Acoust. Sci. & Tech. **25**, 1 (2004)

ACOUSTICAL LETTER

Critical-band based frequency compression for digital hearing aids

Keiichi Yasu*, Masato Hishitani, Takayuki Arai and Yuji Murahara

Department of Electrical and Electronics Engineering, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan

(*Received 1 April 2003, Accepted for publication 20 June 2003*)

Keywords: Hearing aids, Critical band, Speech intelligibility, Real-time processing, DSP
PACS number: 43.66.Yw [DOI: 10.1250/ast.25.61]

1. Introduction

There were many studies on the human auditory filter and the critical band (e.g., Fletcher [1] and Zwicker [2]). Patterson measured an auditory filter using notched-noise method [3]. Glasberg and Moore measured an auditory filter of hearing impaired and normal hearing people with notched-noise masker and reported that hearing impaired people had wider auditory filter than normal hearing people [4]. In the previous studies [5,6], a speech signal was split into 18 critical bands, and a set of odd-numbered bands was presented to the subject's right ear, while the rest was presented to the left ear. The speech signals became clearer for both normal hearing and hearing impaired subjects. This approach, however, is only useful when both ears have similar auditory characteristics. Therefore, we proposed an epochal method in which critical-band was compressed along the frequency axis in light of the shape of the auditory filters of hearing-impaired people [7] (Fig. 1).

In Exp. 1, two hearing-impaired people subjectively evaluated the quality and intelligibility of speech sounds using the Mean Opinion Score (MOS). In Exp. 2, they took an intelligibility test for an objective evaluation.

2. Algorithms

Two approaches were tested in our previous study [7]. In both approaches a speech signal was compressed toward the center of each critical band along the frequency axis. The first approach was based on a filter bank with a set of bandpass filters. The second was based on the fast Fourier transform (FFT). In this paper, we use the FFT-based approach.

First, an input speech signal was divided into frames with a frame length of 512 samples, a frame shift of 128 samples and windowed by the Hamming window. Next, the signal for each frame was transformed from the time domain to the frequency domain by FFT. After the amplitude and phase spectra of the FFT were calculated, a compressed amplitude spectrum was computed for each band. The compression was done for the amplitude spectrum toward the center of each critical band along the frequency axis. The compression rate ranged from 10% to 90%. Next, the amplitude spectrum after piece-wise compression was multiplied by the original phase spectrum. Finally, the overlap add (OLA) technique was applied to the IFFT of the product from the previous step to obtain the final signal. The stimuli were normalized by the RMS. The simulation of compression algorithm implemented by using "SIMULINK." Figure 2 shows the block-diagram of this technique.

3. Experiments and results

Two experiments were conducted. In Exp. 1, the quality and intelligibility of speech sounds were evaluated. In Exp. 2, an intelligibility score was evaluated. Two hearing-impaired subjects participated in the both experiments. Both subjects have hearing levels above 90 dB, are classified as profoundly hearing-impaired people and usually wear hearing aids. Before the experiment, we measured the shapes of critical bands of subjects with the notched-noise method [3], which we implemented with "SIMULINK." By measuring the shape of the auditory filter of hearing impaired, we confirmed that the critical band of hearing-impaired people was wider than for normal hearing people.

3.1. Experiment 1

We processed sounds along with sounds compressed by 20%, 40%, 60% and 80% using an FFT-based approach. 0% compression (appearing in Tables 1 and 2) corresponds to the original speech sounds. We used six sentences (three spoken by males, three by females) for the speech samples from "The Phoneme-Balanced 1000 Sentence Speech Database" by NTT-Advanced Technology. The experiment was controlled by a personal computer and was conducted in a soundproof room. Subjects made pair-wise subjective comparison between the original sounds and processed sounds, and they could play each sound as many times as needed. Then they evaluated the quality and intelligibility of the processed speech sounds using the Mean Opinion Score (MOS). In the MOS test, subjects were asked to evaluate sounds on five-point scale (1–5). Higher numbers indicated a greater degree, and Point 3 was set for the original. They evaluated 48 (4 compression rates $\times$ 6 sentences $\times$ 2 repetitions) times in all. The stimuli were presented in random order. Table 1 shows the average MOS in Exp. 1.

3.2. Experiment 2

Next, we gave subjects on an intelligibility test. We processed each speech sample from 10% to 90% compression in 10% steps. The speech samples were nonsense Vowel-Consonant-Vowel (VCV) syllables embedded in a Japanese carrier phrase. The speech samples were elicited from a native Japanese male. The vowels in each VCV syllable were /a/ and the consonant varied between each of the 14 Japanese consonants. Each stimulus was presented twice, and subjects were forced to choose one of 14 VCV's by clicking a button on the screen with a mouse. The 140 stimuli (14 VCV

*e-mail: k-yasu@sophia.ac.jp

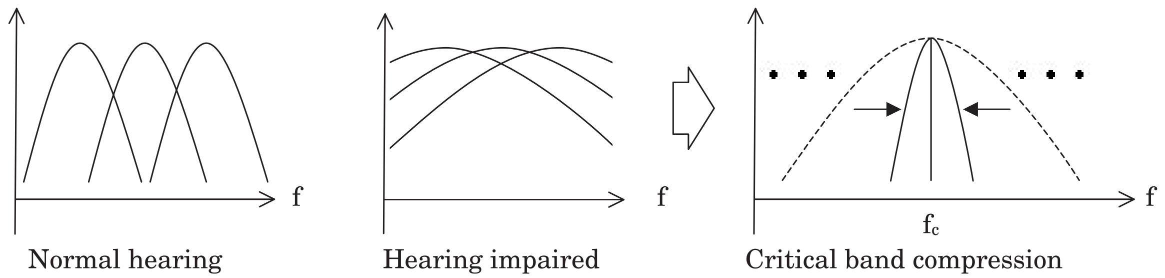

Fig. 1 Schematic figures of the auditory filter shape of the normal hearing (left) and hearing impaired (middle), and the shape of critical-band compressed toward the center frequency f_c.

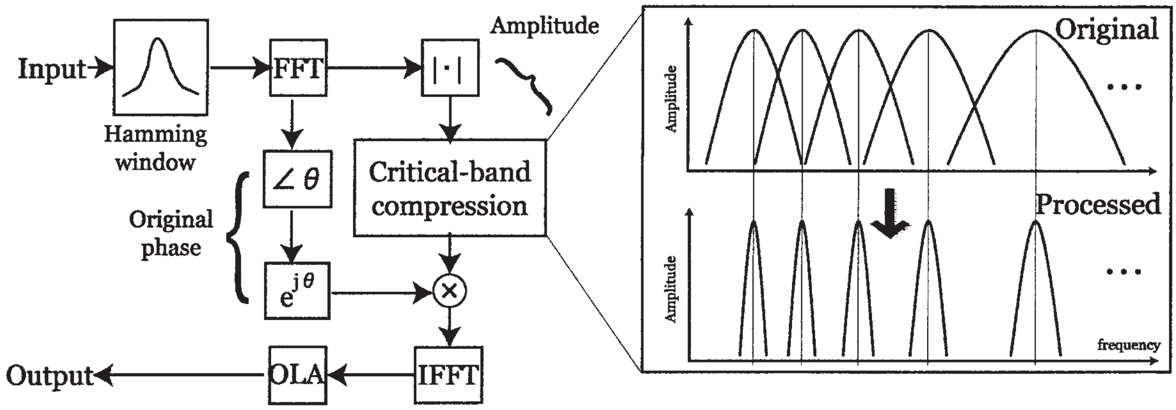

Fig. 2 Block diagram of the FFT-based approach.

Table 1 Average of Mean Opinion Score (MOS) for quality and intelligibility of speech.

Compression	Subject A		Subject B	
rate[%]	Quality	Intelligibility	Quality	Intelligibility
0	3.0	3.0	3.0	3.0
20	3.5	4.3	3.3	3.8
40	2.8	3.7	3.8	4.3
60	2.1	2.5	3.2	3.3
80	1.0	1.2	2.3	2.8

syllables $\times$ 10 compression rates, including original speech sounds of which compression rate is 0%) were presented randomly. Table 2 shows the average intelligibility score.

4. Discussion

Table 1 shows that the best scores for quality and intelligibility are at 20% compression for Subject A and 40% for Subject B. Table 2 shows that intelligibility scores are higher for Subject A above 30% compression. Subject B has the highest scores at 20% and 40% compression. From Exp. 1 and Exp. 2 we can surmise that adequate compression rates are from 20% to 30% for Subject A, and from 20% to 40% for Subject B. Ideal compression rates differ for each subject because each subject has a uniquely individual shaped auditory filter. Our proposed compression technique, then, is effective for individual subject, when the compression rate is adjusted properly.

Table 2 Average of intelligibility score [%].

Compression rate[%]	0	10	20	30	40	50	60	70	80	90
Subject A	57.1	57.1	57.1	64.3	71.4	64.3	71.4	64.3	71.4	64.3
Subject B	14.3	28.6	42.9	21.4	42.9	28.6	28.6	14.3	28.6	35.7

5. Conclusion

According to the results of Exp. 1 and Exp. 2, the compression along the frequency axis using FFT-based approach improved the quality and intelligibility of speech sounds for hearing impaired subjects. For the future, our task is light-weighting the program for the FFT-based approach and to achieve the real time simulation using DSP. Also, it is very important to discuss which critical band is important for quality, intelligibility of speech sounds for hearing impaired people. We are ready to transplant the programs from “SIMULINK” to DSP. Our goal is to develop the hearing assisted system. Confirming the usefulness of the frequency compression algorithms is a great step toward developing a hearing assisted system for hearing impaired.

References

[1] H. Flecther, “Auditory patterns,” *Rev. Mod. Phys.*, **12**, 47–65 (1940).

[2] E. Zwicker, G. Flottorp and S. S. Stevens, “Critical band width in loudness summation,” *J. Acoust. Soc. Am.*, **3**, 548–557

(1957).
[3] R. D. Patterson, "Auditory filter shapes derived with noise stimuli," *J. Acoust. Soc. Am.*, **59**, 640–654 (1976).
[4] B. R. Glasberg and B. C. J. Moore, "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.*, **79**, 1020–1033 (1986).
[5] D. S. Chaudhari and P. C. Pandey, "Critical band splitting of speech signal for reducing the effect of spectral masking in bilateral sensorineural hearing impairment," *Proc. ISSPA*, pp. 22–25 (1999).
[6] D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," *Proc. IEEE ICASSP 98*, Vol. 6, pp. 3601–3604 (1998).
[7] K. Yasu, K. Kobayashi, K. Shinohara, M. Hishitani, T. Arai and Y. Murahara, "Critical-band compression method for digital hearing aids," *Proc. Forum Acusticum*, Sevilla (2002).