Proc. 14th Int. Nat. Conf. on Digitial Signal Processing (DSP 2002), Santorni, Greece, July 1-3, 2002

DESIGN OF COMB FILTERS BASED ON AUDITORY FILTER BANDWIDTHS FOR BINAURAL DICHOTIC PRESENTATION FOR PERSONS WITH SENSORINEURAL HEARING IMPAIRMENT

Alice N. Cheeran[1], Prem C. Pandey[2], and Dakshayani S. Jangamashetti[3]

[1]BME Group, IIT Bombay, Powai Mumbai-400 076, India
alicenc@ee.iitb.ac.in
[2,3]EE Dept., IIT Bombay, Powai Mumbai-400 076, India
[2]pcpandey@ee.iitb.ac.in
[3]shetty@ee.iitb.ac.in

Abstract: Splitting of speech into two signals by comb filters having complementary magnitude response and presenting these signals to the two ears has helped in improving the perception for persons with bilateral sensorineural hearing loss. Investigation is carried out to improve the comb filters based on auditory critical bands, with the objective of minimizing the perceived spectral distortion. Listening tests were conducted to find the difference in intensity with monaural and binaural presentations for equal loudness perception. Based on the results obtained, 256-coefficient linear phase FIR comb filters were designed using frequency sampling technique, to obtain magnitude response with pass band ripple of 1 dB, stop band attenuation of 30 dB, and crossovers adjusted to lie between –4 dB and –6 dB with respect to pass band response. Listening tests involving closed set identification of 12 vowel-consonant-vowel syllables were conducted, to compare the performance of the new comb filter with the filter with sharp transitions. The new comb filters resulted in a higher improvement of recognition scores and relative information transmission.

1. INTRODUCTION

The characteristics of sensorineural hearing loss, which occurs due to the damage of hair cells in the cochlea or degeneration of auditory nerve fibers, or both, are frequency dependant increase in hearing threshold, reduction in dynamic range, reduced frequency resolution associated with increase in spectral masking and decrease in temporal resolution associated with increase in temporal masking [1],[2]. Sensorineural hearing loss exhibits widening of auditory filters due to increased spectral masking, resulting in severe smearing of spectral envelope. Normally vowels are characterized by formant frequency cues, which are widely separated from each other, hence their perception is not much affected. However, perception of consonants is severely degraded, since it requires discrimination of sub-phonemic segments like formant transitions and noise bursts.

Splitting of speech into two signals, such that the frequency components that are likely to get masked are separated and presented to different ears has helped in reducing the effect of spectral masking. The information from the signals presented to the two ears gets integrated at higher levels in the auditory process. In an early reported effort for splitting of speech spectrally for binaural dichotic presentation, an analog delay based design was used to obtain two complementary comb filter magnitude responses with constant bandwidth pass and stop bands [3]. The improvement of dichotic over diotic was insignificant. Later, Lunner *et al.* [4] reported an overall improvement of 2 dB in speech-to-noise ratio for dichotic with respect to diotic, with the use of comb filters with eight channel filter bank (constant bandwidth of 700 Hz) realized using complementary interpolated linear phase FIR filters. Focus in the design was on efficiency, and not on separation of bands and undistorted perception for spectral components at band crossovers.

Chaudhari and Pandey [5],[6],[7] investigated the use of 18-band comb filters with complementary magnitude responses based on auditory critical bands described by Zwicker [8]. 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 odd and even bands which form a pair of comb filters, split the speech into two such that the spectral components that are likely to get masked are presented to different ears. However, with filters with finite crossovers in magnitude response, the spectral components lying in the transition region are presented to different ears. An imbalance in loudness was perceived at the crossovers between adjacent bands. In the present investigation comb filters are designed with three considerations; adjustment of magnitude response at the transition crossovers to minimize the changes in intensity perception, reduction in pass band ripple, and increase in stop band attenuation.

2. COMB FILTERS WITH SHARP TRANSITION BETWEEN PASS AND STOP BANDS

In the scheme investigated by Chaudhari and Pandey [5],[6],[7] for real time processing, comb filters were designed as 128-coefficient linear phase FIR filters using frequency sampling technique. To obtain proper separation while splitting, the filters were designed with sharp transition between bands. In the present investigation, 128-coefficient linear phase FIR comb filters were realized with sharp transitions for each of the band pass filters (i.e. transition width of one sample). With a sampling rate of 10 k samples/s, the transition width (Δf) corresponds to 78 Hz. The signal processing has been carried out off-line, with filters implemented using floating-point arithmetic, so that the effects of coefficient quantization and calculation errors are not significant. The filters have maximum ripple of 4 dB and minimum stop band attenuation of 10 dB.

As a first step towards designing better filters, comb filters were designed with 256 coefficients, taking into consideration the present developments in DSP processors. The magnitude response of the pair of comb filter is shown in Fig. 1. The filters have sharp transition between bands, pass band ripple of 3 dB, and stop band attenuation of 11 dB. Doubling of filter coefficients halved the transition width (Δf = 39 Hz), but it did not result in significant improvements in pass band ripple and stop band attenuation.

Relatively small stop band attenuation may lead to inadequate separation between bands. The just noticeable difference (JND) for overall intensity in case of synthetic vowel is reported to be 1.5 dB, while for first and second formants it is 1.5 dB and 3 dB respectively [9]. The JND remains very near to 1 dB for a wide range of levels for wide band of noise [10]. Perceptual distortion was noticed when a slowly sweeping sine tone with frequency variation limited to lie in one pass band with ripple of 4 dB, was processed and perceived. Hence it was decided to limit the maximum pass band ripple to 1 dB. Another serious problem with these filters is that the crossovers between adjacent bands are at different levels with respect to the pass band response, and this may result in perceptual distortions due to decrease or increase in the perceived intensity of spectral components lying in the transition region. When a sine wave with frequency sweeping slowly between 0 and 5 kHz was processed with the pair of comb filters and presented binaurally, a change in intensity was clearly perceived at the transition between bands.

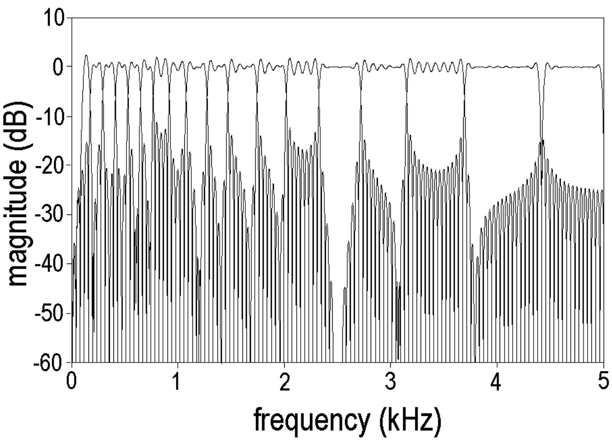

Fig. 1. Magnitude response of the pair of comb filters (256-coefficients) designed for sharp transitions.

3. COMB FILTERS WITH IMPROVED PASS BAND RIPPLE AND STOP BAND ATTENUATION

Rabiner *et al.* [11] reported a method, involving frequency sampling technique for FIR filter design, for increasing the side lobe attenuation of prototype filters by trading sharp transition between bands. The magnitude of the transition samples was considered unconstrained and was adjusted to make the required changes in the response. They used linear programming technique to find the optimal magnitude for the transition samples. For the design of the critical band based comb filters, use of automated design techniques based on optimization criteria was not found suitable, and hence the filters were designed by using iterative procedure.

For adjusting the magnitude response of the comb filters, the samples in the pass band are considered as constrained samples taking a magnitude of 1. Samples lying close to the edge of the pass band are taken as unconstrained (transition samples), and remaining samples of the stop band are constrained with value 0. The magnitude of the unconstrained samples lying in the transition region was varied to modify the magnitude response and the interpolated response was observed. This iterative process was continued for each transition sample, until optimization is obtained for parameters under consideration. The number of transition samples (0, 1, 2) was dependent on the available stop bandwidth, increasing from low to high frequencies.

The filter with 256 coefficients provided stop band attenuation of 38 dB with pass band ripple constrained to 1 dB. With sampling rate of 10 k samples/s, the filters have a transition band of 78 Hz at lower frequencies and 117 Hz at higher frequencies. Listening tests did not show any change in intensity perception due to pass band ripple, when the frequency of a sinusoidal tone was swept over the pass band with maximum ripple. When a sine wave with its frequency slowly swept over 0 to 5 kHz was processed with these comb filters and was presented binaurally, a change in intensity was perceived, for frequencies in the transition region. Thus, even though this comb filter provided relatively flat pass bands and adequate band separation, it was necessary to modify the magnitude response at the transitions to balance the perception at all frequencies.

4. COMB FILTERS ADJUSTED WITH INTER-BAND CROSSOVERS

In ideal splitting, any spectral components would be presented to one ear. However, with the filters with finite crossover in magnitude response, the components lying in the pass band are presented to one ear, whereas those lying in the transition region are presented to both ears. With the same intensity, binaurally presented components will be louder than monaurally presented components, however the loudness is generally less than double [12]. Loudness is related to intensity in a very complex way, since intensity is not the sole determinant of loudness. If the magnitude response is not properly adjusted at the transitions, the components lying in the overlapped region will be perceived with different loudness and will reduce the speech quality. Loudness evaluation test was conducted to determine the difference in intensity for the same perception in monaural and binaural presentations. Comb filters were designed with different magnitudes at crossovers between adjacent bands. These tests and results are presented in the following subsections.

4.1. Perceptual Balance of Monaural and Binaural Intensity levels.

Listening tests were conducted to determine the difference in the intensity in monaural and binaural presentations, such that they evoke the same perceived loudness level. Stimuli used were pure tones in four different frequencies (0.5, 1, 2, 4 kHz), sustained vowel /*a*/, and broad-band noise. Five normal hearing subjects participated in the tests. The stimuli of 1 s duration were presented through headphones monaurally and binaurally one after the other, with an inter-stimulus interval of 1 s. Monaural intensity was kept constant at 85 dB and binaural intensity was varied from 84 dB to 70 dB in steps of 1dB, to establish the monaural versus binaural intensity balance, following a 2-step matching procedure.

The results from these tests for different stimuli on the five subjects are given in Table 1. The perceived levels match when the binaural level was 4–9 dB lower than monaural level. An interesting observation here is that for tones, there is significant inter-subject variation. Also for each subject there is a significant inter-frequency variation. However, for vowel and broadband noise, the level difference for balance is about 9 dB for all the subjects.

Table 1. The intensity difference (in dB) between monaural and binaural presentations for same loudness perception.

Subject	STIMULI					
	Pure Tone (kHz)				Vowel /*a*/	Noise
	0.25	1	2	4		
AC	8	7	4	7	8	9
DJ	8	6	5	7	8	7
VK	12	7	9	7	10	9
AJ	5	5	6	6	9	9
MD	12	7	9	9	9	9

4.2. Comb Filters with Crossovers Adjusted for Perceptual Balance

Based on the results obtained from the listening tests for determining intensity level difference for perceptual balance between monaural and binaural presentation, pairs of comb filters were designed with different crossovers between adjacent bands and listening tests were conducted with swept sine waves. Seven pairs of comb filters were designed with different crossover points varying between –3 dB and –9 dB at the crossover region of the first two auditory critical bands (100–200 Hz, 200–300 Hz). A sine wave with frequency swept between 100 Hz to 300 Hz over an interval of 30 s was processed with these comb filters and listening tests were conducted. For swept sine waves processed with comb filter pairs with crossover points between –4 dB and –6 dB, change in intensity perception was not noticeable as the swept sine wave moved from one ear to the other. To verify the effect in high frequency range, comb filter pairs were designed with crossover points varying between –3 dB and –9 dB at the overlapping of the pass bands of the 15th and 16th auditory filters. Listening tests were conducted using sine wave with frequency swept between 3 kHz and 3.5 kHz, so as to cover the overlapping region of these bands. Similar results were obtained as those obtained for lower frequency bands.

Next the transition regions of the pair of comb filters were modified to obtain crossovers to lie between –4 dB and –6 dB with respect to the pass band response at every adjacent band. The pass band ripple was constrained to 1 dB and stop band attenuation was maximized. The comb filters had 256 coefficients. The transition width was varied from 78 Hz to 117 Hz. The adjustment of magnitude response was done, by adjusting the magnitude of the different transition samples iteratively. Fig. 2 shows the magnitude response of the pair of comb filters.

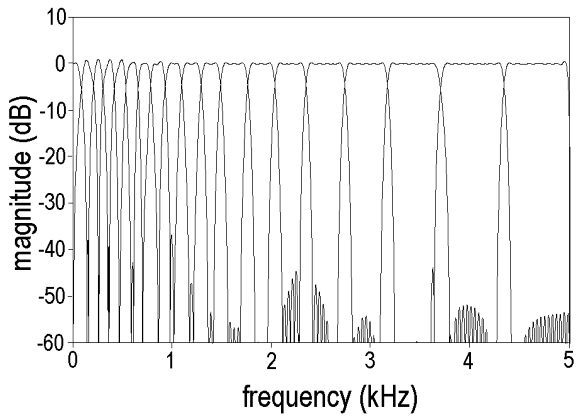

Fig. 2. Magnitude response of the pair of comb filters (256-coefficients) designed for low perceived spectral distortion.

5. EVALUATION AND RESULTS

Listening tests were carried out for comparing the two types of comb filter pairs namely (i) with sharp transition between bands (denoted as A) and (ii) with minimum perceived spectral distortion. (denoted as B). The filters were implemented for off-line processing and tests were conducted on three normal hearing subjects with hearing loss simulated by adding broad-band noise with constant short-time signal-to-noise ratio (SNR). SNRs used were ∞, 6, 3, 0, –3, –6, –9, –12, and –15 dB. In the listening tests the subjects were asked to identify a closed set of 12 English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in a vowel-consonant-vowel context with vowel /*a*/ as in "f*a*ther", using a computerized test administration system [6],[7], and responses were recorded in the form of stimulus-response confusion matrices for each of the test conditions.

Percentage recognition score and relative information transmitted for different features for unprocessed and processed speech were analyzed. Improvement was found in recognition scores and transmission of features in all the test conditions for processing with both the pairs of comb filters (A and B), with higher improvement for comb filter B. Table 2(a) shows the recognition score for unprocessed and processed speech with comb filters; namely with sharp transitions (A) and with minimum spectral distortion (B), for different levels of hearing loss simulation. The average relative improvements in recognition score at –15 dB were 15.5 and 21.8% for filter sets A and B respectively.

Stimulus-response confusion matrices were subjected to information transmission analysis for overall information and for speech features of voicing, place, frication, manner, and duration. The relative information transmission scores are given in Table 2(b) for overall information. In their improvements, the maximum contribution was due to place feature.

6. CONCLUSIONS

It has been earlier reported that the use of comb filters, based on auditory filter bandwidths, for binaural dichotic presentation can help in reducing the effect of spectral masking related to sensorineural hearing impairment. Filters with sharp inter-band transitions (38 Hz) had large pass band ripple (4 dB), and relatively low stop band attenuation (10 dB). New filters have been designed to reduce the perceived spectral distortion. Parameters for the inter-band crossovers were selected on the basis of listening tests. In designing the new filters, transition bandwidth has been increased in order to reduce the pass band ripple, to increase the stop band attenuation, and to adjust the inter-band crossovers. The final filters have transition bandwidths of 78 Hz to 117 Hz, pass band ripple of 1 dB, stop band attenuation of 30 dB, and inter-band crossovers between –4 dB and –6 dB. These filters do not result in intensity variations for swept sine waves indicating that the problem of perceived spectral distortion has been solved. Listening tests indicated that under simulated hearing loss the new filter gave better speech recognition scores and relative information transmission than the earlier filter. Therefore the comb filter design can be used for binaural aids for persons with bilateral sensorineural hearing impairment, for reducing the effect of spectral masking.

Table 2. Recognition scores and overall relative information transmitted for Unprocessed Speech (Su) and speech processed with comb filters with (i) sharp transition between bands (SpA), and (ii) minimum spectral distortion (SpB).

(a) Recognition scores

S	∞ SNR			-3 dB SNR			-6 dB SNR			-9 dB SNR			-12 dB SNR			-15 dB SNR		
	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB
S1	100.0	98.7	100.0	90.3	99.0	99.3	92.7	98.3	99.0	82.3	99.0	98.7	73.3	87.3	90.7	75.7	94.3	94.0
S2	100.0	100.0	100.0	91.0	98.0	100.0	93.0	98.3	99.7	87.7	87.3	96.7	83.3	87.7	89.0	71.7	73.3	77.7
S3	100.0	99.7	100.0	86.3	99.0	99.3	76.7	89.0	96.0	72.0	87.3	89.7	64.0	83.0	86.7	57.0	68.0	75.7
Avg.	100.0	99.5	100.0	89.2	98.7	99.5	87.5	95.2	98.2	80.7	91.2	95.0	73.5	86.0	88.8	68.1	78.5	82.5

(b) Overall Relative information transmitted

S	∞ SNR			-3 dB SNR			-6 dB SNR			-9 dB SNR			-12 dB SNR			-15 dB SNR		
	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB	Su	SpA	SpB
S1	100.0	98.0	100.0	90.0	98.0	99.0	93.0	98.0	99.0	84.0	98.0	98.0	76.0	85.0	89.0	76.0	92.0	91.0
S2	100.0	100.0	100.0	93.0	97.0	100.0	93.0	98.0	99.0	88.0	90.0	96.0	85.0	87.0	88.0	73.0	70.0	78.0
S3	100.0	99.0	100.0	88.0	98.0	99.0	82.0	89.0	95.0	77.0	88.0	88.0	73.0	82.0	85.0	66.0	73.0	75.0
Avg.	100.0	99.0	100.0	90.3	97.7	99.3	89.3	95.0	97.7	83.0	92.0	94.0	78.0	84.7	87.3	71.7	78.3	81.3

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 using critical filter band for bilateral sensorineural hearing impairment," *Proc. 16th Int. Congress Acoust. (ICA)*, Seattle, Washington, USA, 1998, pp. 213-214.

[6] 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.

[7] 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.

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

[9] J.L.Flanagan, *Speech Analysis, Synthesis and Perception*, Springer-Verlag, New York, 1972.

[10] W.A. Yost, *Fundamentals of Hearing*, Academic Press Inc., San Diego, California, USA, 1994.

[11] 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.

[12] B. Scharf, "Dichotic summation of loudness," *J. Acoust. Soc. Am.*, Vol. 45, 1969, pp. 1193-1205.