

REAL TIME SPEECH PROCESSING FOR DICHOTIC PRESENTATION FOR BINAURAL HEARING AIDS

D.S. Chaudhari and P.C. Pandey
Biomedical Engineering & Electrical Engineering
Indian Institute of Technology, Bombay
Powai, Mumbai 400 076

Abstract

In case of sensorineural impairment, the reduced frequency resolution in the peripheral auditory system due to spread of masking of a frequency component by adjacent frequency components degrades speech perception. Speech processing by a bank of critical band filters and adding signals from alternate bands for presenting odd numbered critical bands to one ear and even numbered ones to the other is likely to reduce the effect of spread of masking along the cochlear partition, and therefore it may help in improving speech intelligibility. This scheme has been implemented for real-time processing for use as a binaural hearing aid. The implementation consists of two TI/TMS320C50 DSP processor based boards, each having 14-bit ADC and DAC. Speech signal after amplification and signal conditioning is connected as input to both the boards. One board uses filter coefficients corresponding to the even numbered critical bands, and the other board uses filter coefficients for the odd numbered critical bands. For filtering, the FIR filter coefficients for the two channels are loaded to the corresponding DSP boards, using serial port from a PC. Twelve English consonants were used in carrying out listening tests in vowel-consonant-vowel (VCV) syllables presented to a normal hearing subject with simulated sensorineural hearing loss. The test results show an increase of 8 % in the recognition score under adverse listening condition. Information transmission analysis of confusion matrices, for various features, show maximum improvement for the place feature.

Introduction

Loss of frequency resolution due to spread of masking is one of the characteristics of the sensorineural hearing impairment. This loss of frequency resolution results

from masking of a frequency component by adjacent components during the auditory processing along the basilar membrane (or along the cochlear partition). For persons with bilateral sensorineural impairment with residual hearing in both ears, the effect of spectral masking can be reduced by splitting the speech into two different signals for presentation to the two ears, a scheme known as dichotic presentation. The ability of humans to receive and perceptually combine signals from both the ears for improving recognition of speech is well established [1]. Lunner et al [2] tested the scheme of splitting speech using 8-channel constant bandwidth for dichotic presentation and found improvements in speech perception.

We are investigating the effectiveness of speech processing by a bank of critical band filter and adding the odd numbered and even numbered critical bands for dichotic presentation for improving speech reception in cases of bilateral sensorineural hearing impairment with some residual hearing. We implemented this scheme using off-line processing of speech with eighteen critical bands with normal hearing subjects with simulated sensorineural hearing loss of varying degrees on five subjects in age group of 21 to 40 years [3] and ten hearing impaired subjects in the age group of 18 to 58 with mild-to-very severe bilateral sensorineural hearing loss [4]. For experimental evaluation, listening tests were carried out for measuring confusion among the set of twelve consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in vowel-consonant-vowel (VCV) context with vowel /a/. The scheme was found helpful in improving speech quality, recognition scores, and transmission of features, particularly the place feature indicating the usefulness of the scheme for better reception of the spectral characteristics.

On the basis of these results, from off-line processing, we implemented the scheme for real time processing for use as a binaural hearing aid. In this paper, we present

the scheme of implementation and preliminary results of listening tests with a normal hearing subject with simulated sensorineural hearing loss.

Implementation

The 3 dB cutoff frequencies of the two filters correspond to the critical bands selected as per the auditory filter bands described by Zwicker [5], and these are given in Table 1.

The real time implementation has been done by using two TI/TMS 320C50 “starter kit” DSP boards. Each board consists of processor along with an analog interface circuit (AIC) with 14-bit ADC and DAC and a programmable timer which can be used for setting the sampling rate. The filter program and coefficients can be loaded into the program RAM on the DSP chip using serial port interface. Our processing set-up, as shown in the Figure 1, consists of an input low pass filter ($f_p = 4.6$ kHz, $f_s = 5.0$ kHz, pass band ripple < 0.3 dB, stop band attenuation > 40 dB), two DSP boards operating with sampling rate = 10 k samples/sec, and two audio amplifiers. Each filter magnitude response (with alternating channel bands as given in Table 1) is approximated as a linear phase FIR filter response with 128 coefficients by windowing technique [6]. The program for the filter with appropriate filter coefficients are loaded on to the DSP boards. It is to be noted that no data transfer takes place between the two boards. The response of two filters has been tested by applying swept sine wave and by obtaining averaged magnitude spectra of the filter outputs for random noise input.

Experimental Method

The scheme was tested on a normal hearing subject by

TABLE 1: Pass bands (the 3 dB cutoff frequencies, in kHz) for multiple band pass filtering for the two channels

Channel 1		Channel 2	
Band #	Pass band	Band #	Pass band
1	- -0.20	2	0.20-0.30
3	0.30-0.40	4	0.40-0.51
5	0.51-0.63	6	0.63-0.77
7	0.77-0.92	8	0.92-1.08
9	1.08-1.27	10	1.27-1.48
11	1.48-1.72	12	1.72-2.00
13	2.00-2.32	14	2.32-2.70
15	2.70-3.15	16	3.15-3.70
17	3.70-4.40	18	4.40- -

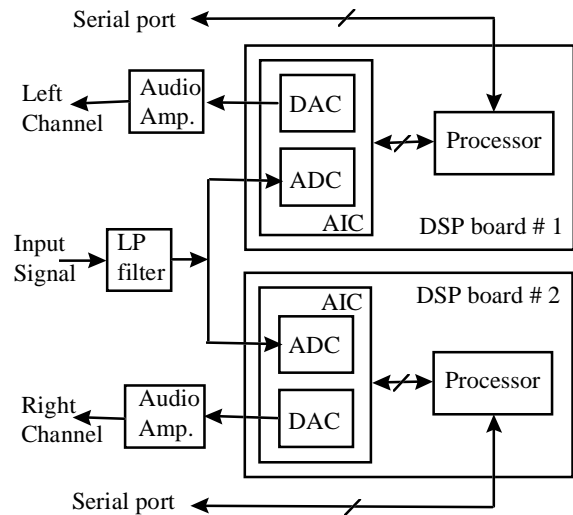


FIG. 1: Speech processing using two DSP boards for dichotic presentation

simulating sensorineural hearing impairment. Broad band noise can be used for simulating different aspects of sensorineural hearing loss in normal hearing subjects [7,8,9]. In our experiments, we have used Gaussian white noise limited to band of speech signal as masking noise at signal-to-noise ratio of 3 and 0 dB. The noise was added in such a way that the overall level remains unchanged. The presentation level was kept constant and at subject’s most comfortable listening level. Due to repetitive nature of the test, an automated test administration system was used [3,4]. The test was administered for (a) unprocessed speech diotically presented and (b) processed speech dichotically presented. The subject was seated in acoustically isolated room during the testing. The stimuli were presented using a pair of Telephonics TDH-39P headphones.

The listening tests were carried out for obtaining stimulus-response confusion matrices among the set of twelve consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in vowel-consonant-vowel (VCV) context with vowel /a/.

Results and Discussion

The recognition scores obtained from the confusion matrices averaged over five tests are given in Table 2. Improvement in the recognition score was observed in case of processed speech as opposed to unprocessed speech. For studying the reception of specific consonant features, the stimuli-response matrices were subjected to

TABLE 2: Percentage recognition scores for subject RBK, averaged across five tests, for 12 consonants in VCV context, for unprocessed and processed speech for SNR = 3 dB and 0 dB.

SNR	Unproc. Speech		Proce. Speech		p value for 2-tailed test [10]
	mean	s.d.	mean	s.d.	
3 dB	51.3	1.4	59.3	1.9	< 0.001
0 dB	43.0	3.8	51.0	1.9	< 0.01

TABLE 3: Percentage relative information transmitted for subject RBK, for set of 12 consonants in VCV context.

Features	SNR = 3 dB		SNR = 0 dB	
	Unproc. speech	Proc. speech	Unproc. speech	Proc. speech
Voicing	69	70	66	59
Manner	63	61	52	55
Place	10	27	9	16
Overall	57	63	55	55

information transmission analysis [11], the results are given in Table 3. Under both SNR conditions, we see an improvement in the reception of place features due to processing. As the place feature depends on frequency resolving capacity of the auditory processing, it can be concluded that the implemented scheme has improved the reception of spectral characteristics.

Acknowledgments

This work is part of an R&D project sponsored by AICTE. We are grateful to Dr. M. N. Nagaraja, Dy. Director (T), AYJ National Inst. Hearing Handicapped, Mumbai for help in setting up the experimental facilities for the listening tests. The DSP boards used were provided by Texas Instruments as part of Texas Instruments DSP Elite Lab.

References

1. B. C. J. Moore, *An Introduction to Psychology of Hearing*, New York: Academic, 1982.
2. T. Lunner, S. Arlinger, and J. Hellgren, "8-channel digital filter bank for hearing aid use: preliminary results in monaural, diotic and dichotic modes", *Scand. Audiol. Suppl.* 38, pp. 75-81, 1993.
3. D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," *Proc. of Int. Conf. on Acoust. Speech and Signal Processing (ICASSP-98)*, Seattle, Wash., 1998 (to be published).
4. D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of speech signal using critical filter bank for bilateral sensorineural hearing impairment," *Proc. of 16th Int. Congress on Acoust. (ICA)*, Seattle, Wash., 1998 (to be published).
5. E. Zwicker, "Subdivision of audible frequency range into critical bands (Frequenzgruppen)", *J. Acoust. Soc. Am.*, vol. 33, p. 248, 1961.
6. Preethi Kasthuri, *Real Time Digital Filter with Tunable Frequency Response Using TMS320C50 Processor*, project report, Signal Processing and Instrumentation Lab., EE Dept., IIT Bombay, 1997.
7. S. DeGennaro, L. D. Braid, and N. I. Durlach "Study of multi-band syllabic compression with simulated sensorineural hearing loss", *J. Acoust. Soc. Am.*, vol. 69, S16, 1981.
8. H. Fletcher, "The perception of sound by deafened persons", *J. Acoust. Soc. Am.*, vol. 24, pp. 490-497, 1952.
9. W. Jesteadt, Ed., *Modeling Sensorineural Hearing Loss*, Mahwah, New Jersey: Lawrence Erlbaum, 1997.
10. G. W. Snedecor and W. G. Cochran, *Statistical Methods*, Ames, Iowa: The Iowa State University Press, 1980.
11. G. A. Miller and P. E. Nicely, "An analysis of perceptual confusions among some English consonants", *J. Acoust. Soc. of Am.*, vol. 27 (2), pp. 338-352, 1955.