

Hearing Aid Method by Equalizing Frequency Response of Phoneme Extracted from Human Voice

Kohei Arai¹

1 Graduate School of Science and Engineering
Saga University
Saga City, Japan

Takuto Konishi¹

1 Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Hearing aid method by equalizing frequency response of phoneme which is extracted from human voice is proposed. One of the problems of the existing hearing aid is poor customization of the frequency response compensation. Frequency response characteristics are different by the person who need hearing aid. The proposed hearing aid is based on frequency response equalization by phoneme by phoneme. Frequency characteristics of phoneme are to be equalized. This is the specific feature of the proposed hearing aid method. Through experiments, it is found that the proposed hearing aid by phoneme is superior to the conventional hearing aid.

Keywords—Hearing aid; phoneme; frequency response; equalization filter; hidden markov model (HMM)

I. INTRODUCTION

In general, hearing capability of human voices is getting bad for elderly persons since a high-frequency response of elderly persons' ears is getting poor. Hearing capability is defined with the well-known averaged hearing capability level that is defined as Averaged value of hearing capability for human voices regarding frequency components ranged from 500 Hz to 4000 Hz. In accordance with the definition, 25-40 dB of loudness of human voices are difficult to hear slightly when human voice is not loud while 40-70 dB of loudness of human voices are difficult to hear when human voice is normal level.

Earlier devices, known as ear trumpets or ear horns [1], [2], were passive funnel-like amplification cones designed to gather sound energy and direct it into the ear canal. After that not so small number of methods have been proposed so far [3]-[10].

Mobile device based personalized equalizer for improving the hearing capability of human voices for elderly persons are proposed. Through experiments, it is found that the proposed equalizer does work well for improving hearing capability by 2 to 55% of voice Recognition success ratio. According to the investigation of the frequency component analysis and formant detections, most of the voice sounds have the formant frequencies for the first to third frequencies within the range of 3445 Hz. Therefore, a nonlinear equalizing multiplier is better to enhance the frequency components for the first to third formants. The experimental results with the voice above input experiments show that a good Percent Correct Recognition: PCR is required for 0 to more than 8000 Hz of frequency components. Also, 8162 Hz cut off frequency

would be better for both noise suppressions and keeping a good PCR [11].

As I described above, hearing capability is getting deteriorated for aged persons. It is called "Senile deafness". In Japan, around 18% of peoples whose age ranged from 65 to 74 have a trouble on hearing capability while 40% of peoples whose age is more than 74 have a trouble on hearing capability. There are some young peoples who have a trouble on hearing capability for some specific frequency component. Although they need a hearing aid, most of they do not like to have such conventional hearing aid due to some reasons. It does not look good. Hearing capability, frequency response varied for time being. Hearing capability is different by person. There are some other reasons.

Because of these reasons, a customization of hearing aid is required. Also, equalization of specific spectrum components is required. Furthermore, it would be better to equalize specific frequency component by phoneme by phoneme if they would like to hear human voices. Therefore, human voice hearing capability improvement method by equalizing frequency response equalization by phoneme by phoneme is proposed. This is the specific feature of the proposed hearing aid method.

The following section describes the proposed method for equalization followed by some experiments. Then conclusions are described together with some discussions and future research works.

II. PROPOSED METHOD

A. Frequency Response Model

Fig. 1 shows the cochlea of human ear model.

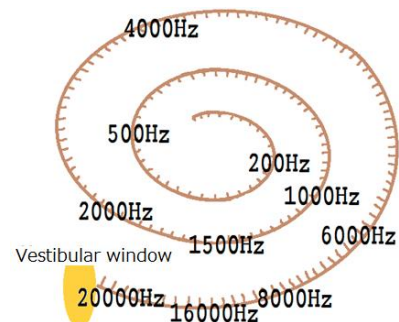


Fig. 1. Cochlea of human ear model.

From the vestibular window to the end of the cochlea, frequency response is varied from high to low frequency components ranged from zero to 20KHz around. Usually, high frequency response is getting deteriorated by age. On the other hand, some frequency response is degraded for the young generation's deafness.

B. Japanese Phoneme Model

Frequency range of each Japanese phoneme is shown in Table 1. In Japanese, there are just 23 of phoneme. The number of phoneme is different by Language. The number of phoneme of Japanese is smallest followed by Germany (the number of phoneme is 25).

It is considerably certain that it would better to equalize by phoneme by phoneme because frequency component of each phoneme is different each other. This is the fundamental idea of the proposed hearing aid. Also, it is realized by using smartphone or i-phone as an application software installed on the mobile devices. Therefore, it can be customized by human and may be changed the equalization characteristics even if their frequency response is changed for time being. Also, it can be worked in a real-time basis because the equalization filter can be created in prior to use.

Auditory Steady-State Response: ASSR [12] allows to measure frequency response of human ear objectively (Galambos et al. (1981) [13], Rickards et al. (1994) [14], Kuwada et al. (1986) [15]). During sleep, frequency response can be measured using ASSR.

TABLE I. FREQUENCY RANGE OF EACH JAPANESE PHONEME¹

	Frequency Range(Hz)
Vowel	a 0~1500
	i 0~1000,4000~5000
	u 0~700
	e 0~1000,2000~3000
	o 0~1000
Consonant	k 900~3500
	s 4000~5100
	t 4000~5100
	c ~
	n 3900~4900
	m 0~400
	r 0~1000,4000~4500
	g 0~500,2000~5000
	z 0~500,4000~5000
	d 0~500,3900~5000
	b 0~1000
	p 200~700
	j 0~1000,4000~5000
	w ~
Special Mora	n 0~400
	q ~
	H ~

¹ <http://www.geocities.jp/myonseil/>

It is proposed to measure responses by input 23 of different phoneme to human ear using ASSR. Then appropriate equalizer for each phoneme is designed and installed it to smartphone or i-phone in prior to use.

C. Procedure of the Proposed Design of Equalization

Before using the proposed equalizer, customization of the equalization is required. The most appropriate equalization filter response is designed as follows:

- 1) Frequency response characteristic of each phoneme is measured with ASSR.
- 2) Equalization filter is designed by each phoneme.

Phoneme is extracted from the acquired voice signals based on Hidden Markov Model: HMM² which is shown in Fig. 2. "Julius" software which is developed by Julius development team composed with Kyoto University, Nagoya Institute of Technology, etc. which allows speech recognition.³

First, input voice signals are divided into the frames (25 ms in this case) with the pre-assigned short term shift of the signals (10 ms in this case) as shown in Fig. 3.

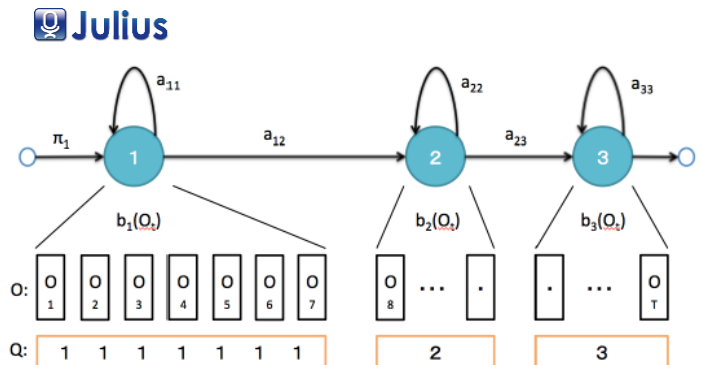


Fig. 2. Example of the well known Hidden Markov Model: HMM.

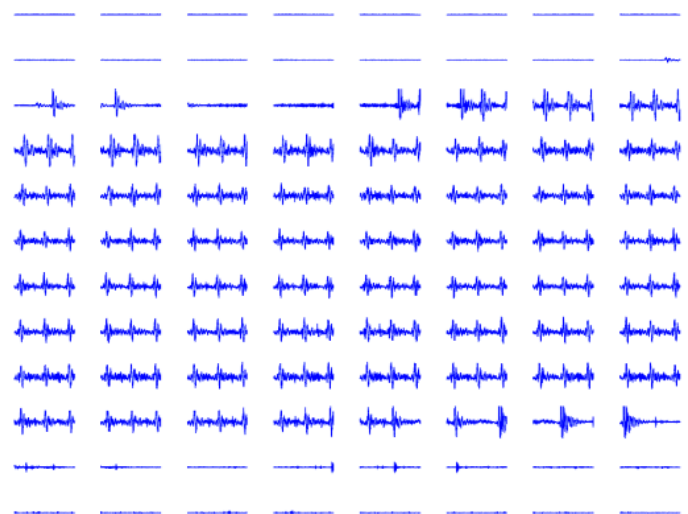


Fig. 3. Divided input voice signal into frames with 25ms of width.

² https://en.wikipedia.org/wiki/Hidden_Markov_model

³ <https://julius.osdn.jp/juliusbook/ja/julius.html>

After that, phoneme is extracted from the frame signal with the quality assessed results “n_score” as shown in Fig. 4. The frames are attached frame ID and assessed frames are attached “unit number”. These are candidates of the phoneme. The most reliable phoneme is selected from the candidates. In the case of Fig. 4, #2 of units are selected depending on the assessed “n_score”.

Input voice signals are equalized using previously designed equalizing filters by each phoneme. The equalizing filter is designed as a bandpass filter as shown in Fig. 5. Such bandpass filter can be synthesized by composing low-pass, bandpass and high-pass filters. The low-pass filter suppresses the existing noises while bandpass filter enhances the required frequency response.

The high-pass filter suppresses a low frequency noise. Another method for creating equalizing filter is a composition of low-pass and high-pass filter which are shown in Fig. 6. By combine the two low-pass and high-pass filters, an arbitrary frequency response of equalizing filter can be designed.

The filter responses are candidates of the low-pass filters (see Fig. 7). From these candidates, calm frequency response of filter is selected.

id:	from	to	n_score	unit
[0	13]	-20.464859	silB #1	
[14	14]	-25.480347	silB #2	
[15	19]	-28.562841	silB #3	
[20	20]	-38.377106	ky #1	
[21	31]	-36.201962	ky #2	
[32	35]	-37.868408	ky #3	
[36	40]	-34.720387	o: #1	
[41	55]	-25.332415	o: #2	
[56	56]	-31.801636	o: #3	

Fig. 4. Phoneme is extracted from the frame signal with the quality assessed results “n_score”.

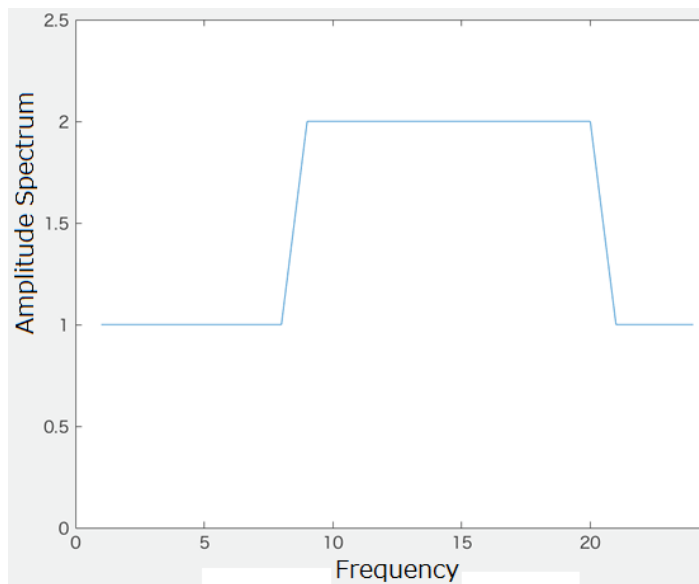
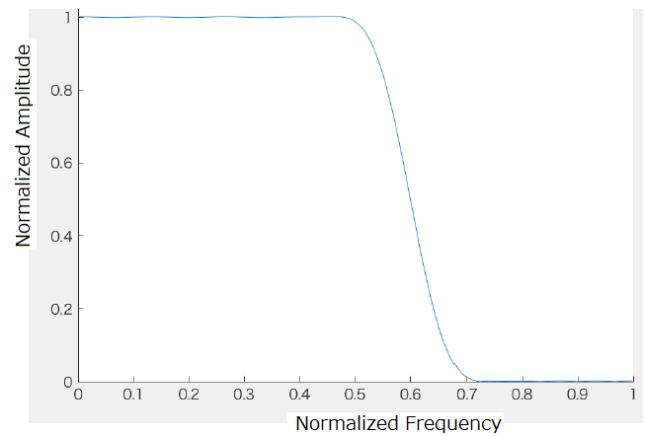
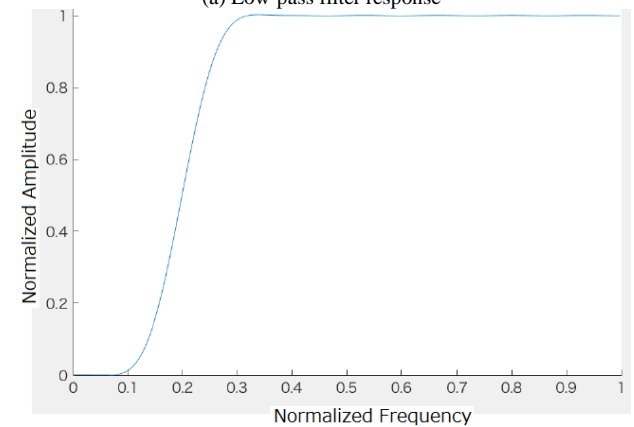


Fig. 5. Frequency response of the required equalizing filter.

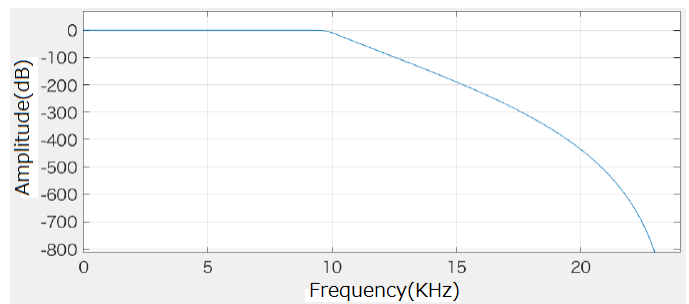


(a) Low-pass filter response

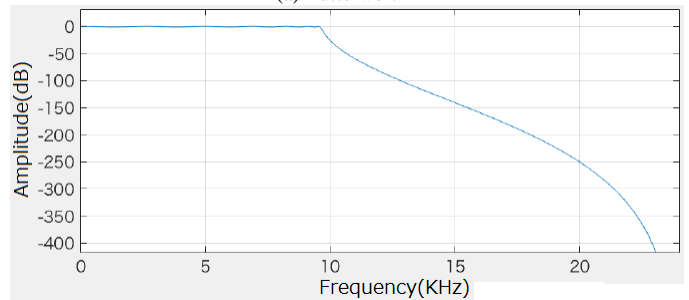


(b) High-pass filter response

Fig. 6. Arbitrary bandpass filter frequency response creation with low-pass and high-pass filters.



(a) Butterworth⁴



(b) Cevichef I⁵

⁴ https://en.wikipedia.org/wiki/Butterworth_filter.

⁵ <https://ja.wikipedia.org/wiki/%E3%83%81%E3%82%A7%E3%83%93%E3%82%B7%E3%82%A7%E3%83%95%E3%83%95%E3%82%A3%E3%83%A B%E3%82%BF>.

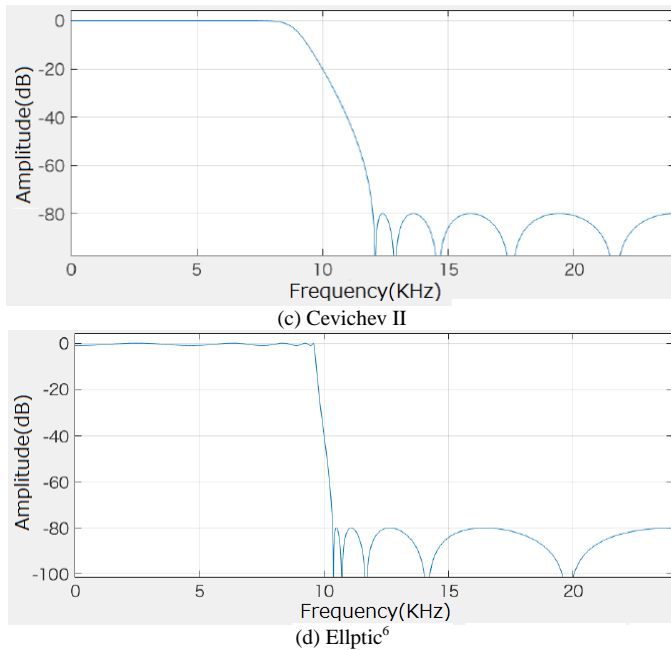


Fig. 7. Candidates of the low-pass filter of frequency responses for equalizing filter.

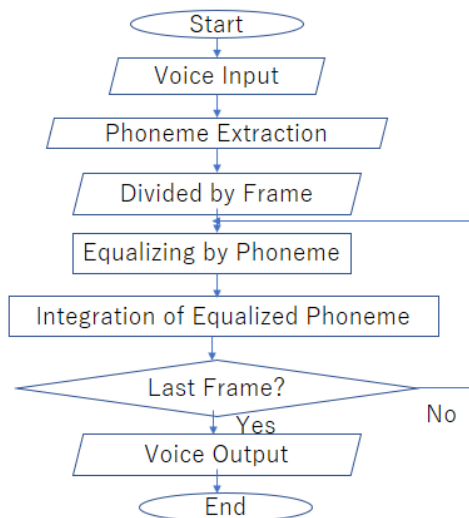


Fig. 8. Detailed flow chart of the proposed procedure.

The detailed flow chart of the proposed procedure is shown in Fig. 8.

After the voice is input in the PC with microphone, phoneme is extracted from the input voice signal followed by division of phoneme by 25ms of frame. Then equalization filter is retrieved by phoneme database followed by integration of the equalized phoneme until the end of the divided frames. After that, the equalized voice signal is output from the PC with speaker.

III. EXPERIMENTS

A. Experimental Environment

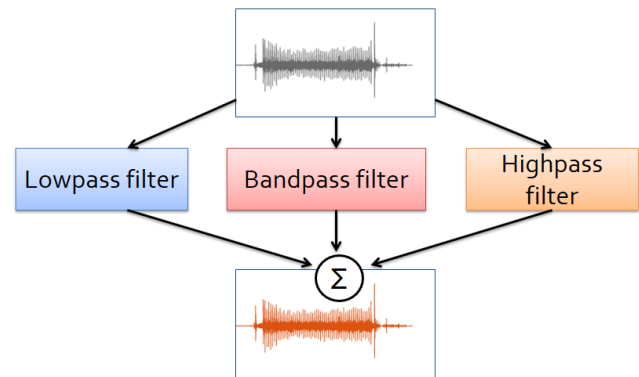
Experimental environment is shown in Table 2. The entire program used for the experiment is based on Matlab.

TABLE II. EXPERIMENTAL ENVIRONMENT

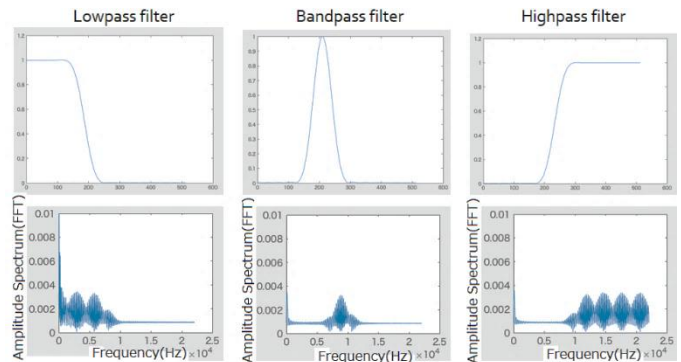
PC	MacBook_Pro_Retina_Mid_2014
OS	MacOS_X_10.9.4
CPU	2.6GHz_Intel_Core_i5
Main_Memory	8GB_1600MHz_DDR3
Programing_Lanbguage	Matlab
Software	MATLAB_R2015b

B. Preliminary Experiment

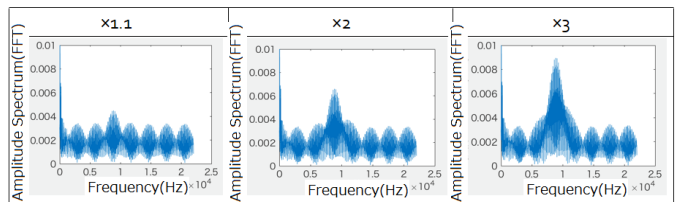
The basic idea behind the proposed equalizing filter is illustrated in the Fig. 9(a). Example of the designed low-pass, high-pass and bandpass filters are shown in Fig. 9(b). Meanwhile, specific frequency ranges can be enhanced as shown in Fig. 9(c).



(a) Illustrative view of the basic idea of the proposed equalizing filter



(b) Three filters



(c) Arbitrary frequency ranges are enhanced

Fig. 9. Basic idea of the proposed equalizing filter and example of the frequency responses of the designed low-pass, high-pass and bandpass filters.

⁶ https://en.wikipedia.org/wiki/Elliptic_filter

C. Experimental Results

One of the examples of actual spectrum of phoneme is shown in Fig. 10. This is an example of “a”. There are peaks which are named as Formants (from the first to n-th formants) which represent features of the input voices.

Appropriate frequency ranges which must be enhanced are determined with the formants. These formants are estimated with envelopes of frequency spectrum of each phoneme. Then appropriate filter response can be designed by the method.

Fig. 11(a) shows the frequency responses with frequency enhancement while Fig. 11(b) shows the frequency responses without enhancement. #2 in Fig. 4 must be enhanced while #1 and #3 has not to be enhanced. The left image shows #1 and #3 of frequency response while the right image shows #2 of frequency response which must be equalized.

The processed voice signals by the proposed frequency response equalization are shown in Fig. 12. The left image is the original voice input signal while the right image shows the reconstructed output voice signal after the frequency equalization. These are corresponding to the voice signals which are shown in Fig. 11(a) of the left and the right images, respectively.

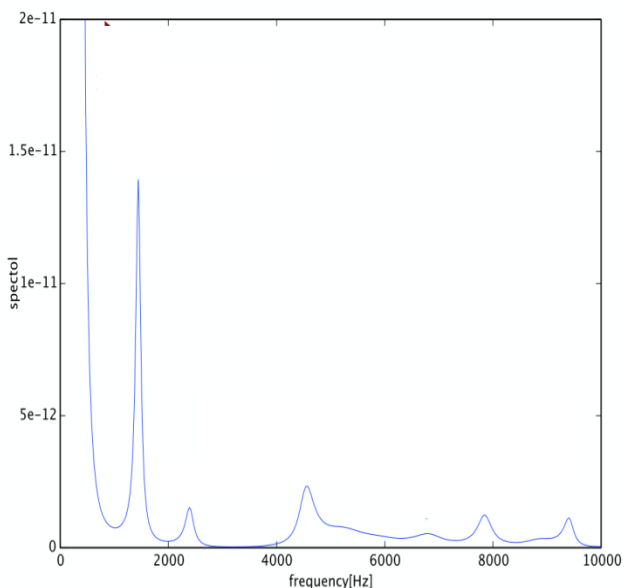
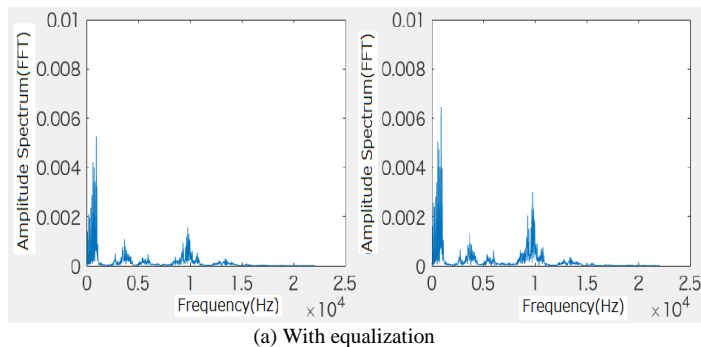
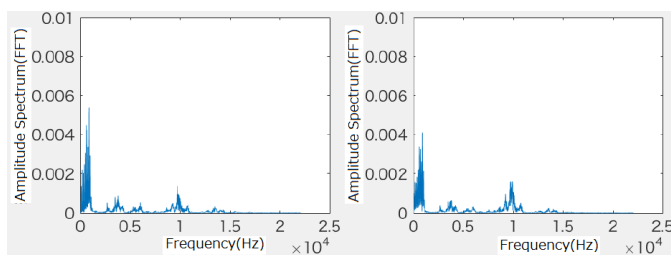


Fig. 10. Example of formant of “a”.



(a) With equalization



(b) Without equalization

Fig. 11. Frequency responses with and without frequency enhancement.

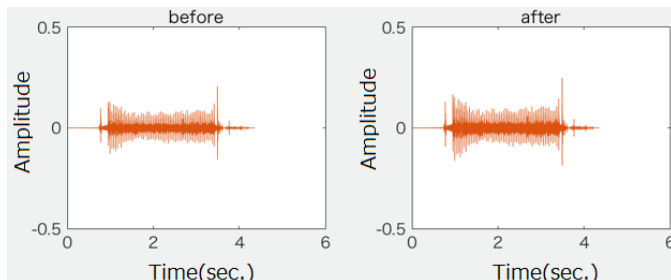


Fig. 12. Example of the processed voice signals by the proposed frequency response equalization.

Four of patients participate a validation test for the proposed system. “Kyo-wa Ii-Tenki-da” in Japanese (“It is fine day” in English) is pronounced by the user. 44,100Hz (Sampling frequency) / 16bit (Quantization bit) / monaural voice signal is created. Also, a degraded input voice signal is created by using low-pass filter with “butter-worse” filter with the cut-off frequency at the 5 KHz. This is called as #1 input voice signal hereafter. The #2 input voice signal is also created with conventional frequency equalization with high-pass filter (the cut-off frequency is at 5 KHz). Another #3 input voice signal is created with the proposed method of frequency equalizer. The four patients hear these three input voice signals and then evaluate the quality of voice with 5 grades. Table 3 shows the evaluation results.

As the results from the evaluation experiments for three input voice signals, it is found that the proposed method shows superior performance to the other two degraded voice signals and the restored voice signal with conventional high-pass filter about 10 points. It is noticed that some of the consonances are not clear enough though. Also, it is noticed that #3 input voice is not so natural since reconstruction is made some sound defects caused by the combining the different frame signal peace of phoneme for the proposed frequency equalization method. In comparison to the conventional method, the reconstructed voice signal by the proposed method is not so noisy. This is one of the features of the proposed method.

TABLE III. EVALUATED RESULTS FOR THREE INPUT VOICE SIGNALS

Input_Voice_Signal	Score	Comments
#1 Voice_Signal	2.67	Relatively_unclear
#2 Conventional	3.5	Consonance_is_not_clear_enough, Noisy
#3 Proposed	3.83	Comparatively_consonance_is_clear, Not_so_noisy

IV. CONCLUSION

Human voice hearing capability is improved by equalizing frequency response equalization by phoneme by phoneme. One of the problems of the existing hearing aid is poor customization of the frequency response compensation. Frequency response characteristics are different by the person who need hearing aid. The proposed hearing aid is based on frequency response equalization by phoneme by phoneme. Through experiments, it is found that the proposed hearing aid by phoneme is superior to the conventional hearing aid by the factor of 9.4 %.

It is found that the proposed method shows superior performance to the other two degraded voice signals and the restored voice signal with conventional high-pass filter about 10 points. It is noticed that some of the consonances are not clear enough though. Also, it is noticed that #3 input voice is not so natural due to the fact that reconstruction is made some sound defects caused by the combining the different frame signal piece of phoneme for the proposed frequency equalization method. In comparison to the conventional method, the reconstructed voice signal by the proposed method is not so noisy. This is one of the features of the proposed method.

Further investigations are required for simultaneous estimation of cornea curvature center and cornea radius, noise removal of the depth image.

ACKNOWLEDGMENT

The authors would like to thank the fourth group members of the Department of Information Science, Saga University for their contribution to the experiments.

REFERENCES

- [1] Bentler Ruth A., Duve , Monica R. (2000). "Comparison of Hearing Aids Over the 20th Century". *Ear & Hearing* 21 (6): 625–639. doi:10.1097/00003446-200012000-00009.
- [2] Bentler, R. A.; Kramer, S. E. (2000). "Guidelines for choosing a self-report outcome measure". *Ear and hearing* 21 (4 Suppl): 37S–49S. doi:10.1097/00003446-200008001-00006. PMID 10981593. edit
- [3] Jack Katz; Larry Medwetsky; Robert Burkard; Linda Hood (2009). "Chapter 38, Hearing Aid Fitting for Adults: Selection, Fitting, Verification, and Validation". *Handbook of Clinical Audiology* (6th ed.). Baltimore MD: Lippincott Williams & Wilkins. p. 858. ISBN 978-0-7817-8106-0.
- [4] K. Sickel, Shortest Path Search with Constraints on Surface Models of In-ear Hearing Aids 52. IWK, Internationales Wissenschaftliches Kolloquium (*Computer science meets automation Ilmenau 10. - 13.09.2007*) Vol. 2 Ilmenau : TU Ilmenau Universitätsbibliothek 2007, pp. 221-226.
- [5] K. Sickel et al., Semi-Automatic Manufacturing of Customized Hearing Aids Using a Feature Driven Rule-based Framework *Proceedings of the*

Vision, Modeling, and Visualization Workshop 2009 (Braunschweig, Germany November 16–18, 2009), pp. 305-312.

- [6] Dave Fabry, Hans Mülder, Evert Dijkstra (November 2007). "Acceptance of the wireless microphone as a hearing aid accessory for adults". *The Hearing Journal* 60 (11): 32–36. doi:10.1097/01.hj.0000299170.11367.24.
- [7] Hawkins D (1984). "Comparisons of speech recognition in noise by mildly-to-moderately hearing-impaired children using hearing aids and FM systems". *Journal of Speech and Hearing Disorders* 49 (4): 409. doi:10.1044/jshd.4904.409.
- [8] Ricketts T., Henry P. (2002). "Evaluation of an adaptive, directional-microphone hearing aid". *International Journal of Audiology* 41 (2): 100–112. doi:10.3109/14992020209090400.
- [9] Lewis M Samantha, Crandell Carl C, Valente Michael, Horn Jane Enrietto (2004). "Speech perception in noise: directional microphones versus frequency modulation (FM) systems". *Journal of the American Academy of Audiology* 15 (6): 426–439. doi:10.3766/jaaa.15.6.4.
- [10] Exemption from Preemption of State and Local Hearing Aid Requirements; Applications for Exemption, Docket No. 77N-0333, 45 Fed. Reg. 67326; Medical Devices: Applications for Exemption from Federal Preemption of State and Local hearing Aid Requirements, Docket No. 78P-0222, 45 Fed. 67325 (Oct. 10, 1980).
- [11] Kohei Arai, Takuto Konishi, Mobile device based personalized equalizer for improving hearing capability of human voices in particular for elderly persons, *International Journal of Advanced Research on Artificial Intelligence*, 4, 6, 23-27, 2015.
- [12] American Academy of Pediatrics. Task Force on Newborn Hearing Screening. Newborn and infant hearing loss: detection and intervention. *Pediatrics* 103:53-68., 1999.
- [13] Galambos R, Makeig S, Talmachoff PJ. A 40 Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A*, 1981.
- [14] Rickards FW, Tan LE, Cohen LT, et al. Auditory steady-state evoked potentials in newborns. *Br J Audiol* 28:327-337., 1994.
- [15] Kuwada S, Batra R, Maher VL. Scalp potentials of normal and hearing-impaired subjects in response to sinusoidal, 1986.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 37 books and published 570 journal papers. He received 30 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. <http://teagis.ip.is.saga-u.ac.jp/index.html> .