third way is to allow the subject to record their answers. The voice would then be sent to the server for speech recognition or manually recognized by professionals. Our speech audiometry section provides the above three ways for the subject to conduct speech audiometry. After finishing the speech audiometry, the corresponding word recognition rate will be plotted on the speech audiogram as a function of Performance Intensity.

3. The THear Application

THear is developed based on our multimodal framework mentioned above for audiometry on mobile devices.

At the server side, we use MongoDB as our database to store user information, audiometry results, the sound intensity calibration parameters and the word lists for speech audiometry. The audio clips of the speech audiometry are saved as files on the server. Apache Tomcat is used as the application server. All the services which includes user information service, user audiometry history service, sound intensity calibration service and speech audiometry service are exposed as RESTful services by Apache Tomcat.

At the client side, PhoneGap is used for providing cross-platform mobile application. Playback Module Plugin and Recording Modules Plugin are developed to provide the necessary functionality for conducting the audiometry. HTML5, CSS3 and JavaScript are used for building the user interface. The demonstration of each page is shown in Figure 2 and the functionalities of each page are described as follows:

1. User account page (Figure 2a): It allows users to create or login their accounts to our system. Users are required to fill in their basic information such as name, gender, year of born, occupation, whether he/she is wearing hearing aids, etc. This information is useful for statistical purpose.
2. User audiometry history page (Figure 2b): All the historical audiometry can be reviewed on this page. Users can track the changes in their hearing levels over time. They can also check the dates and the detailed results of each audiometry.
3. Noise-level measurement page (Figure 2c): The background noise certainly affects the accuracy of audiometry. The background noise is sampled and analysed by Recording Module on this page. The result of the analysis is used for determining the accuracy of audiometry.
4. Pure-tone audiometry page (Figure 2g): A series of pure-tone at different frequencies and sound intensities will be played accordingly. Users give their responses by clicking the “Hear” or “Not hear” button. Playback Module with calibration parameter is used on this page to ensure the accuracy of the output sound intensity.
5. Speech audiometry page (Figure 2d, Figure 2e, Figure 2f): A list of audio clips in the speech audiometry word list will be played at various sound intensities. According to the design of the speech audiometry, users can respond to the words they heard by 1)Selecting the answers from multiple choices, 2)Typing the answers 3)Recording the answers. The answers are streamed to the server to compute the speech audiometry results.
6. Audiometry result page (Figure 2h, Figure 2i): The pure-tone audiometry results and the speech audiometry are reported on this page. Pure-tone audiometry results with hearing threshold across different frequencies are reported as audiogram. Word recognition rate and the performance-intensity function are the results of the speech audiometry.

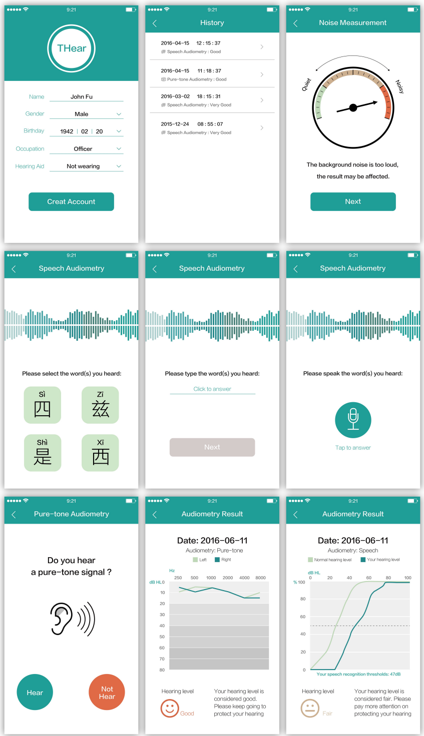

Figure 2: *The pages of (a, top-left) user account, (b, top-middle) audiometry history, (c, top-right) noise-level measurement, (d, middle-left) speech audiometry by multiple choices, (e, middle-middle) speech audiometry by typing answer, (f, middle-right) speech audiometry by recording, (g, button-left) pure-tone audiometry, (h, button-middle) pure-tone audiometry results, (i, button-right) speech audiometry results.*

4. Pilot Evaluation of THear

12 university students aged 20-25 were invited to conduct pilot evaluation of THear. They were all with normal hearing. In the first two evaluations, pure-tone audiometry and speech audiometry were evaluated individually. THear was installed on Apple

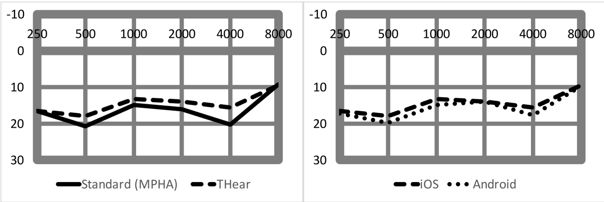

Figure 3: *(a, left)The average hearing thresholds of 12 subjects by using MPHA and THear. (b, right)The average hearing thresholds of 12 subjects by using THear on iOS and Android.*

iPhone6 Plus and subjects were required to wear the Apple Earpod (come with stock Apple iPhone6) during the evaluations. In the third evaluation, THear was installed on Sony Xperia Z (Android 5.1.1) and the phone was used for conducting pure-tone audiometry. The results were compared with the results of using Apple iPhone to conduct the evaluation.

4.1. Pure-tone Audiometry Evaluation

Firstly, 12 subjects conducted the pure-tone audiometry with MPHA, which is a calibrated iPad audiometry application published by Wu et al.[15], in a quiet office room (background noise is 30 - 35dB SPL) to obtain their hearing thresholds. Then they conducted the pure-tone audiometry with THear in the same place (Figure 2g). In this experiment, the hearing thresholds at 6 frequencies (250, 500, 1000, 2000, 4000, 8000 Hz) were measured.

The average of the two pure-tone audiometry results are shown in Figure 3. The differences of the two approaches were -1.7dB, -2.7dB, -1.6dB, -1.95dB, -4.6dB, +0.3dB at 6 frequencies (250, 500, 1000, 2000, 4000, 8000 Hz) respectively. The differences were not greater than 5dB which was the smallest interval unit in the audiometry. In addition, the shapes of the minimal audibility curves were also the same. It proved the accuracy and validity of pure-tone audiometry of THear.

4.2. Speech Audiometry Evaluation

The speech audiometry using the multiple choices approach was carried out (Figure 2d). All the subjects were required to conduct the speech audiometry with 20 different sets of word lists. For each of the word list, there were 20 questions and each question consisted of 4 choices. For each of the question, a monosyllable Mandarin speech was played at a specific sound intensity once. The subjects needed to select the word they heard on the screen. The speech audiometry experiment was repeated with 5 sound intensities, which were 1dB HL, 5dB HL, 10dB HL, 20dB HL and 30dB HL. The experiment was conducted in a quiet office room (background noise is 30 - 35dB SPL)

The average correct rates in different sound intensities were calculated. We randomly separated all the subjects' results into two groups. We treated the first group as the standard group and the second group as the testing group. The average correct rate for each subject of the first group was calculated and plotted on a speech audiogram. The points were connected and acted as the standard normal-hearing performance-intensity (PI) function. For the testing group, we plotted each of their PI function on the speech audiogram. Since all of them were with normal hearing, their PI function were similar to the standard normal-hearing PI function as expected. Figure 4 shows the performance-intensity function of standard group and the 6 subjects in testing group. The result proved that THear is accurate in speech audiometry.

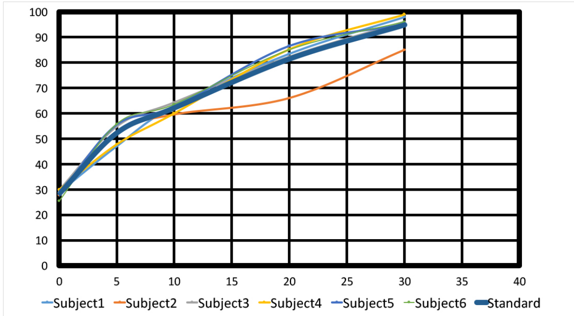

Figure 4: *The performance-intensity function of standard group and the 6 subjects in testing group.*

4.3. Cross-platform Evaluation

To prove the reliability of our application across Android phone and Apple iPhone, we compiled and installed the THear on an Android phone (Sony XperiaZ, Android 5.1.1). We invited the same 12 subjects mentioned above to conduct pure-tone audiometry in a quiet office room(background noise is 30 - 35dB SPL, same as above) by using the Android phone. In this experiment, the same earphone (Apple Earpod) was used.

Figure 3b shows the average hearing thersholds measured by THear on Sony Xperia Z (Android) and Apple iPhone 6 plus (iOS). The differences of two version of THear were 0.7dB, 1.7dB, 1.5dB, 0.0dB, 2.2dB, 0.4dB at 6 frequencies (250, 500, 1000, 2000, 4000, 8000 Hz) respectively and the average difference was 1.1dB. In addition, the shapes and the trends of the minimal audibility curves were also the same. It proved the reliability of THear across different mobile OS.

5. Conclusion

This paper proved the possibility of developing an automatic pure-tone audiometry and speech audiometry for mobile devices. The cross-platform issue was discussed and resolved by using client-server based multimodal framework. In this framework, open standards such as HTML5, CSS5 and JavaScript are used. A multimodal, cross-platform audiometry application, THear, is developed for mobile devices based on the above framework. Comparing THear with the traditional audiometry, professional equipment is not required. This creates the possibility of conducting pure-tone audiometry and speech audiometry all around the clock for the general public. This application is a good tool that helps to reduce the workload of doctors or audiologists by screening the patients.

In the future, the application will be launched to the general public for helping people who have concerns about their hearing. More subjects will be invited to improve the pilot application.

6. Acknowledgements

This work is supported by the National Basic Research Program (973 Program) of China (2013CB329304), National Natural, and Science Foundation of China (61370023). This work is also partially supported by the major project of the National Social Science Foundation of China(13&ZD189)

7. References

[1] D. Shankweiler, W. Strange, and R. Verbrugge, "Speech and the problem of perceptual constancy," *Perceiving, acting, and knowing: Toward an ecological psychology*, pp. 315–345, 1977.

[2] E. Gerrits and E. de Bree, "Early language development of children at familial risk of dyslexia: Speech perception and production," *Journal of Communication Disorders*, vol. 42, no. 3, pp. 180–194, 2009.

[3] K. Nagao, M. Paullin, V. Livinsky, J. B. Polikoff, L. D. Vallino, T. G. Morlet, N. C. Schanen, and H. T. Bunnell, "Speech production-perception relationships in children with speech delay," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.

[4] B. J. Fligor and L. C. Cox, "Output levels of commercially available portable compact disc players and the potential risk to hearing," *Ear and hearing*, vol. 25, no. 6, pp. 513–527, 2004.

[5] A. S.-L.-H. Association *et al.*, "Guidelines for manual pure-tone threshold audiometry," 2005.

[6] C. W. Newman, B. E. Weinstein, G. P. Jacobson, and G. A. Hug, "The hearing handicap inventory for adults: psychometric adequacy and audiometric correlates." *Ear and hearing*, vol. 11, no. 6, pp. 430–433, 1990.

[7] I. J. Hirsh, H. Davis, S. R. Silverman, E. G. Reynolds, E. Eldert, and R. W. Benson, "Development of materials for speech audiometry," *Journal of Speech and Hearing Disorders*, vol. 17, no. 3, pp. 321–337, 1952.

[8] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *Communications Magazine, IEEE*, vol. 48, no. 9, pp. 140–150, 2010.

[9] P. Pelegris, K. Banitsas, T. Orbach, and K. Marias, "A novel method to detect heart beat rate using a mobile phone," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 5488–5491.

[10] S. Gradl, P. Kugler, C. Lohmuller, and B. Eskofier, "Real-time ecg monitoring and arrhythmia detection using android-based mobile devices," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 2452–2455.

[11] P. Liu, K.-W. Yuen, W.-K. Leung, and H. Meng, "menunciate: Development of a computer-aided pronunciation training system on a cross-platform framework for mobile, speech-enabled application development," in *Chinese Spoken Language Processing (ISCSLP), 2012 8th International Symposium on*. IEEE, 2012, pp. 170–173.

[12] D. A. F. Ayag, R. C. Bautista, M. A. T. D. Eala, and R. P. Feria, "Acceptability of hear-o as a mobile hearing screening tool," in *Information, Intelligence, Systems and Applications (IISA), 2013 Fourth International Conference on*. IEEE, 2013, pp. 1–6.

[13] F. Larrosa, J. Rama-Lopez, J. Benitez, J. M. Morales, A. Martinez, M. A. Alañon, D. Arancibia-Tagle, A. Batuecas-Caletrio, M. Martinez-Lopez, N. Perez-Fernandez *et al.*, "Development and evaluation of an audiology app for iphone/ipad mobile devices," *Acta oto-laryngologica*, vol. 135, no. 11, pp. 1119–1127, 2015.

[14] S. Peer and J. J. Fagan, "Hearing loss in the developing world: Evaluating the iphone mobile device as a screening tool," *SAMJ: South African Medical Journal*, vol. 105, no. 1, pp. 35–39, 2015.

[15] Y. Wu, J. Jia, W. Leung, Y. Liu, and L. Cai, "Mpha: A personal hearing doctor based on mobile devices," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 2015, pp. 155–162.

[16] C. Smits, T. S. Kapteyn, and T. Houtgast, "Development and validation of an automatic speech-in-noise screening test by telephone," *International journal of audiology*, vol. 43, no. 1, pp. 15–28, 2004.

[17] S. Jansen, H. Luts, K. C. Wagener, B. Frachet, and J. Wouters, "The french digit triplet test: A hearing screening tool for speech intelligibility in noise," *International Journal of Audiology*, vol. 49, no. 5, pp. 378–387, 2010.