

OPTIMIZING THE SWEEP CYCLE OF TIME-VARYING COMB FILTERS FOR BINAURAL DICHOTIC PRESENTATION IN SENSORINEURAL HEARING IMPAIRMENT

Alice N. Cheeran¹ and Prem C. Pandey²

¹BME Group, IIT Bombay, Powai Mumbai-400 076, India
alicenc@ee.iitb.ac.in

²EE Dept., IIT Bombay, Powai Mumbai-400 076, India
pcpandey@ee.iitb.ac.in

Abstract: Splitting of speech into two using time-varying comb filters has helped in reducing the effect of increased temporal and spectral masking simultaneously. The time varying comb filters contained pre-calculated sets of coefficients, which were selected in steps, such that a cyclic sweeping of magnitude responses occur. Presently investigation has been carried out to find the best sweep cycle. The two sets of coefficients selected at any instant of time formed a pair of complementary comb filters with 18 auditory critical bands which provided spectral separation of components that are likely to get masked. Sweeping of bands provides temporal separation and reduces the effect of temporal masking. As the number of magnitude responses swept increases, the sweeping becomes smoother. The sweep cycle duration determines the time for which the sensory cells of the basilar membrane remain stimulated. Perception evaluation tests were conducted with slowly sweeping sine wave and running speech from a male and a female speaker for sweep cycles of 10, 20, 40, 50, 60, 80, and 100 ms. Best results were obtained for a sweep cycle of 50 ms.

1. INTRODUCTION

Sensorineural hearing loss which occurs due to the malfunctioning of cochlea, auditory nerves or both, is characterized by elevated hearing threshold, loudness recruitment, reduced temporal and frequency resolution, and increased temporal and spectral masking [1],[2]. Increased spectral masking causes smearing of spectral peaks and valleys due to broader auditory filters. Reduction in spectral contrasts leads to difficulty in the discrimination of consonantal place feature. Forward and backward masking of weak segments by strong ones causes reduction in the discrimination of sub-phonemic segments like noise bursts, voice-on-set time, and formant transitions. Masking takes place at the peripheral level of the auditory system. Information received at both the ears gets integrated at higher levels of auditory perception. Hence splitting of speech into two complementary signals for binaural dichotic presentation can be used to relax the hair cells of the basilar membrane periodically and reduce the effect of masking.

Lunner *et al.* [3] implemented an 8-channel digital filter bank with constant bandwidth of approximately 700 Hz. The alternate bands were combined to form two complementary comb filters for spectral splitting in binaural dichotic presentation. An improvement of 2 dB in speech-to-noise ratio is reported for dichotic over diotic. Further, the comb filters were alternated after every 10 ms for obtaining combined spectral and temporal splitting. They reported no improvement over spectral splitting and poor sound quality due to switching of bands.

Recently, Chaudhari and Pandey [4],[5] investigated a scheme of spectral splitting using a pair of comb filters with complementary magnitude responses. Each of these comb filters had nine pass bands based on

auditory critical bands described by Zwicker [6]. The bandwidths were constant at 100 Hz for center frequencies below 500 Hz and were 15-17% of the center frequencies in the range of 1-5 kHz. The pair of comb filters formed by the odd and even bands, split the speech signal into two such that the spectral components that are likely to get masked are presented to the two different ears. The comb filters were linear phase FIR filters with 128 coefficients designed with sharp transitions between bands, using frequency sampling techniques. Real time implementation of the scheme was done on two TI/TMS320C50 DSP processors. Listening test evaluation on bilaterally hearing impaired persons showed improvement in the perception of consonantal features particularly the place feature, which lead to the conclusion that spectral splitting helped in reducing the effect of increased spectral masking for sensorineural hearing impaired. Further, a scheme of temporal splitting was experimented, in which inter-aural switching of speech signal was done using trapezoidal fading function [7]. The evaluation of the scheme was conducted on normal subjects with simulated hearing loss. The scheme helped in improving the consonantal duration feature, showing reduction in the effect of increased temporal masking.

In spectral splitting sensory cells corresponding to alternate bands are always relaxed, while in temporal splitting all the sensory cells of the two ears are alternately relaxed. Subsequently a scheme of combined splitting was developed, so that all the sensory cells of the basilar membrane are periodically relaxed from stimulation for some time. The implementation was done using a pair of time varying comb filters with pre-calculated sets of coefficients, which were selected in steps such that a cyclic sweeping of magnitude responses occur. Experimental evaluation on normal

subjects with simulated hearing loss was conducted for a constant sweep cycle of 20 ms [8]. The scheme provided improvement in response times and recognition scores. Information transmission analysis showed increase in the perception of all features especially the place and duration features. The scheme was successful in simultaneously reducing the effects of increased temporal and spectral masking. Presently investigation is carried out to find the optimal time duration of a single cycle of sweeping in the time varying comb filters.

2. IMPLEMENTATION

In our earlier work [8], investigation was carried out with a constant sweep cycle of 20 ms. With constant sweep cycle the time duration of stimulation of the sensory cells remains constant. For the perception of consonants, temporal cues related to rapid formant transitions, voice-on-set time etc. are important. The duration of formant transitions in stop consonants is 15–30 ms [9]. The present investigation is involved in finding the optimal value for a sweep cycle in the time varying comb filters. The schematic representation of the scheme of combined spectral and temporal splitting using time varying comb filters is shown in Fig. 1(a) The digitized input signal $s(n)$, is passed through a pair of time varying comb filters to produce two digitized outputs, $s_1(n)$ and $s_2(n)$ to be fed to the two ears (binaural dichotic presentation). Each of these time varying comb filters contained m comb filters (shiftings) which have magnitude responses such that the pass bands and stop bands of each of these comb filter will be slightly shifted along the frequency axis with respect to the pass and stop bands of the previous comb filter, as shown in Fig. 1(b). If the comb filters in the time varying comb filter for the left ear are numbered in the order of sweeping as $[1], [2], \dots, [m/2], [m/2 + 1], \dots, [m]$, then the same for the right will be $[m/2 + 1], [m/2 + 2], \dots, [m], [1], [2], \dots, [m/2]$. At any instant of time two complementary magnitude responses corresponding to odd and even bands processes the speech, which provides separation of spectral components that are likely to get masked. $[1]$ and $[m/2 + 1]$, $[2]$ and $[m/2 + 2]$, $\dots, [m/2]$ and $[m]$ forms these complementary pairs.

Each of these comb filters has nine pass bands corresponding to auditory critical bands described by Zwicker [7] and were designed as 256-coefficient linear phase FIR filters with minimum spectral distortion. To minimize the perceived spectral distortion at the transition between any two adjacent bands of a pair used simultaneously, the crossovers were adjusted to lie between -4 dB and -6 dB with respect to pass band gain [10]. This was obtained by iteratively adjusting the magnitude of the transition samples [11]. The transition bandwidth of the comb filters was 78 Hz at low frequencies and 117 Hz at higher frequencies with a sampling rate of 10 k Samples/s. The stop band attenuation was greater than 30 dB with pass band ripple constrained to 1 dB. A swept sine tone processed with each of these complementary pair of comb filters that need to be used together, produce negligible perceived variation in loudness.

The number of magnitude responses (shiftings) in a time varying comb filter was chosen as 2, 4, 8, and 16. An idealized representation of magnitude response of the pair of time varying comb filters with 16 shiftings in a sweep cycle of 50 ms is shown in Fig. 2 As the number of shiftings increases, more magnitude responses take part in the sweeping, hence the sweeping becomes smoother and the stimulation of the sensory cells in the basilar membrane also will move smoothly from one band to another. Sweeping of bands provides temporal splitting and helps in reducing the effect of increased forward and backward masking.

3. EVALUATION AND RESULTS

For initial investigation, a sine wave linearly swept from 100 Hz to 5 kHz and back with a sweep period of 1s was used as test signal. The processing was done for filters with 2, 4, 8, and 16 shiftings. Filters with sweep cycle of 10, 20, 50, 80, and 100 ms were used. Then there were 20 (4 x 5) types of filters. Test signals processed by each of these filters were presented in a randomized order for comparison with the unprocessed test material. The processed signals with 8 and 16 shiftings, for 50 ms sweep cycle were found to be almost similar to the unprocessed swept sine tone.

For the second round of evaluation, 6 types of filters were used: filters with sweep cycle time of 40, 50, 60 ms and 8 and 16 shiftings. Listening tests were conducted on three normal subjects with the test signal processed by these six comb filter pairs, for comparison with the unprocessed test signal. Maximum similarity with the original was obtained for 16 shiftings with a sweep cycle time period of 50 ms. The difference in speech clarity for 8 and 16 shiftings was negligible.

Further, running speech from a male and a female speaker was recorded using the microphone connected to the PC sound card. Perception test was conducted with the two speech stimuli processed by the time varying comb filters with different sweep cycle time periods of 10, 20, 40, 50, 60, 80, and 100 ms for 2, 4, 8, and 16 shiftings. Bilateral sensorineural hearing impairment was simulated by adding broad-band noise with constant short-time signal-to-noise ratio (SNR). For obtaining different levels of hearing impairment SNRs of ∞ , 6, 3, 0, -3, -6, and -9 were used. Since the number of processed test materials was too many, the first author performed an informal perception evaluation and about 8 processed stimuli were short-listed for each of the SNRs with male and female running speech material. Perception evaluation of these processed signals were conducted on two normal subjects other than the first author as in the previous case. Improvement with processing was clearly perceived for all the processed test materials, with an increasing trend from higher SNR to lower SNR. Speech intelligibility of processed stimuli with 8 and 16 shiftings was almost the same. For very low sweep cycle (10 and 20 ms) a ringing sound could be perceived along with processed speech stimuli. At sweep cycle of 100 ms an echo was heard, as the speech stimulus was perceived. Finally, the highest perceptual quality ranking was given for 16 shiftings with a sweep cycle of 50 ms.

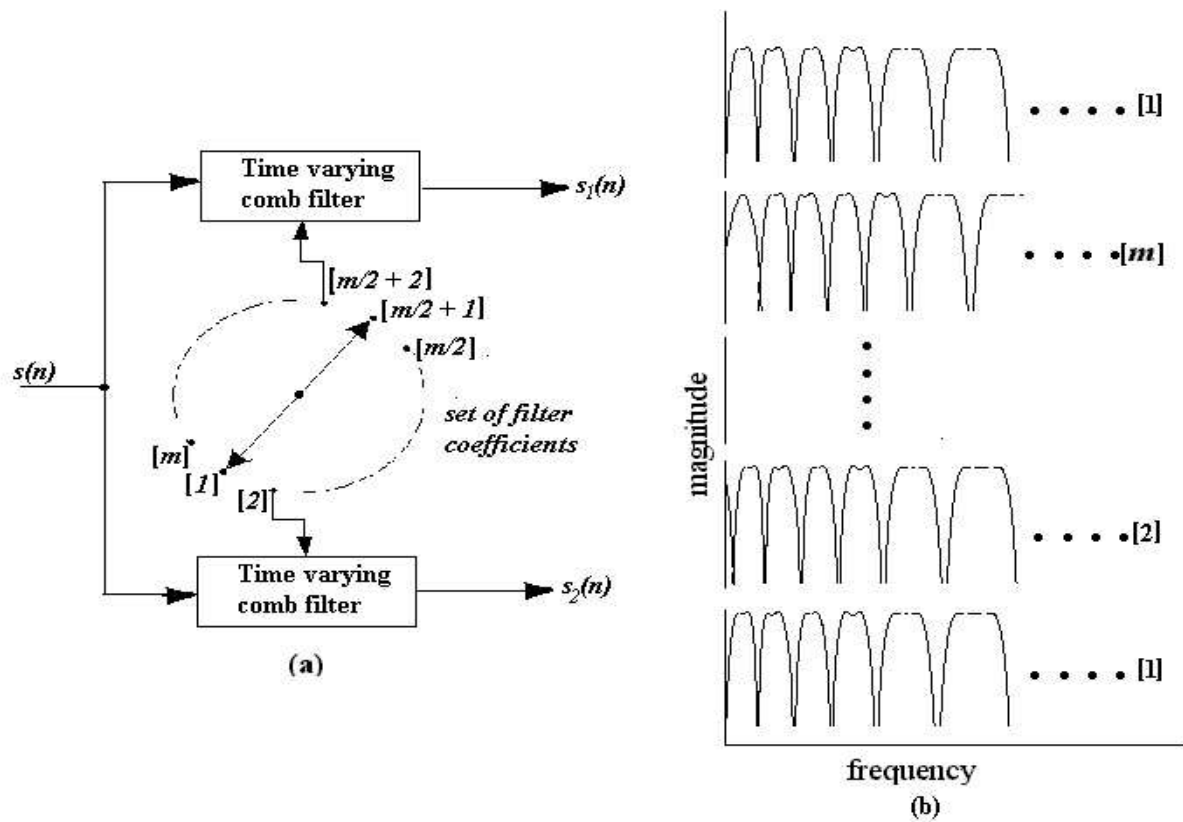


Fig. 1. (a) A schematic representation of the scheme of combined temporal and spectral splitting using time varying comb filters and (b) The representation of magnitude response of one of the time varying comb filters which includes the magnitude response of the comb filters, which are swept over one after the other cyclically.

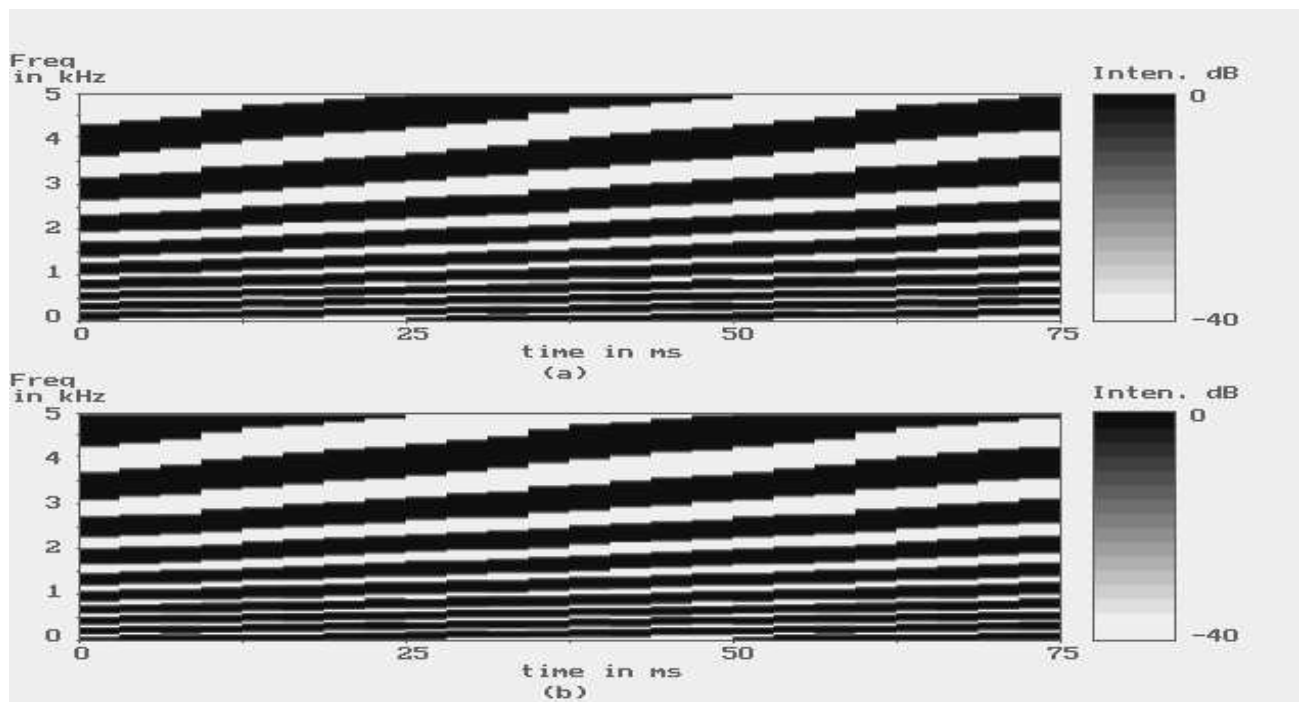


Fig. 2. An idealized representation of magnitude response of the pair of time varying comb filters with 16 shiftings and 50 ms sweep cycle (a) for the left ear and (b) for right ear.

4. CONCLUSIONS

The scheme of time varying comb filters for splitting the speech temporally and spectrally has helped in improving the perception of speech, under adverse listening conditions, by reducing the effects of increased temporal and spectral masking. As the number of shiftings increases, the sweeping of bands becomes smoother. This helps in reducing the spectral distortion associated with switching. The improvements appear to stabilize at about 8-16 shiftings. As the number of shiftings increases, the number of coefficients to be stored and retrieved increases. Hence it is concluded that 8-16 shiftings are adequate.

All the sensory cells of the basilar membrane get activated for almost the same time duration in a particular sweep cycle for all the shiftings. As the sweep cycle increases, each of the sensory cells gets stimulated and relaxed for a longer time in one sweep cycle. Certain minimum time of stimulation may be required for perceiving some of the temporal cues. Also if the stimulation time exceeds certain limit, the scheme may not help in reducing the effect of increased forward and backward masking in persons with sensorineural hearing impairment. With the present investigation, it could be concluded that a sweep cycle of 50 ms provides the maximum improvement.

Work need to be continued on testing the scheme on bilateral sensorineural hearing impaired subjects. The problems due to elevated hearing threshold and loudness recruitment also need to be considered, by implementing multi-band compression in the scheme along with time varying comb filters, which may help in reducing all the characteristics of sensorineural hearing impairment.

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] 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.
- [4] 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.
- [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] E. Zwicker, "Subdivision of the audible range into critical bands (Frequenzgruppen)," *J. Acoust. Soc. Am.*, Vol. 33, 1961, pp. 248.
- [7] D.S. Jangamashetti and P.C. Pandey, "Inter-aural switching with different fading functions for binaural dichotic presentation to reduce the effect of temporal masking in sensorineural hearing loss," *Proc. 4th World Multi conf. on Systemics, Cybernetics, and Informatics (SCI'2000)*, Orlando, Florida, USA, 2000, pp. 434-439.
- [8] P.C. Pandey, D.S. Jangamashetti, and A.N. Cheeran, "Binaural dichotic presentation to reduce the effect of increased temporal and spectral masking in sensorineural hearing impairment," *J. Acoust. Soc. Am.*, Vol. 110(5), 2001, pp. 2705.
- [9] C.W. Turner, S.J. Smith, P.L. Aldridge, and S.L. Stewart, "Formant transition duration and speech recognition in normal and hearing impaired listeners," *J. Acoust. Soc. Am.*, Vol. 101(5), 1997, pp. 2822-2825.
- [10] A.N. Cheeran, P.C. Pandey, and D.S. Jangamashetti, "Comb filters for binaural dichotic presentation to improve speech perception by persons with bilateral sensorineural hearing impairment," *J. Acoust. Soc. Am.*, Vol. 110(5), 2001, pp. 2705.
- [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.