CRITICAL-BAND COMPRESSION METHOD FOR DIGITAL HEARING AIDS

PACS: 43.66.Ts

Yasu, Keiichi; Kobayashi, Kei; Shinohara, Kohshi; Hishitani, Masato; Arai, Takayuki; and Murahara, Yuji;
Dept. of Electrical and Electronics Engineering, Sophia University
7-1 Kioi-cho, Chiyoda-ku, Tokyo
102-8554
Japan
Tel: +81-3-3238-3417
Fax: +81-3-3238-3321
E-mail: k-yasu@sophia.ac.ac.jp, arai@sophia.ac.jp

ABSTRACT

A new method of processing speech for digital hearing aids was proposed. In this method, each critical band was compressed along the frequency axis. To implement this, we tested two algorithms based on a filter bank (FB) approach and the fast Fourier transform (FFT). Five profoundly hearing-impaired subjects participated in a perceptual experiment. In both FB-based and FFT-based approaches, we confirmed that the processed speech became more intelligible and clearer for the subjects when using our method with the compression rate of 50%-80%. We also discussed the feasibility of implementing a real-time simulation system using "Simulink".

INTRODUCTION

Several techniques and algorithms have been reported for digital hearing aids [1,2]. We have focused research on the wider critical bands of hearing impaired people compared to normal hearing people [3]. Our goal is to develop the hearing aids in light of the critical band characteristics of hearing impaired people. In a former experiment [4,5], 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. In the present study, we propose a method for processing speech signals within a single channel to reduce interference between adjacent frequency bands. This compression is intended to compensate for the frequency selectivity of hearing impairments, as has been reported in previous studies. Our algorithm is unique in that it operates on a single rather than dual channel. Two approaches were tested. In each case the 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). Real-time processing using DSP is necessary to achieve our goal. For the first step, we created real-time simulation models for those approaches using "Simulink". In Section 2 we describe the principle behind each approach. In Section 3, experimental results are presented. Discussion follows in Section 4.

TWO APPROACHES FOR FREQUENCY COMPRESSOION

Filter-bank approach

The Filter-bank (FB) approach is based on a filterbank with a set of band pass filters. An input signal was passed through a FIR filter by the Kaiser window at each of 21 channel bands. The signal is given by $x_{org}[k,n]$ (k = 1, 2, ..., 21), and transformed into the Hilbert envelope $E_{org}[k,n]$, which contains the most essential acoustic information. The carrier component $\cos\theta_{org}[k,n]$ is given by

$$\cos\theta_{org}[k,n] = x_{org}[k,n] / E_{org}[k,n].$$

The same input signal was also passed through another FIR filter at each of 21 critical bands. The signal is given by $x_{cmp}[k,n]$. The bandwidth of each FIR filter ranged from 50% to 90% compared to the original bandwidth. The Hilbert envelope $E_{cmp}[k,n]$ and the carrier component $\cos\theta_{cmp}[k,n]$ are given by

$$\cos\theta_{cmp}[k,n] = x_{cmp}[k,n] / E_{cmp}[k,n].$$

Next, the product of the Hilbert envelope $E_{org}[k,n]$ obtained from the first step and the carrier component $\cos\theta_{cmp}[k,n]$ obtained from second step was taken for each band. The output signal $y[k,n]$ for channel k was given by

$$y[k,n] = E_{org}[k,n]\cos\theta_{cmp}[k,n].$$

Finally, a summation of all the outputs from the 21 bands was calculated to arrive at the final signal. The spectrograms before and after the FB-based processing are shown in Fig. 1 (left column). This figure shows the speech information is compressed into the center frequency of each critical band.

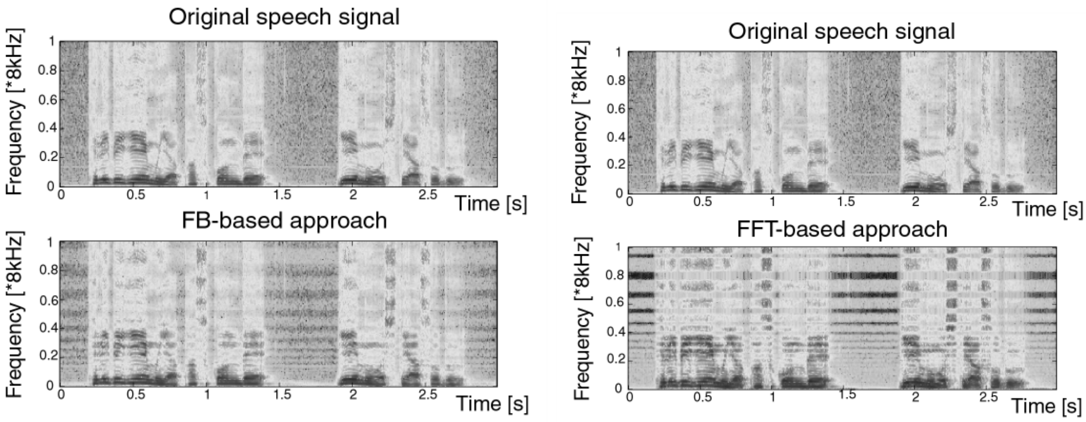

Fig. 1 - Spectrograms of the FB-based (left) and the FFT-based (right) approaches. Top panel: original and bottom panel: processed with compression rate of 50%.

FFT-based approach

The second approach is based on the fast Fourier transform (FFT). 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, and the compression rate ranged from 50% to 90%. Next, the partially compressed amplitude spectrum was multiplied by the original phase spectrum to re-synthesize a band-limited signal. Finally, the overlap add (OLA) technique was applied to the IFFT of the product from the previous step to obtain the final signal. Fig. 1 (right column) shows the spectrum before and after using this technique (compression rate is 50%).

EXPERIMENTS

Experiment I: FB-based and FFT-based approaches

Two MATLAB programs were written according to the approaches described in Section 2. The compression rate varied from 10% to 100% in 10% steps. A hearing impaired listener participated in a preliminary experiment, and we obtained the best performance with a compression rate of 70% for the filter-bank based approach and 60% for the FFT-based approach.

Five hearing-impaired subjects participated in the main experiment. All subjects have hearing levels above 90dB classified as profoundly hearing-impaired and usually wear hearing aids. We used five sentences (three spoken by males, two by females) from "The Phoneme-Balanced 1000 Sentence Speech Database Vol.2" by NTT-Advanced Technology. Subjects followed the same procedure for all 5 sentences and executed the set of sentences twice, once for the FB-based and once for the FFT-based approach. Before playing each sentence, we indicated which sentence was spoken on the list. Then, they were asked to evaluate each token for naturalness, clarity, and intelligibility with a five-point scale (1-5). Higher numbers indicated a greater degree, and point 3 was set for the original. We also asked them to make comments for each sound. Speech sounds were played through a pair of loudspeakers (BOSE Speaker 402) with an amplifier (SONY V777E) rather than headphones because of the dynamics of hearing range of the subjects.

Result

The results of FB-based and FFT-based approaches are given in Table 1.

Table. 1 - Result of Experiment I

Evaluation	FB-based approach						FFT-based approach					
	Subjects						Subjects					
	A	B	C	D	E	Average	A	B	C	D	E	Average
Natural	4.0	3.0	4.0	5.0	4.0	4.0	3.0	3.0	2.0	4.0	4.0	3.2
Clear	3.5	3.0	4.0	4.0	4.0	3.7	4.0	3.0	3.0	4.0	5.0	3.8
Intelligible	3.0	3.0	4.0	4.0	4.0	3.6	4.0	3.0	2.0	3.0	2.0	2.8

Experiment II: Experiment of a real-time system using "Simulink"

We created an interface by using "Simulink" to allow us to implement a real-time system simulation. In this system, output sounds were computed in real time, while MALTAB programs created waveform. We conducted an experiment of the FB-based approach for real-time system. Two hearing-impaired subjects participated in the experiment of the FB-based approach simulation. During this experiment, compression rates could be changed for each model from 10% to 100%, in 10% steps, by means of a graphical user interface. After pressing a "play" button they listened to the original and compressed sounds. They could change compression rate freely and listen to the original sounds at any time. Next, they were asked which compression-rate was best for listening and were asked to evaluate each token for naturalness, clarity, and general impression using a five-point scale (1-5) and point 3 was set for the original. We also asked them to make comments for each sound.

Result

The result of real-time simulation is given in Table 2.

Table. 2 - Result of Experiment II (Real-time simulation)

Sentences	Subject A				Subject B			
	Compression rate[%]	Evaluations			Compression rate[%]	Evaluations		
		Natural	Clear	General		Natural	Clear	General
No.1	90	3.0	4.0	3.0	50	4.0	5.0	4.0
No.2	60	3.5	3.0	4.0	60	4.0	4.0	4.0
No.3	80	3.0	4.0	3.0	40	4.0	5.0	4.0
No.4	60	3.0	5.0	3.0	50	5.0	5.0	5.0
No.5	60	3.0	4.0	4.0	50	4.0	4.0	4.0
Average	70	3.1	4.0	3.4	50	4.2	4.6	4.2

DISCUSSION

<u>Experiment I: FB-based and FFT-based approaches</u>

The results show that FB-based approach yields more natural, clear and intelligible signals than the original. Some of the more common comments from the subjects were that the outline of the signal becomes clearer and that the process clears up the signal especially in higher frequencies. We believe this is because the interference between adjacent critical bands was reduced in this approach. Furthermore, the asymmetry of the simultaneous masking effect on the frequency domain might help to make signals clearer in higher frequency range.

The FFT-based approach also yields clearer signals than the original. The subjects reported that the FFT-based sounds were clearer than the FB-based sounds, although they were less natural and less intelligible. This can be explained from the different spectrograms in Fig. 1. The FFT-based approach has more distinct contrast than the FB-based approach because zeros were padded on the frequency domain. In the case of the FB-based approach, some energy is leaking between the bands. Thus, for clarity, the average score of the FFT-based is higher than for the FB-based, although that was not the case for naturalness and intelligibility (possibly for the same reason).

<u>Experiment II: Real-time simulation</u>

From the real-time simulation, we observed a similar tendency as in Experiment I (the MATLAB simulation). Furthermore, the processed signals were more preferable than the original (the range of compression rate that the subjects selected varies form 50-80%). This suggests that the effective system (the FB-based system, at least) can be implemented on a digital signal processor (DSP). One of the biggest advantages of real-time simulation is that it enables us to change the compression rate at any time. We need to continue to develop a real-time system for the FFT-based approach as well as the FB-based approach, and conduct experiments for more subjects.

CONCLUSION

We proposed a new method in which critical-band was compressed along the frequency axis. To implement this, two approaches, the FB-based and the FFT-based approaches, were tested. As a result of the experiments for hearing impaired people, there was the improvement in the quality of sound. For the FB-based approach, naturalness, clarity and intelligibility were improved. For the FFT-method, clarity was improved. According to this result, both the FB-based and the FFT-based approaches, are effective methods for creating speech sounds for hearing impaired people. We also successfully developed a real-time simulation system with variable compression rate for the experiment. This achievement is the big step toward developing a real-time system using a DSP for researching hearing aids.

ACKNOWLEDGEMENTS

I would like to thank for Terri Lander for the useful comments on writing this paper.

BIBLOGRAPHIACAL REFERENCES

[1] T. Stetzler, N. Magotra, P. Gelabert, P. Kasthuri and S. Bangalore "Low power real-time programmable DSP development platform for digital hearing aids," *Proc. of IEEE ICASSP,* Vol.4 , pp.2339-2342, 1999.

[2] N. Magotra, P. Kasthuri, Y. Yang, R. Whitman and F. Livingston, "Multichannel adaptive noise reduction in digital hearing aids," *Proc. of IEEE ISCAS*, Vol.6 , pp.582-585 1998.

[3] B.C.J. Moore, *An Introduction to the Psychology of Hearing*, 3rd edition, Academic Press, London, 1989.

[4] 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. of ISSPA*, pp.22-25, 1999.

[5] D.S. Chaudhari and P.C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," *Proc. of IEEE ICASSP*, Vol.6, pp.3601-3604, 1998.