

CRITICAL BAND SPLITTING OF SPEECH SIGNAL FOR REDUCING THE EFFECT OF SPECTRAL MASKING IN BILATERAL SENSORINEURAL HEARING IMPAIRMENT

D. S. Chaudhari¹ and P. C. Pandey²

¹School of Biomedical Engineering

²Department of Electrical Engineering
Indian Institute of Technology, Bombay

Powai, Mumbai 400 076, India

Email: pcpandey@ee.iitb.ernet.in

ABSTRACT

Persons with sensorineural hearing impairment face a particular problem in view of decrease in frequency resolving capacity of the auditory system due to spread of spectral masking along the cochlear partition. Filtering speech signal by a filter bank and adding signals from alternate bands and presenting it to the two ears is likely to reduce the effect of spectral masking and thus improving the speech intelligibility. The scheme was implemented in real time processing for use as a binaural hearing aid. Processing for each channel is done by linear phase FIR filter having a magnitude response with pass bands corresponding to alternate critical bands, the magnitude response for the two channels being complemented. For both the filters, the magnitude response was approximated with 128 coefficients using frequency sampling technique of linear phase FIR filter design. The implementation was done on two TI/TMS320C50 based DSP boards, each having 14-bit ADC and DAC. Twelve English consonants were used for carrying out listening tests in vowel-consonant-vowel (VCV) and consonant-vowel (CV) syllables presented to six hearing impaired subjects with bilateral sensorineural hearing loss. Information transmission analysis of confusion matrices for various features show maximum improvement for the place feature.

1. INTRODUCTION

The sensorineural hearing impairments are characterized by increase in thresholds of hearing, reduction in dynamic range of hearing, degradation of temporal resolution and increase in temporal masking, and degraded frequency selectivity due to increase in spectral masking. Many hearing aids incorporate frequency compensation and amplitude compression. Some of the techniques currently being investigated are based on signal processing schemes such as spectral transposition, speech enhancement using the properties of "clear" speech. These are likely to

increase the performance of hearing aids for persons with residual hearing as well as that of other sensory aids like cochlear prostheses and vibro-tactile aids used by profoundly hearing impaired.

The ability to perceptually combine the binaurally received signals from the two ears improve speech reception under adverse listening conditions [10]. Binaural listening offers better overall sound quality, clear speech intelligibility, more relaxed listening, and it helps in source localization [10] [11] [8]. Many studies have supported the use of binaural hearing aid over monaural under different listening conditions [5] [1] [7]. There is a possibility that splitting speech into two complementary parts on the basis of frequency and presenting it binaurally might increase its intelligibility. Thus, the binaural hearing aid which can split the speech signal on the basis of frequency components for presenting over two ears can be helpful to hearing impaired persons with residual hearing.

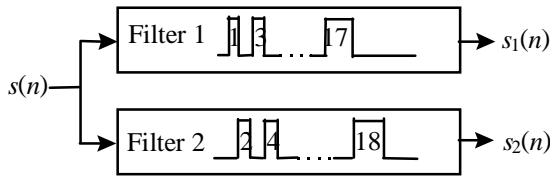
The hearing aids, based on speech processing for binaural dichotic (different signal in each ear) presentation can reduce the effect of spectral masking. The objective of the present study is to investigate a scheme of splitting speech into two signals, with complementary spectra, for binaural dichotic presentation, in order to lessen the effect of reduced auditory frequency selectivity and thereby to improve speech intelligibility.

We implemented this scheme using off-line processing of speech with eighteen critical bands corresponding to auditory filters, on seven normal hearing subjects with simulated sensorineural hearing loss of varying degrees in age group of 21 to 40 years and ten hearing impaired subjects in the age group of 18 to 58 with 'mild'-to-'very severe' bilateral sensorineural hearing loss [3] [4]. 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, the scheme was implemented for real-time processing speech signal for use as a binaural hearing aid. The listening tests were carried out with subjects having bilateral sensorineural impairment.

2. IMPLEMENTATION

The processing scheme was implemented in which the gain of all the filter bands was kept constant. In our implementation, 18 critical bands corresponding to auditory filters, were used as described by Zwicker [14] which covers the main part of the speech spectrum. The 3-dB cut-off frequencies of the two comb filters correspond to critical bands along with



Band	Passband frequency	Band	Passband frequency
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- --

Figure 1. Splitting of speech signal using two comb filters. The filter magnitude response is shown in each block (the table shows 3-dB cut-off frequencies).

the scheme of implementation are shown in Fig. 1. The real-time processing was done using two DSP boards based on 16-bit fixed point processor, TI/TMS320C50 [12] [13]. 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 AIC also has a low pass filter at the input of ADC and at the output of DAC. The processing set-up as shown in Fig. 2 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 of 10 k samples/s, and two audio

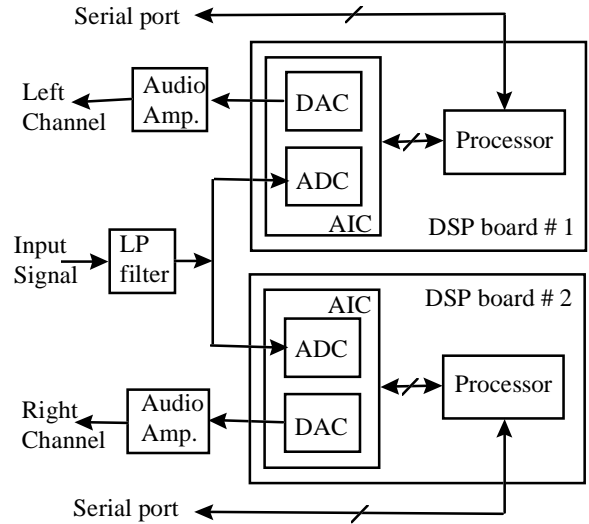


Figure 2. Speech processing using two TI/TMS320C50 based DSP boards for dichotic presentation.

amplifiers. For both the comb filters, the desired magnitude response was approximated with 128 coefficients using frequency sampling technique of linear phase FIR filter design. The filter program and coefficients can be loaded into the program RAM on the DSP chip using serial port interface. No data transfer takes place between the two boards. The filter responses were verified by applying swept sine wave and by obtaining averaged magnitude spectra of the filter outputs for random noise input.

3. EXPERIMENTAL METHOD

The scheme was tested on bilateral sensorineural hearing impaired subjects. The subjects who participated were from different part of India and they had no difficulty in clearly recognizing the test stimuli. Six hearing impaired subjects participated in the experiments. The subjects were right handed and familiar with the English. The subjects had 'mild'-to-'very severe' bilateral sensorineural hearing loss. The PTA (pure tone average hearing threshold level, test frequencies: 0.5, 1, and 2 kHz) difference between right and left ear was from 4 to 30 dB. The presentation level was kept to the subject's most comfortable listening level and presentation were made using a pair of headphones (Telephonics TDH-39P). Tests were carried out using an automated test administration system [2], to evaluate advantages of the processed signal presented dichotically over the unprocessed signal diotically (same signal in both the ears) presented. During the testing, the subject was seated in acoustically isolated room.

The listening tests were carried out for obtaining stimulus-response confusion matrices among the close set of twelve English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in VCV and CV contexts, the vowel being /a/ as in 'father.' The number of stimuli were restricted to 12, so that they can be conveniently accommodated on subject's screen in the computerized test administration system. The syllables were spoken by a male speaker.

4. RESULTS AND DISCUSSION

Listening tests were conducted with six hearing impaired subjects in VCV and CV contexts, for comparing the diotic presentation with unprocessed speech and the dichotic presentation with constant gain filter implementation. A compilation of subjects' qualitative assessment about the test stimuli indicated that perceptual quality improved due to processing for dichotic presentation.

The subjects showed decrease in response time due to processing which indicates an improvement in listening condition with processing. Improvement in the recognition score was observed in the case of processed speech as opposed to unprocessed, indicating the usefulness of the implemented scheme. The confusion matrices were subjected to information transmission analysis [9] in order to study the

Table 1. Percentage recognition scores for subject SG, averaged across five tests, for 12 consonants, for unprocessed and processed speech.

Con-text	Unproc. Speech		Proc. Speech		p value for 1-tailed test
	mean	s.d.	mean	s.d.	
VCV	60.7	3.8	75.0	1.7	< 0.0005
CV	74.3	3.0	86.7	3.4	< 0.0005

reception of specific consonant feature. All subjects have shown improvements in relative information transmission in manner, place, and voicing features and overall information transmission.

The recognition scores, for a typical subject in both the contexts, obtained from the confusion matrices averaged over five tests are given in Table 1. The percentage improvement in scores are 14.3 and 12.4 in VCV and CV context respectively. In both the contexts, the subject showed highly significant improvement ($p < 0.0005$). The results of information transmission analysis for the same subject are given in Table 2. In both the contexts, the overall

Table 2. Percentage relative information transmitted for subject SG, for set of 12 consonants.

Features	VCV		CV	
	Unproc. Speech	Proc. Speech	Unproc. Speech	Proc. Speech
Voicing	60	78	100	100
Manner	52	64	66	79
Place	25	48	43	62
Overall	65	72	78	84

improvements are contributed by all the three features of voicing, manner, and place. The maximum improvement is observed in case of place feature which is 23% and 19% in VCV and CV contexts respectively. As the place information is related to frequency resolving capacity of the auditory processing, one can say that the implemented scheme has reduced the effect of spectral masking.

5. REFERENCES

- [1] Cox R. M. and Bisset J. D. "Relationship between two measures of aided binaural advantage". *Journal of Speech and Hearing Disorders*, 49:399-408, 1984.
- [2] Chaudhari D. S. and Pandey P. C. "Splitting of speech signal by critical band filtering for bilateral sensorineural hearing impairment". *Proceedings of the 18th Conference of Indian Association of Biomedical Scientists*, New Delhi, 1997, pages V/4-5.
- [3] Chaudhari D. S. and Pandey P. C. "Dichotic presentation of speech signal with critical band filtering for improving speech perception". *Proceedings of International Conference on Acoustic, Speech, and Signal Processing*, Seattle, Washington, 1998, paper AE3.1.
- [4] Chaudhari D. S. and Pandey P. C. "Dichotic presentation of speech signal using critical filter bank for bilateral sensorineural hearing impaired". *Proceedings of 16th International Congress on Acoustics*, Seattle, Washington, 1998, vol. 1, pages 213-214.
- [5] Dirks D. and Carhart R. "A survey of reactions from users of binaural and monaural hearing aids". *Journal of Speech and Hearing Disorders*, 27(4):311-322, 1962
- [6] Dubno J. R. and Dirks D. D. "Auditory filter characteristics and consonant recognition for hearing impaired listeners". *Journal of the Acoustical Society of America*, 85:1666-1675, 1989.

- [7] Hawkins D. B. and Yacullo W. S. "Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation". *Journal of Speech and Hearing Disorders*, 49:278-286.
- [8] Kollmeier B., Hohmann V., and Peissig J. "Digital signal processing for binaural hearing aids". *Proceedings of International Congress on Acoustics (ICA)*, 1992, Beijing, pages H3-4.
- [9] Miller G. A. and Nicely P. E. "An analysis of perceptual confusions among some English consonants". *Journal of the Acoustical Society of America*, 27(2):338-352, 1955.
- [10] Moore B. C. J. *An introduction to the psychology of hearing*. Academic, London, 1982.
- [11] Pickles J. O. *An introduction to the physiology of hearing*. Academic, London, 1982.
- [12] Texas Instruments. *User's guide, TI-TMS320C5X digital signal processor products*. Texas Instruments, USA, 1993.
- [13] Texas Instruments. *Starter kit user's guide, TI-TMS320C2X digital signal processor products*. Texas Instruments, USA, 1994.
- [14] Zwicker E. "Subdivision of audible frequency range into critical bands (Frequenz-gruppen)". *Journal of the Acoustical Society of America*, 33:248, 1961.