33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Study on Self Hearing Assessment Using Speech Sounds

Junil Sohn(1), Dongwook Kim(1), Yunseo Ku(1), Kyungwon Lee(2), Junghak Lee(2)

Abstract— **In this study, we proposed new self assessment of hearing loss in mobile phones and realized a function of compensation for hearing impaired person. The results of experiments on mobile phone showed that the proposed hearing test is sufficient to check hearing loss and the compensation based on the result of the proposed hearing test can improve speech intelligibility of hearing impaired persons.**

I. Introduction

A lot of people have mobile phones and needs for better sound quality in mobile phones have increased especially in various noisy environments. The improved sound quality gives elderly people great benefit like the enhanced speech intelligibility. For enhancing speech intelligibility, a self assessment of user's hearing loss is a very essential process, but it is difficult that the traditional test with high fidelity like using various pure tones is applied in mobile phones. Because users have to spend a long time to respond to stimulation with various frequencies and amplitude repetition, mobile phone is not good system for the hearing test with high fidelity, and users do not have knowledge of audiology.

In this study, we proposed new self assessment of hearing loss and realized function of compensation for hearing impaired person and showed that the proposed hearing test is sufficient to check hearing loss and the compensation based on the result of the proposed hearing test can improve speech intelligibility of hearing impaired persons, through experiments using mobile phones.

II. The Assessment of Hearing Loss

A. The traditional hearing test using pure tones

Hearing impaired persons take a hearing test for checking their hearing loss and fitting hearing aid. Normal persons also take a hearing test for checking whether or not their hearing is normal.

The traditional hearing test using pure tones is conducted in a sound-proof testing booth by audiologist. In general, the sequence of frequencies is 1000Hz-2000Hz-4000Hz-8000Hz and then 1000Hz -500Hz-250Hz-125Hz [1]. There are a few methods to give stimuli to subject like descending method, ascending method and mixed method according to the sequence of sound levels [2, 3]. For example, in mixed method, an audiologist gives pure tone with 1000Hz at a comfortable sound level. If a subject can hear this tone, the sound level of the next stimulus is adjusted by -10dB SPL, otherwise, the sound level of the next stimulus is adjusted by +5dB SPL. The sound level at which subject can hear two times in the three same stimuli is recorded as hearing threshold. The audiologist conducts the same process in 2000Hz, 4000Hz, and so on. Therefore the traditional hearing test using pure tones requires long time to measure degree of hearing loss because users should repeatedly response to so many stimuli with various frequencies and sound levels.

B. The proposed self assessment of hearing loss

We propose a new self assessment of hearing loss that requires very short time to test so that non-specialists can easily check their hearing loss using mobile phones and finally hear the enhanced speech or music by signal processing based on user's hearing loss in mobile phones.

In the proposed self assessment of hearing loss, pure tones that are generally used in traditional hearing test are replaced by four phonemes. It seems to be reasonable that phonemes are used as test sound, because phonemes have various formants and spectral energy distribution and have much more information than pure tone.

We chose four phonemes, /a/, /i/, /sh/ and /s/ as test sounds. The /a/ has three formants at 710Hz, 1100Hz and 2640Hz and /i/ has also three formants at 400Hz, 1900Hz and 2550Hz [4], while /s/ and /sh/ has distinguishable area from other phonemes between 3000~4000Hz and 2500~4000Hz, respectively. Hearing loss of users can be obtained by combination of hearing level for these phonemes.

III. Realization of Multi-Channel Wide Dynamic Range Compression

Hearing impaired persons cannot hear silent sounds, while they can hear loud sounds as well as normal person. This means that hearing impaired persons have relatively narrow dynamic range when compared with normal person. Further, degrees of hearing loss are different in each frequency bands. Therefore, wide dynamic range of input signal has to be frequency-selectively compressed for compensating hearing loss. Figure 1 shows block diagram that realized in this study.

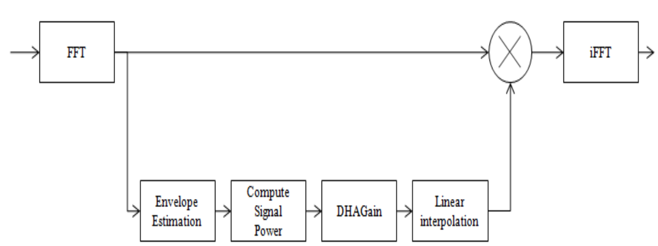

Figure 1. The block diagram of wide dynamic range compensation

Junil Sohn, Dongwook Kim and Yunseo Ku are with Samsung Advanced Institute of Technology (SAIT), Korea(e-mail :ji.sohn@samsung.com)

Kyungwon Lee and Junghak Lee are with Hallym University of Graduate Studies, Korea(e-mail : leejh@hallym.ac.kr)

In order to separate input signal into multi-channel signals, we used Modified Discrete Cosine Transform (MDCT) which is based on a DCT of overlapping data. MDCT has been used in sub-band processing due to advantage like guarantee of perfect reconstruction [5, 6, 7]. Output of MDCT as analysis filter, $X(k)$ is given by

$$X(k) = \sum_{n=0}^{2M-1} w(n)x(n)\cos\left(\frac{(2n+M+1)(2k+1)\pi}{4M}\right) \quad (1)$$

Where block size L = 2M and $k = 0, 1, \cdots, M-1$ and window function $w(n)$ have to satisfy following condition.

$$w^2(n) + w^2(n+M) = 1, \quad n = 0, 1, 2, \cdots, M-1 \quad (2)$$

Based on the outputs of MDCT, envelopes in each channel and signal power in each band are calculated. Additionally, we defined a band is a bundle of channels and the optimal numbers of channels for each band can be previously designed according to hearing loss.

Actually we used 64 channels MDCT as analysis filter in our system. A number of channels which is the most distinct key point of our system give us great benefits, such as flexibility of band structure and minimization of fitting error. Generally, too much difference of gains between two neighboring bands can cause distortion of the reconstructed signal by synthesis filter. But we can change the number of channels in each band and gradually set the gain of each channel in two neighboring bands because our system has a lot of channels.

Based on the input signal power and user's hearing loss in each band, we determine gain for each band and then apply the extended gains for 64 channels by linear interpolation into 64 outputs of MDCT. Finally, inverse MDCT generates the reconstructed signal from the amplified 64 signals. The output of inverse MDCT as synthesis filter is given by

$$y(n) = w(n)\left(\frac{1}{M}\right)\sum_{k=0}^{M-1} X(k)\cos\left(\frac{(2n+M+1)(2k+1)\pi}{4M}\right) \quad (3)$$

Where $n = 0, 1, 2, \cdots, 2M-1$.

IV. Experiments

A. *Feasibility verification of the proposed hearing test*

Before we realize the proposed hearing test on mobile phones, we executed feasibility test in a soundproof room. The four phonemes were played back to subjects using headphone through an audiometer. The sound level of each phoneme has range between 20 and 70 dB SPL with 5dB SPL interval. We measured minimum sound levels that subjects can hear for each phoneme and extracted relations between audiograms of subjects and the measured threshold levels for four phonemes. The extracted relations can be explained by

$$\begin{aligned} HL_{500} &= 0.91(TH_a+TH_i)/2+2.4 \\ HL_{1000} &= 0.92(TH_a + TH_i + TH_s)/3-3.09 \\ HL_{2000} &= 0.84(TH_a + TH_{sh} + TH_s)/3+2.01 \quad (4) \\ HL_{4000} &= 0.91(TH_{sh} + TH_s)/2+3.33 \end{aligned}$$

where HL_f is the estimated audiogram at frequency f and TH_a, TH_i, TH_s and TH_{sh} means the measured threshold level for test sounds /a/, /i/, /s/ and /sh/ , respectively. We tested 24 normal ears and 32 ears with mild or moderate hearing loss using audiometer in soundproof room. Figure 2 shows that the estimated hearing loss using these relations is close to audiogram which was previously measured using pure tones. Through this figure, we can know that mean error was 8.7dB and the average time required was as short as 57 seconds. Therefore, the proposed hearing test is sufficient to simply check hearing loss.

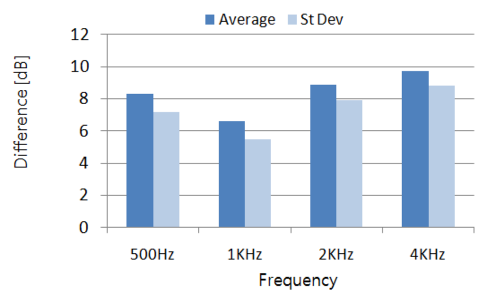

Figure 2. Difference between estimated audiogram and real audiogram

B. *Implementation and experiments of the proposed hearing test on mobile phones*

We realized the proposed hearing test on mobile phones, Samsung SCH-M175. Four phonemes, /a/, /i/, /sh/ and /s/ were separately recorded from 30 dB HL to 65 dB HL with an interval of 5 dB at 44.1KHz sampling rate by female voices.

TABLE I
REFERENCE EQUIVALENT THRESHOLD SOUND PRESSURE LEVELS FOR INSERT EARPHONES RECOMMENDED BY ANSI S3.6-1996

Frequency (Hz)	Coupler Type		
	Occluded Ear simulator	HA-2 with Rigid Tube	HA-1
125	28.0	26.0	26.5
250	17.5	14.0	14.5
500	9.5	5.5	6.0
1000	5.5	0.0	0.0
2000	11.5	3.0	2.5
4000	15.0	5.5	0.0
8000	15.5	0.0	-3.5

We chose earphone with relatively flat frequency response and adjusted sound level of output in each band to fit Reference Equivalent Threshold Sound Pressure Levels(RETSPLs) recommended by ANSI S3.6-1996.

65 ears were included in the experiment using mobile phones. Of these, 38 ears had normal hearing sensitivity (mean age 28.89 yrs, SD 4.28) and 27 ears had hearing loss (mean age 59.22 yrs, SD 19.64).

Regression formulae were selected based on the highest correlation between phoneme-estimated thresholds and pure tone estimated thresholds for each test frequency as shown in table 2. Where, a_{th} and sh_{th} means phoneme-estimated threshold of /a/ and /sh/, respectively. The combinations of /a/ and /sh/ showed relatively higher correlation than all other possible combinations.

TABLE II
REGRESSION FORMULAE

Frequency (Hz)	Regression formula
500	$0.73a_{th} + 4.28$
1000	$0.91a_{th} - 0.84$
2000	$0.97(a_{th}+sh_{th})/2 - 2.48$
4000	$0.98sh_{th} - 2.37$

Figure 3 shows the accuracy of the proposed hearing test in screening for hearing loss. In this figure, the phoneme-estimated thresholds show the highest accuracy in screening the hearing loss at the criterion level of 35 dB HL with respect to the sensitivity and specificity. Positive predictive value (PPV) and negative predictive value (NPV) were also high at this level.

The differences between phoneme-estimated threshold and pure tone thresholds are shown in the Figure 4. From this result, we can know that the differences between phoneme-estimated thresholds and pure tone thresholds were less than 5 dB for all test frequencies and the average difference was 3.29 dB.

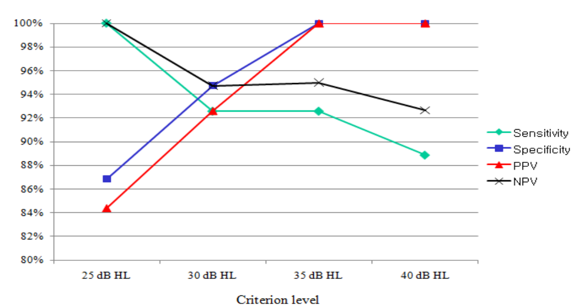

Figure 3. The accuracy of the proposed hearing test in screening for hearing loss

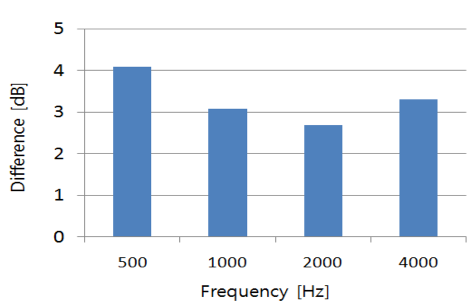

Figure 4. Differences between phoneme-estimated thresholds and pure tone thresholds for 4 test frequencies

We measured test time that is required for subjects to perform the hearing test for the only 2 phonemes, /a/ and /sh/. In this experiment, each subject carried out hearing test twice for left and right ear, if both ears are not deaf. Figure 5 and Figure 6 show the average test time for the first test ear, according to age categories and hearing loss categories respectively. From these figures, we can know the older and the worse hearing loss requires the longer test time. Actually, young subjects in the under 50 age group well understood the given task and rapidly adapted to the new hearing test, while a few subjects in the over 50 age group felt hard to perform the hearing test using mobile phone.

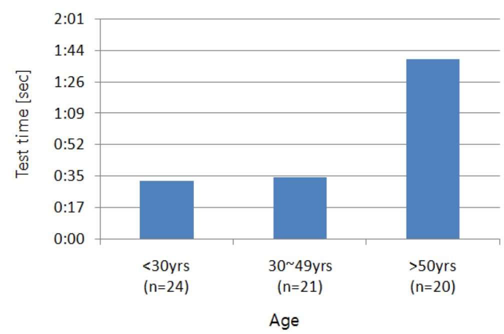

Figure 5. The average test time for the first ear according to age categories

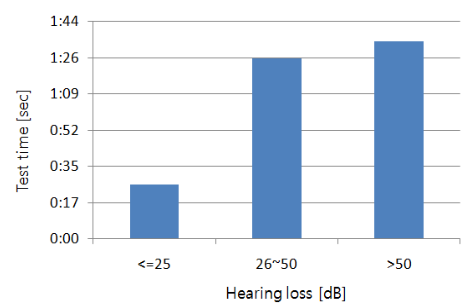

Figure 6. The average test time for the first ear according to hearing loss categories

In the comparison of test time between the first test ear and the second test ear, the average test time were 34 sec and 27 sec, respectively. Figure 7 shows the average test time of each test sound in two tests. In general, the test time for the second ear is shorter than the test time for the first ear because of learning effect. The test time for each phoneme decreased according to sequence of phonemes, because the initial sound level for the first phoneme is set 35dB SPL but the initial sound levels for the other phonemes are set by -5dB SPL than the threshold level of the previous phoneme. For example, if a subject could hear /i/ at 50dB SPL, the initial sound level of /sh/ was set by 45dB SPL to reduce test time.

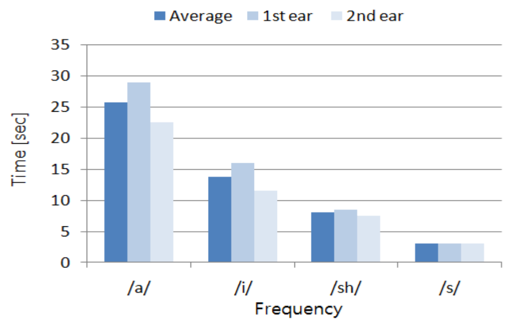

Figure 7. The average test time of each phoneme in the first and second ear test

C. *Compensation of hearing loss using multi-channel wide dynamic range compression*

Based on the measured hearing loss, we played the compensated voices back to four subjects. In order to compensate for hearing loss of each subject, we used multi-channel wide dynamic range compression that was explained at chapter 3. Figure 8 shows hearing losses of the four subjects. One subject(subject 1) has hearing threshold of 25dB in high frequency, another subject(subject 2) has hearing threshold of 45dB at 4000Hz and the other two subjects(subject 3 and subject 4) can be classified normal hearing group.

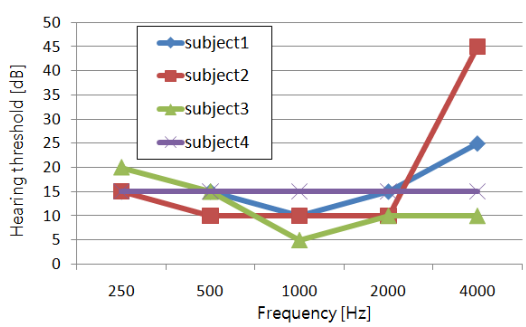

Figure 8. Hearing thresholds of 4 subjects

Figure 9 shows the result of the phoneme-estimated threshold using mobile phones. In this figure, we can know that the phoneme-estimated thresholds resemble the pure tone estimated thresholds shown in Figure 8.

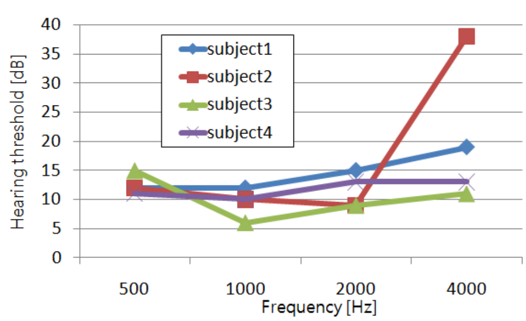

Figure 9. The phoneme-estimated threshold of 4 subjects

Based on the phoneme-estimated thresholds, we set gains of multi-channel wide dynamic range compression for each subject. In order to show the effect of compensation, we used 3 test sounds, /pa/, /ta/, /ka/. In the first test, subjects randomly heard the original version of three test sounds and said what they heard. In the second test, they randomly heard the compensated version of these test sounds and said what they heard. In each test, initial sound level was 60dB SPL and if subjects gave correct answers more than 5 times in 10 attempts, sound level is adjusted by -5dB SPL. Otherwise, the sound level is recorded as hearing threshold. Table 3 shows the effect by multi-channel wide dynamic range compression. In this result, the proposed method can give benefits to user.

TABLE III

THE IMPROVEMENTS OF SPEECH INTELLIGIBILITY

	Subject 1	Subject 2	Subject 3	Subject 4
Original	30dB	25dB	20dB	25dB
Compensated	20dB	20dB	20dB	15dB

V. CONCLUSION

In this study, we proposed new method using speech for hearing test and showed that users can check their hearing loss using mobile phone by themselves. We expect that application of the proposed self assessment of hearing loss and multi-channel dynamic range compensation into mobile phone can bring real benefits such as improvement of speech intelligibility in telephone conversation or a better listen to music or movies for users. Further, hearing loss of users can be managed by mobile phones and test results can be transmitted to remote hospital or hearing center through mobile phones.

REFERENCES

[1] D.D. Dirks, Handbook of clinical audiology: Bone-conduction threshold testing, Lippincott Williams & Wilkins, pp. 132-146(1994).

[2] B.A. Goldstein and C.W. Newman, Handbook of clinical audiology, Clinical masking: a decision–making process, Lippincott Williams & Wilkins, pp. 109-131(1994).

[3] R.J. Roeser, K.A. Buckley and G.S. Stickney GS, Pure tone test, Thieme New York, pp. 227-256(2000).

[4] G.E. Peterson & H.H. Barney, Control methods used in a study of the vowels, Journal of the Acoustic Society of America, Vol. 24, pp. 175-184(1952).

[5] J. Prince and A. Bradley, Analysis/Synthesis filter bank design based on time domain aliasing cancellation, IEEE Transactions, ASSP-34, No. 5, pp. 1153-1161(1986).

[6] J. Princen, A. Johnson and A. Bradley, Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation, Proc. Of the ICASSP, pp. 2161-2164(1987).

[7] T. Sporer, K. Brandenburg and B. Edler, The Use of Multirate Filter Banks for Coding of High Quality Digital Audio, 6th European Signal Processing Conference, Vol. 1, pp. 211-214(1992).