7th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2003), Orlando, USA, July 27-30, 2003

Dichotic Presentation for Binaural Hearing Aids Using Perceptually Balanced Critical Bandwidth Based Comb Filters

Alice N. Cheeran(1) and Prem C. Pandey (2)
(1)Biomedical Engineering Group, (2)Dept. of Electrical Engg.
Indian Institute of Technology, Bombay
Powai, Mumbai-400 076, India
E-mail: {alicenc, pcpandey}@ee.iitb.ac.in

ABSTRACT

Splitting of speech using comb filters with complementary magnitude response and presenting them to the two ears has helped in improving the perception of speech in sensorineural hearing impaired persons. The pair of comb filters with 18 bands based on auditory critical bandwidths over 5 kHz band, were designed initially with sharp transition between pass and stop bands. Further these filters were optimized to obtain minimum perceived spectral distortion and perceptual balance of intensity at band crossovers. Such comb filters, realized as 256-coefficient linear phase FIR filters, with pass band ripple of 1 dB, stop band attenuation of 30 dB, transition width of 78 to 117 Hz and filter magnitude responses at crossovers adjusted to lie between 4 and 6 dB, provided better recognition scores than that with sharp transitions, when tested on normal subjects with simulated hearing loss. Present investigation involves evaluation of spectral splitting scheme with comb filters with sharp transitions and with perceptually balanced comb filters on bilateral sensorineural hearing-impaired subjects. Further evaluation was conducted by incorporating an additional filter, which partially compensates for the frequency dependent shift in hearing thresholds. The results show improvement of recognition scores, and reception of consonantal features particularly place, frication and duration, with maximum improvement for the combined scheme.

Keywords: Sensorineural Hearing Loss, Spectral Masking, Binaural Dichotic Presentation.

1. INTRODUCTION

Sensorineural hearing impairment is characterized by frequency dependent shifts in hearing threshold, loudness recruitment, reduced frequency and temporal resolution, and increased spectral and temporal masking [1], [2]. Increased spectral masking leads to reduction in spectral contrast due to smearing of peaks and valleys in the speech spectrum and a reduction in speech clarity. Masking takes place primarily at the peripheral level, while integration of information takes place at higher levels in the auditory system. It has been reported earlier that splitting speech in a complementary manner for binaural dichotic presentation is a possible solution to reduce the effect of increased masking for persons with residual hearing in both the ears, i.e. for persons who can wear binaural hearing aids [3]-[6].

No significant improvement is reported in an earlier work where splitting of speech was carried out by constant bandwidth complementary comb filter implemented using analog delay and adder [3]. An overall improvement of 2 dB in speech-to-noise ratio for dichotic over diotic was obtained with comb filters with eight channel filter bank having constant bandwidths of 700 Hz, designed with complementary interpolated linear phase FIR filters [4]. Later, in an investigation by Chaudhari and Pandey [5] a pair of comb filters with complementary magnitude responses based on auditory critical bandwidths [7] were used. The bandwidths were constant at 100 Hz for center frequencies below 500 Hz and were 15-17 % of the center frequency in the range of 1-5 kHz. The filters were designed with the consideration of relatively flat response in the pass band, high attenuation in stop bands and sharp inter-band transition. Evaluation of the scheme on bilateral sensorineural hearing-impaired subjects showed significant improvement in recognition scores and perception of consonantal features particularly the place feature. Further, magnitude response of the comb filters was adjusted to partially match the audiogram of the hearing impaired subjects [6]. Additional improvement was reported but no specific contribution from place feature.

A pair of comb filters based on 18 critical bands over 5 kHz range was designed as 256-coefficient linear phase FIR filters with sampling rate of 10 kSa/s, using frequency sampling technique, resulting in transition width of 39 Hz, pass band ripple of 3 dB, stop band attenuation of 11 dB, and inter-band crossovers ranging over 0 to 10 dB. When a slowly swept sinusoidal tone was processed with this pair of comb filters, intensity variation could be noticed due to pass-band ripple.

Further, at band transitions, change in intensity was perceived. This indicated a need for designing filters with low pass band ripple, high stop band attenuation and band transitions with gains adjusted for perceptual balance of intensity. For design of such filters, frequency sampling technique was used in an iterative manner, treating one or two transition samples as unconstrained [8]. Listening tests with these filters established that inter-band crossover gain adjusted within 4-6 dB of the pass band gain resulted in perceptual balance and 1 dB ripple in the pass band was found to be perceptually acceptable [9]. Based on these results, perceptually balanced comb filters were designed. These filters have transition width of 78 to 117 Hz and stop band attenuation of 30 dB.

In experimental evaluation of binaural dichotic presentation using spectral splitting, on normal hearing subjects with simulated sensorineural hearing loss, the perceptually balanced filters showed higher improvement in recognition scores and information transmission of consonantal features compared to comb filters with sharp transition. Listening tests were subsequently carried out on persons with moderate bilateral sensorineural loss. For these subjects, effects of cascading a linear phase filter with magnitude response shaped to partly match the audiogram as a way of partial compensation for frequency dependent shifts in hearing thresholds is also investigated.

2. PROCESSING SCHEMES

The processing schemes for spectral splitting were implemented for evaluation through listening tests: (i) comb filters with sharp transition between bands (denoted as SpA), (ii) perceptually balanced comb filters (denoted as SpB) and (iii) perceptually balanced comb filters cascaded with filters to partly match the audiogram. (denoted as SpC).

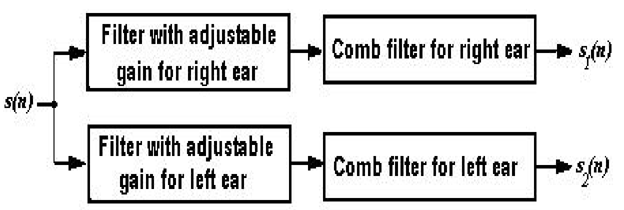

Fig. 1. Schematic representation of the spectral splitting scheme: filters to partially compensate for frequency dependent shift in hearing threshold cascaded with comb filters to overcome the effect of increased spectral masking.

Fig. 1 shows the block diagram of the processing scheme SpC. For processing schemes SpA and SpB only the comb filters are used. All the filters used are linear phase FIR filters with 256 coefficients designed using frequency-sampling techniques.

The magnitude response of the comb filter used in SpA is shown in Fig. 2. The magnitude response of the perceptually balanced comb filters designed for low perceived spectral distortion used in scheme SpB is shown in Fig. 3.

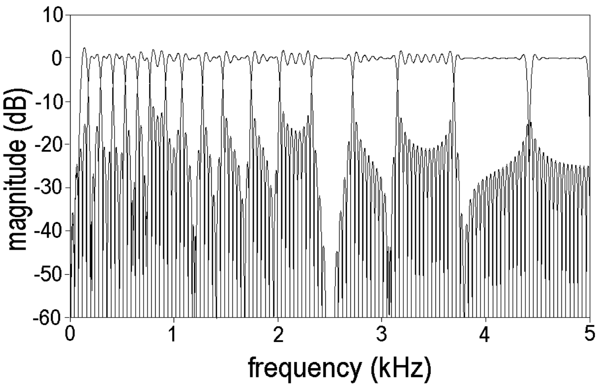

Fig. 2. Magnitude response of the pair of comb filters with sharp transition between bands. S.R. = 10 kSa/s, pass band ripple < 3 dB, stop band atten. > 10 dB, transition width = 39 Hz.

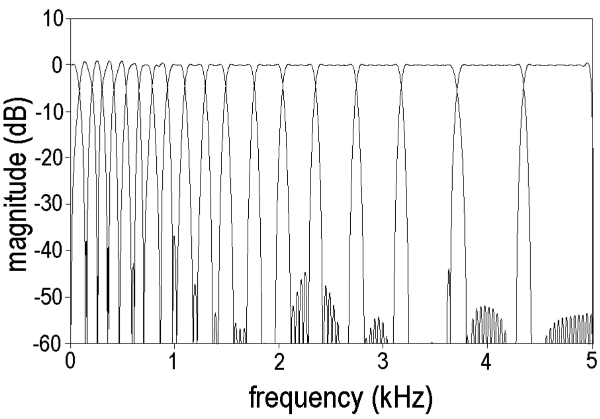

Fig. 3. Magnitude response of the pair of perceptually balanced comb filters. S.R. = 10 kSa/s, pass band ripple < 1 dB, stop band atten. > 30 dB, transition width = 78 – 117 Hz.

The third processing scheme SpC had two filters with gain adjustment to partially compensate for the frequency dependent shift in hearing threshold, which is cascaded with the comb filters used in the scheme SpB. The frequency response of these filters with adjustable gain in the range of –3 to +3 dB was obtained by interpolating the pure tone audiogram of the subject for the two ears. The magnitude in dB at different frequencies are obtained as

$$A_a(f) = -3 + 6\frac{(\alpha(f) - \alpha_{min})}{(\alpha_{max} - \alpha_{min})} \qquad (1)$$

where α (*f*) is the value of hearing loss in dB, α_{min} and α_{max} are respectively the minimum and maximum hearing loss over 125 Hz to 5 kHz frequency range. The frequency compensation was restricted between –3 and +3 dB, keeping in view the limited dynamic range of the hearing impaired subjects. Fig. 4 shows the pure tone audiogram of one of the subjects (SM) and the magnitude response of the filter used to partially compensate for the frequency dependent shift in threshold of hearing for the right and left ears.

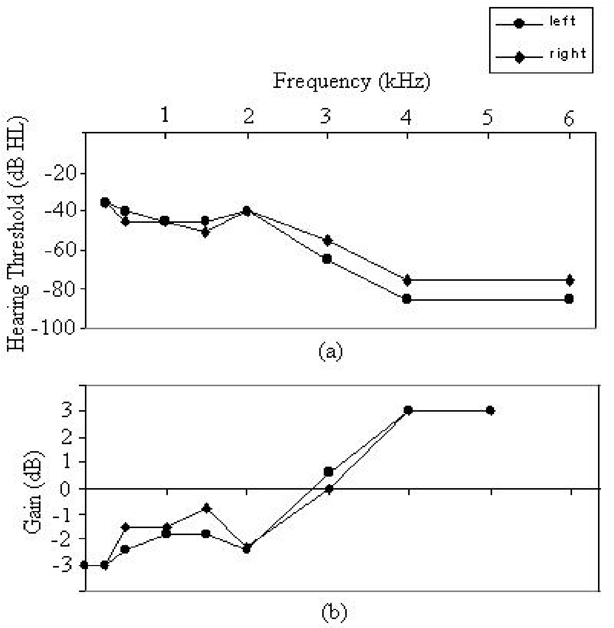

Fig. 4. Example of frequency compensation filter. (a) pure tone audiograms and (b) the filter magnitude response for the two ears (subject: SM).

3. LISTENING TESTS FOR EVALUATION

The experimental evaluation of the schemes was carried out by conducting listening test, on five subjects with bilateral mild to severe sensorineural hearing loss. Subject AB was a 61-year old male, having low frequency loss. Subjects KS and BS were 32 and 38 year old females with high frequency loss. The loss was symmetrical for subject BS, whereas a small asymmetry was noticed for subject KS for frequencies below 1 kHz. Subject SK, 41 year old male had symmetrical loss, progressively increasing from 65 dB HL at low frequency to 105 dB HL at high frequency. Subject SM, 59-year old male had moderate symmetrical loss of about 40 dB HL up to 2 kHz and progressively increasing to 85 dB HL.

Test material consisted of a closed set of 12 English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in vowel-consonant-vowel (VCV) context with vowel /*a*/ as in "father". The test stimuli were recorded at 10 kSa/s and processed offline by the three schemes. A computerized test administration system, consisting of a PC interfaced through RS232 serial port to the subject terminal placed in an acoustically isolated chamber was used. The subject sitting in the acoustically isolated chamber listened the test material presented through headphones (Telephonics TDH-39P) and responded through the terminal placed there. Presentation of the processed sounds to the headphones was done through two DAC ports of a data acquisition card, smoothening low pass filter, and audio amplifiers, with gain adjusted for most comfortable listening level for each ear.

The responses were stored in the form of stimulus-response confusion matrices for each of test conditions. The percentage recognition score, response time and relative information transmitted for different consonantal features were analyzed. The recognition scores of individual subjects were subjected to one-tailed t-tests to obtain statistical significance of processing.

The consonants were grouped according to the articulatory features [10] for obtaining the relative information transmission analysis of features. The features selected for study were voicing (voiced: / *b d g m n z v* / and unvoiced: / *p t k s f* /), place (front: / *p b m f v* /, middle: / *t d n s z* /, and back: / *k g* /), manner (oral stop: / *p b t d k g* /, fricative: / *s z f v* /, and nasals: / *m n* /), nasality (oral: / *p b t d k g s z f v* /, nasal: /*m n* /), frication (stop: / *p b t d k g m n* /, fricative: / *s z f v* /), and duration (short: / *p b t d k g m n f v* / and long: / *s z* /).

4. RESULTS

Improvement was found in response time, recognition score, and reception of consonantal features for all the processing schemes. A plot of response times obtained by the 5 subjects for unprocessed (Su) and the three processing conditions (SpA, SpB, and SpC) are shown in Fig 5. The decrease in response time indicates reduction in load on perception.

The subjects BS (38) and SK (41) had highest response time for unprocessed speech among all subjects. Improvements for these subjects were statistically highly significant. It is observed that elderly subjects do not have much reduction in response time after processing.

Addition of variable gain filter, to compensate for frequency dependant shifts in hearing threshold (SpC), has further reduced the response time of the subject with asymmetrical loss

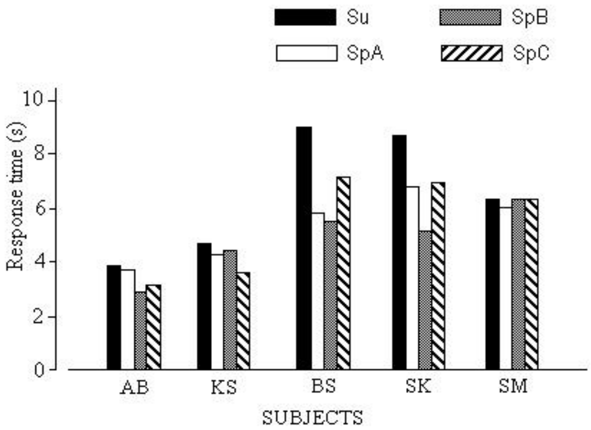

Fig. 5. Response time for unprocessed speech (Su) and processed speech (SpA, SpB, SpC).

Table 1 shows the recognition scores for unprocessed and processed speech for schemes SpA, SpB and SpC along with relative improvement (processed vs. unprocessed), significance level *p* from one-tailed t-tests for individual subjects, and from one-tailed paired t-test. The plot of the percentage recognition scores is shown in Fig. 6.

The relative improvement (%) for SpA ranged from 1.3 to 27.4 with an average of 12.4. For SpB and SpC the relative improvement (%) varied from 8.7 to 39.2 with an average of 19.3 and from 12 to 47.7 with an average of 25.2 respectively. For each of the five subjects, the improvement for SpC was higher than that for SpB, which was higher than that for SpA.

Subject BS with symmetrical high frequency loss showed lowest recognition score for unprocessed speech and maximum relative improvement for all processing schemes. This subject also had reduction in response time with processing. For subject SM who has relatively flat loss (particularly below 2 kHz), the relative improvement of SpC with respect to SpB was maximum (9.3). The relative improvement of SpC over SpB was minimum for subject AB who has low frequency hearing loss.

The recognition score for individual subjects were subjected to one-tailed t-test to obtain statistical significance due to processing. The improvement for processing scheme SpC was statistically highly significant for all the subjects. For other two schemes (SpA and SpB) improvements were statistically significant for all subjects except for subject KS.

Table 1. Recognition scores (%) for unprocessed (Su) and processed speech (SpA, SpB, and SpC). RS = % recognition score, S: subject, s.d. = standard deviation, R.I. =relative improvement in % with respect to unprocessed. p: significance level for one-tailed t-test (processed vs. unprocessed)

S (age)		Su	SpA	SpB	SpC
AB M(61)	RS	88.7	97.0	98.3	99.3
	s.d	4.2	1.4	0.0	0.9
	R.I		9.4	10.9	12.0
	p		0.001	0.004	0.0003
KS F(32)	RS	76.0	77.0	82.7	87.36
	s.d	4.3	3.2	2.8	2.5
	R.I		1.3	8.7	14.9
	p		0.35	0.01	0.0005
BS F(38)	RS	51.0	65.0	71.0	75.3
	s.d	4.3	2.0	3.8	3.8
	R.I		27.4	39.2	47.7
	p		0.0001	0.0000	0.0000
SK M(41)	RS	75.7	82.3	87.4	89.0
	s.d	0.9	3.5	2.5	2.2
	R.I		8.9	15.4	17.6
	p		0.002	0.0000	0.0000
SM M(59)	RS	67.7	77.8	82.7	90.3
	s.d	3.6	4.7	1.9	1.4
	R.I		14.9	22.2	33.5
	p		0.003	0.0000	0.0000
Avg. RS		71.8	79.8	84.4	88.3
Avg. R.I.			12.4	19.3	25.2
p(paired)			0.01	0.0028	0.015

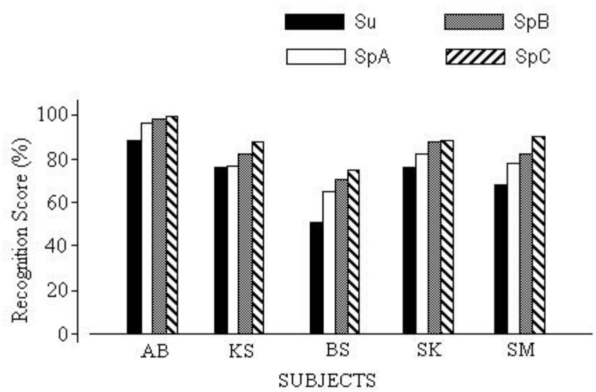

Fig. 6. Percentage recognition score for unprocessed speech (Su) and processed speech (SpA, SpB, SpC).

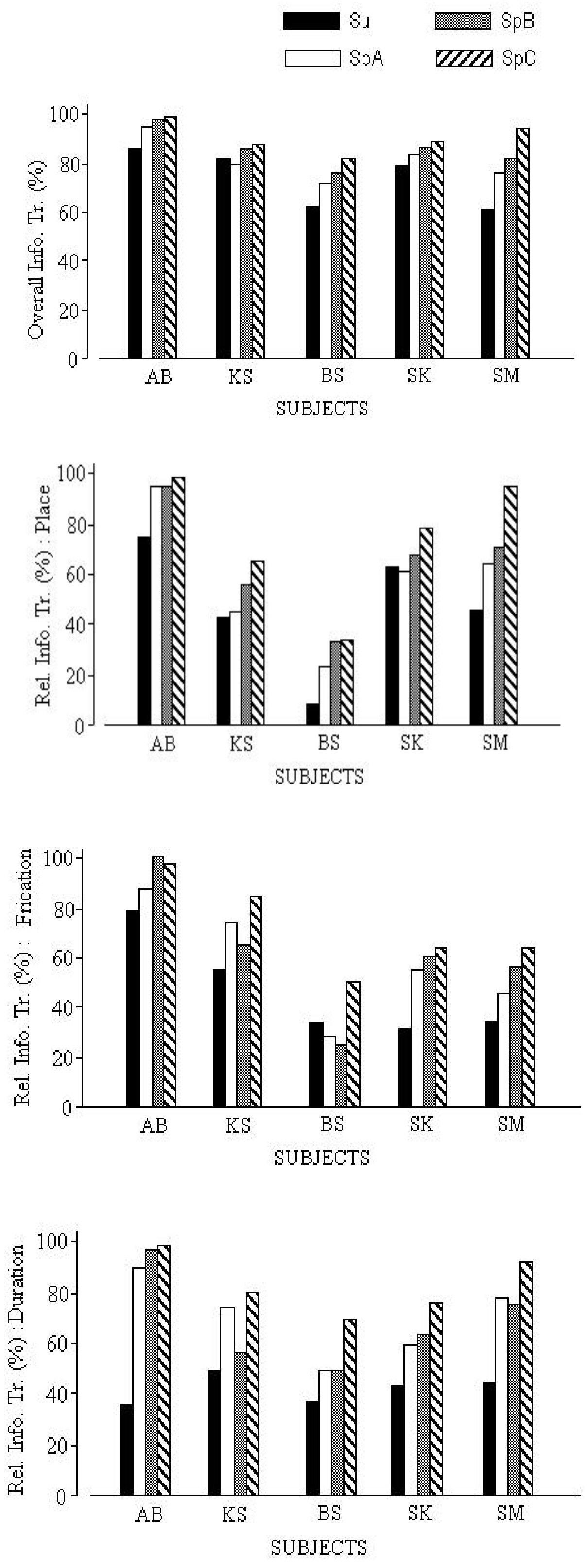

Fig. 7. Percentage relative information transmitted, overall and for features of place, frication, and duration unprocessed speech (Su) and processed speech (SpA, SpB, SpC).

Fig. 7 gives the plot of relative information transmitted for consonantal features: overall, place, duration, and frication. The overall relative information transmitted varied from 61 to 85 for unprocessed speech. The same for processed speech with the schemes SpA, SpB, and SpC are between 71 to 95, 75 to 98, 81 to 99 respectively. The relative improvements varied from −2.5 to 23.0, 4.9 to 32.8, and 7.4 to 54.1 for processed speech with SpA, SpB, and SpC respectively. For place feature, the relative information transmitted for unprocessed speech varied from 8 to 74. The subject BS who has severe high frequency loss, had more relative improvement (SpA : 187, SpB : 312, SpC : 325) compared to other subjects. The relative improvement for SpC varied between 22 and 325. Subject SM had high relative improvement for SpC with respect to SpB (36). For duration feature, the relative information transmitted for unprocessed speech varied from 35 to 48. Subject AB with low frequency loss showed more relative improvement increasing from SpA to SpC. Subjects with severe high frequency loss BS and KS had more relative improvement for SpC with respect to SpB. The relative information transmitted for frication with unprocessed speech varied from 31 to 79. The relative improvement with processing increased for subject SK and SM from SpA to SpC. The subjects BS and KS with severe high frequency loss showed higher relative improvement for SpC with respect to SpB.

5. CONCLUSIONS

All the three processing schemes evaluated in this investigation provided better speech intelligibility in listening tests with persons with mild to severe bilateral sensorineural hearing impairment. Processing with perceptually balanced comb filters provided better results compared to processing with comb filters with sharp transitions. The highest improvement was for the scheme with filters with adjustable gain to partially compensate for the frequency dependent shift in hearing threshold in cascade with perceptually balanced comb filters.

The scheme of compensating for frequency dependent shifts in hearing threshold cascaded with perceptually balanced comb filters provided better intelligibility for all subjects but with varying degrees of improvement. Subject with almost flat loss (particularly below 2 kHz) showed more improvement. He showed more relative improvement in recognition score and relative information transmission of features (overall and place). The subjects with high frequency loss are benefited in the information transmission of features of frication and duration. Subject with high loss at low frequency is least benefited by the combined scheme.

The implementation and evaluation of perceptually balanced comb filters have shown that the scheme can be used in binaural hearing aids for persons with bilateral sensorineural hearing impairment. In order to simultaneously reduce the effects of temporal and spectral masking associated with sensorineural loss, a scheme of dichotic presentation with time varying comb filters has been developed [11]. It provides periodic relaxation to all the sensory cells in the inner ear. Each time varying comb filter consists of a number of comb filters, which are used in a cyclic manner to process the speech for right and left ear. The pass bands of these comb filters are shifted along the frequency axis such that a sweeping of bands take place. This is being evaluated with hearing impaired subjects to find the optimum number of comb filters in each of the time varying comb filter and for the optimum cycle duration for sweeping.

REFERENCES

[1] B.C.J. Moore, “Psychoacoustics of cochlear hearing impairment and design of hearing aids,” **Proc. 16th Int. Congress Acoust. (ICA)**, Seattle, Washington, USA, 1998, pp. 2105-2108.

[2] T. Baer and B.C.J. Moore, “Effects of spectral smearing on the intelligibility of sentences in noise,” **J. Acoust. Soc. Am.**, Vol. 94, 1993, pp. 1229-1241.

[3] P.E. Lyregaard, “Frequency selectivity and speech intelligibility of sentences in noise,” **Scand . Audiol.**, Suppl. 15, 1982, pp.113-122.

[4] 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, 1993, pp. 75-81.

[5] D.S. Chaudhari and P.C. Pandey, “Dichotic presentation of speech signal with critical band filtering for improving speech perception,” **Proc. ICASSP’98**, Seattle, Washington, USA, 1998, paper AE 3.1.

[6] D.S. Chaudhari and P.C. Pandey, “Binaural audio in multimedia systems to improve auditory perception for the hearing impaired,” **Proc. 5th ISSPA ' 99** Brisbane, Australia. 1999.

[7] E. Zwicker, “Subdivision of the audible range into critical bands (Frequenzgruppen),” **J. Acoust. Soc. Am.**., Vol. 33, 1961, pp. 248.

[8] L.R. Rabiner, B. Gold, and C.A. McGonegal, “An approach to the approximation problem for nonrecursive digital filters,” **IEEE Trans. Audio and Electroacoustics**, Vol. AU-16, 1970, pp. 83-105.

[9] A.N. Cheeran, P.C. Pandey and D. S. Jangamashetti, “Design of comb filters based on auditory filter bandwidths for binaural dichotic presentation for persons with sensorineural hearing impairment,” **Proc. 14th Int. Conf. on Digital Signal Processing (DSP2002)**, Santorini, Greece, 2002, pp. 1145-1148.

[10] Miller, G.A., and Nicely, P. E. “An analysis of perceptual confusions among English consonants,” **J. Acoust. Soc. Am.**, Vol. 27, 1955, pp. 338-352.

[11] A.N. Cheeran, P.C. Pandey, and D.S. Jangamashetti, “Optimal sweep cycle for time-varying comb filters for binaural dichotic presentation to improve speech perception in sensorineural hearing impairment,” **J. Acoust. Soc. Am.**, Vol. 111(5), 2001, pp. 2426.