A Hearing Impairment Simulation Method Using Audiogram-based Approximation of Auditory Charatecteristics

Nozomi Jinbo, Shinnosuke Takamichi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura

Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan

{shinnosuke-t, tomoki, neubig, ssakti, s-nakamura}@is.naist.jp

Abstract

Hearing impairment simulation is an effective technique to educate normal-hearing people about auditory perception of the hearing-impaired. Because auditory characteristics of the hearing impaired vary greatly from person-to-person, personalization of the hearing impairment simulation systems is essential to accurately simulate these individual differences. However, measurement of auditory characteristics of individuals is time-consuming work. In this paper, we propose a hearing impairment simulation method that is easily applied to individual hearing-impaired persons. Auditory filter characteristics and gain characteristics are estimated from easily measurable audiograms of each individual. We also implement a method for manually adjusting the hearing impairment level to improve accuracy of the proposed hearing impairment simulation. An experimental evaluation is conducted to compare intelligibility between hearing-impaired and normal-hearing persons with the proposed hearing impairment simulation. The experimental results show that the proposed method effectively makes the word correct rate and phoneme confusion tendency of the normal hearing persons similar to those of the hearing impaired persons.

Index Terms: hearing-impairment simulation, personalization, auditory filter characteristics, gain characteristics, audiogram,

1. Introduction

Hearing impairment, one of the most common diseases causing speech communication disorders, is classified into three types: conductive hearing impairment, sensorineural hearing impairment, and mixed conductive-sensorineural hearing impairment. Conductive hearing impairment is caused by diseases of the external or middle ear and usually causes gain reduction. Sensorineural hearing impairment is caused by diseases of the inner ear. This impairment often causes not only gain reduction but also frequency selectivity reduction. Mixed hearing impairment is a symptom of the development of both conductive and sensorineural hearing impairments. It is well known that auditory characteristics of the hearing impaired vary greatly from person-to-person even if they suffer from the same type of the hearing impairment.

One of the most popular approaches to assist hearing-impaired people in speech communication is hearing aid systems. The gain reduction is alleviated by using digital signal processing. Because auditory characteristics are highly nonlinear to an input sound pressure level, also depending on frequency bands, time-varying filtering process is needed according to the input sound. Moreover, personalization of this process is essential due to large variations in the auditory characteristics among individuals. Although this system is effective for conductive hearing impairment, it is fundamentally difficult to alleviate the frequency selectivity reduction caused by sensorineural hearing impairment and mixed conductive-sensorineural hearing impairment.

Another approach to assist hearing-impaired people in speech communication is to educate normal-hearing people about auditory perception of the hearing-impaired and way of producing speech easily understand by the hearing-impaired people. Hearing impairment simulation is an effective technique for this purpose [1], making it possible for normal-hearing people to experience the auditory perception of the hearing-impaired. HearLoss[1] is one of the hearing impairment simulation systems. A time-invariant gain reduction process and a simple smearing process to reduce the frequency selectivity are implemented in this system, but the accuracy of hearing-impairment simulation is limited to that achievable by these simple processes. To achieve highly accurate hearing-impairment simulation, more sophisticated methods based on auditory filters have been proposed [2, 3, 4, 5]. In these methods, measurement of individual auditory characteristics is essential to develop accurate auditory filters replicating the characteristics of each individual. However, this measurement is very time-consuming work [6]. For the educational purpose mentioned above, it is worthwhile to develop a new hearing impairment simulation system that can be easily personalized to individuals while preserving acceptable simulation accuracy.

In this paper, we propose a new hearing impairment simulation method that is easily personalized to the individual hearing-impaired persons. The proposed method allows normal-hearing people to experience the auditory perception of a target hearing-impaired person under his/her daily conditions in speech communication, *i.e.*, if the target usually wears a hearing-aid system, the proposed method simulates auditory perception of the target wearing it. To approximate auditory characteristics of each individual without the time-consuming measurement of auditory characteristics, we use only easily measurable audiograms, which represent the individual's hearing levels at several frequencies. The gain characteristics and parameters of the auditory filter are roughly predicted from the audiograms. These are used to perform frequency-dependent gain reduction with a time-variant digital filter and smearing based on the auditory filter to simulate frequency selectivity reduction. To improve the simulation accuracy, we further implement manual adjustment of the hearing impairment level for the proposed method. Experimental results show that the proposed method effectively makes the word correct rate and phoneme confusion tendency of normal hearing persons similar to those of the hearing impaired persons.

[1] `http://www.phon.ucl.ac.uk/resource/hearloss/`

2. Basic process of hearing impairment simulation

In this paper, we focus on two basic processes for the hearing impairment simulation: a gain reduction process and a smearing process. The gain reduction process reduces power of the input sound waveform, which is regarded as the inverse process that in a hearing-aid system. It is essential to dynamically control the gain reduction level according to the input sound pressure level due to nonlinearity of the auditory characteristics. Moreover, as the gain reduction characteristics vary from frequency-to-frequency, a frequency-dependent reduction process is necessary. Several gain reduction methods have been proposed based on threshold control of the hearing level [1] and modeling of the nonlinear characteristics using an auditory filterbank [5, 7, 8].

The smearing process reduces the frequency selectivity by smoothing spectral components over the frequency axis. One effective smearing method is the use of the auditory filters [9], such as a roex filter [2], to model the auditory characteristics of the hearing-impaired person and those of the normal-hearing person. The roex filter on the power spectral domain is given by

$$A_c(f) = \begin{cases} \left(1 + p_{u,c}\frac{f-f_c}{f_c}\right)\exp\left(-p_{u,c}\frac{f-f_c}{f_c}\right) & f \geq f_c \\ \left(1 - p_{l,c}\frac{f-f_c}{f_c}\right)\exp\left(p_{l,c}\frac{f-f_c}{f_c}\right) & f < f_c \end{cases} \quad (1)$$

where f is a frequency, f_c is a center frequency of the roex filter, and $p_{u,c}$ and $p_{l,c}$ are parameters to determine the bandwidth in upper and lower frequency sides of the center frequency f_c, respectively. The filter coefficient vector of the center frequency f_c is given by $\boldsymbol{A}(f_c) = [A_c(1), \cdots, A_c(K)]^\top$, where K is the number of points over the frequency axis, *i.e.*, K corresponds to the Nyquist frequency. Let $\boldsymbol{X} = [X(1), \cdots, X(K)]^\top$ and $\boldsymbol{A}_\mathrm{W} = [\boldsymbol{A}_\mathrm{W}(1), \cdots, \boldsymbol{A}_\mathrm{W}(K)]^\top$ be a power spectrum of an input waveform and a matrix consisting of the filter coefficient vectors for a hearing-impaired person, respectively. This input waveform is assumed to be perceived as $\boldsymbol{A}_\mathrm{W}\boldsymbol{X}$ by the hearing-impaired person. In the smearing, the following smeared power spectrum $\boldsymbol{Y} = [Y(1), \cdots, Y(K)]^\top$ is generated:

$$\boldsymbol{Y} = \boldsymbol{A}_\mathrm{N}^{-1}\boldsymbol{A}_\mathrm{W}\boldsymbol{X} \quad (2)$$

where $\boldsymbol{A}_\mathrm{N} = [\boldsymbol{A}_\mathrm{N}(1), \cdots, \boldsymbol{A}_\mathrm{N}(K)]^\top$ is a matrix consisting of the filter coefficient vectors for a normal-hearing person. A smeared waveform is generated from the smeared power spectrum and an original phase spectrum that is extracted from the input waveform assuming that it is perceived as $\boldsymbol{A}_\mathrm{N}\boldsymbol{Y}$ (*i.e.*, $\boldsymbol{A}_\mathrm{W}\boldsymbol{X}$) by the normal-hearing person.

3. Proposed hearing impairment simulation using approximation with audiograms

Although the use of the auditory filters is effective for simulating auditory perception of the hearing-impaired relatively well, it requires time-consuming measurement of auditory characteristics of each individual to determine filter parameters $p_{u,c}$ and $p_{l,c}$ at each center frequency f_c. To address this issue, we propose a hearing-impairment simulation method using only an audiogram of each individual to approximately determine these parameters in addition to the gain reduction characteristics.

Audiograms show the hearing level of a particular person at several frequencies, *e.g.*, 0.25, 0.5, 1, 2, 4, and 8 kHz, as shown in **Figure 1**, where the hearing level corresponds to the threshold at which a sound of that frequency is audible compared with 0 dB of the normal-hearing people. Note that if the hearing-impaired person wears a hearing aid, we use two audiograms: an audiogram measured with a hearing aid is used to determine the gain reduction characteristics and the audiogram measured without a hearing aid is used to determine the auditory filter parameters for the smearing process. The proposed hearing impairment simulation process is shown in **Figure 2**.

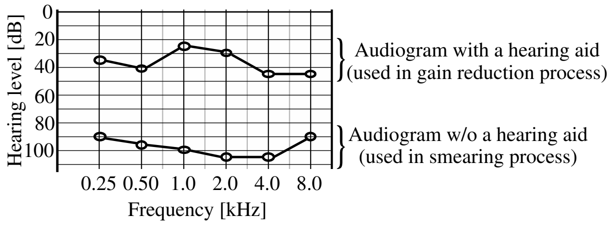

Figure 1: An example of audiograms.

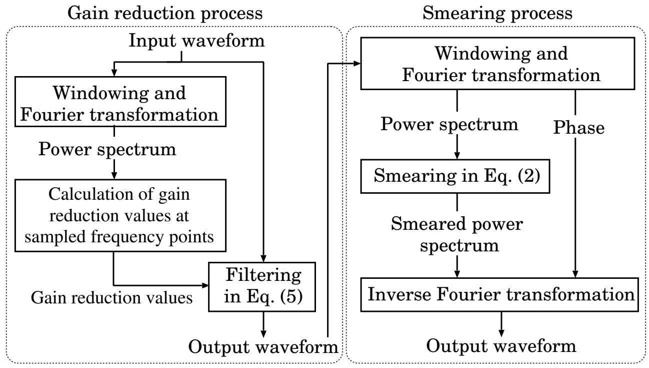

Figure 2: Proposed hearing impairment simulation process.

3.1. Proposed gain reduction process

Let X_k be the power spectrum at frequency k of an input signal. The gain characteristics of the hearing-impaired person are approximated with a piece-wise linear function given by

$$G_k(X_k) = \begin{cases} \dfrac{T_k}{T'_k}X_k & X_k < T'_k \\ \dfrac{120 - T_k}{120 - T'_k}\left(X_k - T'_k\right) + T_k & T'_k \leqq X_k < 120 \\ 120 & 120 \leqq X_k \end{cases} \quad (3)$$

where T_k is the absolute power value corresponding to the hearing level at frequency k of a normal-hearing person and T'_k is that of the hearing-impaired person whose hearing level at frequency k is indicated by o_k, *i.e.*, $T'_k = o_k + T_k$. The uncomfortable hearing level is set to 120 dB. The gain reduction value H_k at frequency k is calculated as

$$H_k = X_k - G_k(X_k). \quad (4)$$

The frequency sampling filter applied to a hearing aid system [10] is employed in the gain reduction process, which is given by

$$H(z) = \frac{1 - r^N z^{-N}}{N} \cdot \left[\sum_{k=1}^{\frac{N}{2}-1} \frac{(-1)^k |H_k| 2\cos\left(\dfrac{\pi k}{N}\right)\left(1 - rz^{-1}\right)}{1 - 2rz^{-1}\cos\left(\dfrac{2\pi k}{N}\right) + r^2 z^{-2}} + \frac{H_0}{1 - rz^{-1}}\right] \quad (5)$$

where H_k is a filter parameter showing frequency response at a sampled frequency point k, which is determined with the gain reduction value. The parameter r is used to make the filter

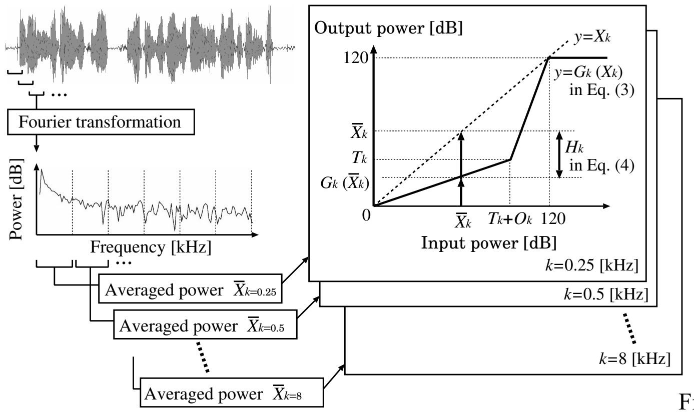

Figure 3: Gain reduction process in the proposed method.

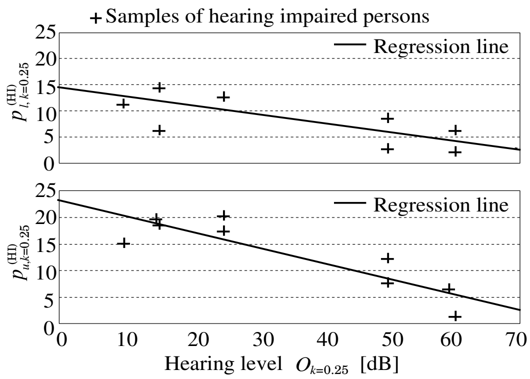

Figure 4: An example of hearing levels and corresponding filter parameters at the center frequency $f_c = 0.25$ kHz among several hearing-impaired persons shown in [12].

stable, and N is the tap length. This process allows us the frequency-specific gain reduction.

Several processes are implemented to alleviate discontinuities caused by rapidly changing the filter parameters. At each analysis frame, the power spectrum averaged in several frequency bands is calculated as X_k in the gain reduction value estimation as shown in **Figure 3**, and then spline interpolation is performed to determine the filter parameter H_k at each sampled frequency point. These parameters are further smoothed between neighboring analysis frames and interpolated sample by sample during filtering.

3.2. Proposed smearing process

It has been reported in [11] that there is a correlation between the hearing level and the roex filter parameters $p_{u,c}$ and $p_{l,c}$. Therefore, we predict these parameters from the hearing level shown in the audiogram. In this paper, linear regression is performed to predict them as follows:

$$p_{u,k}^{(\mathrm{HI})} = a_{u,k} o_k + b_{u,k}, \qquad p_{l,k}^{(\mathrm{HI})} = a_{l,k} o_k + b_{l,k}, \tag{6}$$

where $a_{u,k}$, $b_{u,k}$, $a_{l,k}$, and $b_{l,k}$ are regression parameters and $p_{u,k}^{(\mathrm{HI})}$ and $p_{l,k}^{(\mathrm{HI})}$ are the roex filter parameters at the center frequency k for a hearing-impaired person whose hearing level at frequency k is shown as o_k. The regression parameters are estimated using sample pairs of the hearing level and roex filter parameters among several hearing-impaired persons shown in [12]. **Figure 4** shows an example of the data samples and the estimated linear regression line. The prediction accuracy with the linear regression is shown in **Table 1**. The parameters at the other center frequencies are determined with linear interpolation between the predicted parameters.

On the other hand, the roex filter parameters for the normal-hearing person $p_{u,k}^{(\mathrm{NH})}$ and $p_{l,k}^{(\mathrm{NH})}$ are relatively stable. Therefore, they are set to constant values so that the filter bandwidth is equal to equivalent rectangular bandwidth (ERB) assuming that $p_{u,k}^{(\mathrm{NH})}$ and $p_{l,k}^{(\mathrm{NH})}$ are equal to each other. In such a case, the following equations hold:

$$e = 2\frac{f_c}{p_{u,k}^{(\mathrm{NH})}} + 2\frac{f_c}{p_{l,k}^{(\mathrm{NH})}} \tag{7}$$

$$e = 24.7\,(0.00437 f_c + 1)\,. \tag{8}$$

3.3. Manual adjustment of hearing level

The proposed hearing impairment simulation method is easily personalized to individual hearing-impaired persons but simulation accuracy tends to be degraded by several approximations particularly for the roex filter parameter prediction. To address this issue, we further propose a manually control method to the hearing level. The roex filter parameters in the smearing process are interpolated between those for the hearing-impaired person and those for the normal-hearing person as follows:

$$p_{\mathrm{u},k}^{(\mathrm{HI})\prime} = \alpha p_{\mathrm{u},k}^{(\mathrm{NH})} + (1-\alpha)\, p_{\mathrm{u},k}^{(\mathrm{HI})} \tag{9}$$

$$p_{\mathrm{l},k}^{(\mathrm{HI})\prime} = \alpha p_{\mathrm{l},k}^{(\mathrm{NH})} + (1-\alpha)\, p_{\mathrm{l},k}^{(\mathrm{HI})} \tag{10}$$

where α is an interpolation weight to be controlled manually.

4. Experimental evaluation

4.1. Experimental conditions

We estimated the parameters of the proposed hearing impairment system for a severe hearing-impaired (severe HI) person, who is in her twenties, wearing a hearing aid system in a daily life and a moderate hearing-impaired (moderate HI) person, who is in his fifties, not wearing a hearing aid system. Both persons suffered from the sensorineural hearing impairment. Their hearing levels were measured at 0.25, 0.5, 1, 2, 4, 8 kHz. Note that the hearing level of the severe HI person was measured with/without the hearing aid system.

In frame analysis, we used a Hanning window was used with length of 8 ms and a frame shift of 4 ms. The FFT size was set to 128 in the gain reduction process and 256 in the smearing process. For the frequency sampling filter in Eq. (5), the tap length was set to 32 and the parameter r was set to 0.95. The system was developed with C language.

To evaluate the simulation accuracy of the proposed systems, we conducted word intelligibility tests by manual dictation. We used a speech database of familiarity-controlled word-lists FW07 [13] including 1600 Japanese word utterances classified to four categories in terms of word familiarity. Sampling frequency was set to 16 kHz. Using both 20 familiar words from the familiarity 4 category and 20 unfamiliar words from the familiarity 2 category, the following four types of speech samples were evaluated by listeners: normal speech (N), simulated speech with gain reduction process (G), simulated speech with smearing process (S), and simulated speech with gain reduction and smearing processes (GS). The interpolation weight α was set to 0. The listeners were asked to write down a word by listening to each speech sample. We also conducted another intelligibility test using the simulated speech samples with "GS" by setting α to 0.2, 0.4, 0.6, and 0.8. In each setting, 20 famil-

Table 1: Prediction accuracy of the roex filter parameters from the hearing level at each frequency.

Frequency [kHz]	0.25		0.5		1.0		2.0		4.0	
Filter parameters	p_l	p_u	p_l	p_u	p_l	p_u	p_l	p_u	p_l	p_u
Correlation coefficient	-0.75	-0.88	-0.30	0.08	-0.74	-0.26	0.41	0.74	-0.74	-0.10
Root mean square error	3.5	4.4	8.5	9.4	7.0	12.9	11.5	4.1	2.6	14.2
Number of hearing impaired persons	9		10		10		6		8	

iar words (familiarity 4) and 20 unfamiliar words (familiarity 2) were evaluated. In each intelligibility test, the number of listeners was 8 and each listener evaluated 160 different words in total. All listeners are male and female graduate students in their twenties. Each word was presented to each listener only once. As a target reference, the severe and moderate HI persons evaluated 60 familiar words (familiarity 4) and 60 unfamiliar words (familiarity 2) of normal speech.

4.2. Experimental results

Figure 5 shows the results of the hearing impairment simulation for the severe HI person and the moderate HI person. The smearing process "S" causes larger degradation in the word correct rate than the gain reduction process "G." Moreover, the combination of both processes "GS" causes further degradation. We can see a tendency that the word correct rate of the simulated speech "GS" is significantly lower than the target "HI." This difference is caused by the approximated processes in the proposed method.

Figure 6 shows the result when varying the interpolation weight α. The word correct rate gradually recovers as the interpolation weight increases. We can see that the word correct rate of the simulated speech is close to that of the target "HI" in both the familiarity 4 words and the familiarity 2 words by setting α to 0.4 or 0.6 for the severe HI person and to 0.6 or 0.8 for the moderate HI person. This result suggests that the simulation accuracy is effectively improved by adjusting the interpolation weight in the proposed method according to individual hearing-impaired persons.

We also investigated phoneme confusion tendency of the target HI persons and the normal-hearing persons using the proposed system. As a result, mora confusion rates for the severe HI person are shown in **Table 2**, where a mora is one linguistic unit in Japanese. We can see that the confusion rates of the simulated speech "GS" ($\alpha = 0.6$) are somewhat similar to those of the target HI. We also found a similar tendency for the moderate HI person.

5. Conclusions

In this paper, we have proposed a hearing impairment simulation method based on a gain reduction process and a smearing process using audiogram-based approximation of auditory characteristics, enabling easily application to individual hearing-impaired persons. We have used the proposed system to simulate a severe hearing-impaired person and a moderate hearing-impaired person and evaluated their performance. The experimental results have demonstrated that intelligibility of the simulated speech tends to be lower than that of the target hearing-impaired persons but this difference can be effectively minimized by manually adjusting a single parameter to control the hearing level.

Acknowledgements: Part of this work was supported by JSPS KAKENHI Grant Number 23240023.

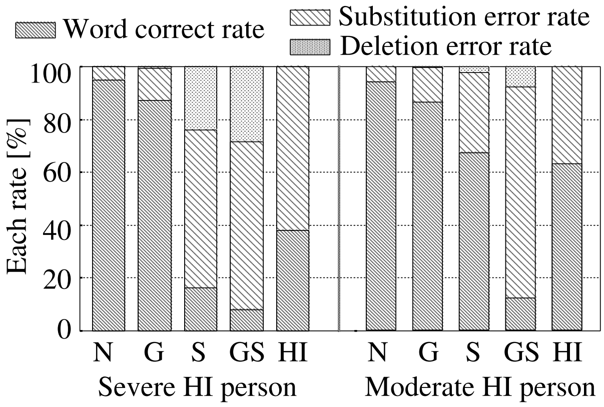

Figure 5: Results of intelligibility tests.

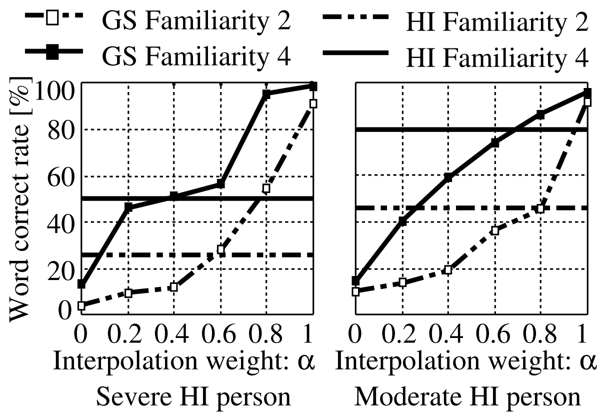

Figure 6: Word correct rate as a function of the interpolation weight α.

Table 2: Mora confusion rates averaged over morae consisting of the same vowel for the severe HI person (upper table) and normal-hearing persons with the proposed system (lower table). Each row shows presented mora and each column shows perceived mora.

Severe HI	/a/	/i/	/u/	/e/	/o/	/N/
/a/	**96.3**	0.0	3.7	0.0	0.0	0.0
/i/	3.8	**57.7**	34.6	0.0	0.0	3.8
/u/	3.1	**68.8**	15.6	9.4	0.0	3.1
/e/	0.0	4.8	4.8	**47.6**	42.9	0.0
/o/	16.7	0.0	0.0	8.3	**50.0**	25.0
/N/	0.0	**87.5**	12.5	0.0	0.0	0.0

GS, $\alpha = 0.6$	/a/	/i/	/u/	/e/	/o/	/N/
/a/	**67.4**	0.0	1.1	2.1	29.3	0.0
/i/	0.0	**71.1**	23.6	0.0	1.1	4.2
/u/	0.0	**74.0**	14.7	6.7	0.9	3.7
/e/	13.6	2.8	10.2	30.6	**42.8**	0.0
/o/	20.5	5.0	5.0	13.6	**55.9**	0.0
/N/	0.0	**87.0**	13.0	0.0	0.0	0.0

6. References

[1] N. P. Erber, "Applications of hearing-loss simulation in education of student clinicians," *Journal of Academy of Rehabilitative Audiology*, Vol. 28, pp. 37–50, 1995.

[2] R. D. Patterson, I. N. Smith, D. L. Weber, and R. Milroy, "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," *The Journal of the Acoustical Society of America*, Vol. 72, No. 6, pp. 1788–1803, Dec. 1982.

[3] B. C.J. Moore, B. R. Glasberg, and A. Simpson, "Evaluation of a method of simulating reduced frequency selectivity," *The Journal of the Acoustical Society of America*, Vol. 91, No. 6, pp. 3402–3423, June 1992.

[4] R. S. Tyler, J. W. Hall, B. R. Glasberg, B. C. J. Moore, and R. D. Patterson, "Auditory filter asymmetry in the hearing impaired," *The Journal of the Acoustical Society of America*, Vol. 76, No. 5, pp. 1363–1368, Nov. 1984.

[5] H. Hu, J. Sang, M. E. Lutman, and S. Bleech, "Simulation of hearing loss using compressive gammachirp auditory filters," *Proc. ICASSP*, pp. 5428–5431, Prague, Czech Republic, May 2011.

[6] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing research*, Vol. 47, No. 1, pp. 103–138, Feb. 1990.

[7] T. Irino and R. D. Patterson. "A time-domain, level-dependent auditory filter: the gammachirp," *The Journal of the Acoustical Society of America*, Vol. 101, No. 1, pp. 412–419, Jan. 1997.

[8] T. Irino and R. D. Patterson. "A compressive gammachirp auditory filter for both physiological and psychophysical data," *The Journal of the Acoustical Society of America*, Vol. 109, No. 5, pp. 2008–2022, May 2001.

[9] T. Baer and B. C. J. Moore, "Effects of spectral smearing on the intelligibility of sentences in noise," *The Journal of the Acoustical Society of America*, Vol. 94, No. 3, pp. 1229–1241, Sep. 1993.

[10] F. Asano, Y. Suzuki, T. Sone, T. Hayashi, M. Satake, K. Ohyama, T. Kobayashi, and T. Takasaka, "A method of designing and realizing a digital hearing aid that compensates loudness for impaired listeners," *The Journal of the Acoustical Society of Japan (in Japanese)*, Vol. 47, No. 6, pp. 373–379, June, 1991.

[11] B. R. Glasberg and B. C. J. Moore, "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *The Journal of the Acoustical Society of America*, Vol. 79, No. 4, pp. 1020–1033, Apr. 1986.

[12] J. G. Desloge, C. M. Read, L. D. Braida, Z. D. Pere, and L. A. Delhorne, "Auditory-filter characteristics for listeners with real and simulated hearing impairment," *Trends in Amplification*, Vol. 16, No. 1, pp. 19–39, Mar. 2012.

[13] T. Kondo, S. Amano, S. Sakamoto, Y. Suzuki, "Development of familiarity-controlled word-lists (FW07)," *IEICE Technical Report (in Japanese)*, Vol. 107, No. 432, pp. 43–48, Jan., 2008.