19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

EFFECT OF BINAURAL DICHOTIC PRESENTATION WITH CRITICAL BANDWIDTH BASED COMB FILTERS ON SOURCE LOCALIZATION

PACS: 43.66.Qp

Kulkarni, Pandurangarao N.; Pandey, Prem C.
EE Dept, Indian Institute of Technology Bombay, Powai Mumbai 400 076, India
{pnkulkarni, pcpandey}@ee.iitb.ac.in

ABSTRACT

Earlier studies have shown that binaural dichotic presentation, using critical bandwidth based spectral splitting with perceptually balanced comb filters, helps in reducing the effect of spectral masking for persons with moderate bilateral sensorineural hearing impairment. It has been reported to provide an SNR advantage of 5 dB for normal hearing subjects under simulated hearing loss, and a relative improvement of 6-121 % in recognition scores for hearing impaired listeners. The objective of the present study is to investigate the effect of this type of binaural dichotic presentation on source localization. Investigations were carried out by processing the stimuli with head related transfer functions (HRTFs) for the two ears with binaural presentation using headphones, for 0 degree elevation and azimuth angles varying from -90 to +90 degrees. The test material included sound of breaking glass, VCV syllable /aba/, and noise band-pass filtered with different bandwidths. A considerable degradation in source localization was observed for narrow band stimuli. But the subjects were able to perceive the source direction for stimuli with bandwidth exceeding one-third octave, indicating that use of dichotic presentation should not impair source localization for speech and broadband environmental sounds.

INTRODUCTION

In cases of sensorineural hearing loss, the auditory filter bandwidth generally increases and frequency selectivity gets reduced due to increased masking [1] - [3]. For reducing the effect of increased spectral masking for listeners with moderate bilateral loss, several studies have examined the use of binaural dichotic presentation, with speech signal split into two parts with complimentary spectra with adjacent bands presented to different ears [4] - [9]. It has been reported that use of auditory critical bandwidth based spectral splitting technique using perceptually balanced comb filters helps in reducing the effect of spectral masking for persons with moderate bilateral sensorineural hearing loss. Using this technique, Cheeran and Pandey [7], [8] reported an SNR advantage of 5 dB for normal hearing subjects under simulated hearing loss, and a relative improvement of 6-121 % in recognition scores and 6.4-312.5 % in relative information transmission of consonantal place feature for hearing impaired listeners. As this signal processing presents the information dichotically, a study of its effect on source localization is very important.

Accurate localization of sound source is a complex perceptual process and it involves integration of multiple acoustic cues [1]. The most important cues for source localization in the horizontal plane are inter-aural time difference (ITD) and inter-aural level difference (ILD) [10], [11]. According to the “duplex” theory of binaural localization, ITD is important for localization at low frequencies, while high frequencies are localized from ILD [12]. Localization in vertical plane and discrimination of back and front primarily depends upon high frequency (> 5 kHz) spectral cues that are created by reflection and diffraction of sound by the external ear (particularly pinna) [13]. ITD can be calculated from the path length difference between the two ears and it varies from a minimum value of 0 for a sound coming from straight ahead, to a maximum value of 690 µs for a sound coming from a source located directly opposite to one ear [13]. ILD varies over 0-20 dB [14]. Best *et al*. [11] in their study on the role of high frequencies in source localization showed that preserving the information above 8 kHz helps in accurate localization of elevation. In spectral splitting for dichotic presentation, each band gets presented monaurally and hence source localization ability may get highly impaired for narrow band signals. For speech and broadband environmental sounds, listeners may be able to use ITD and ILD cues across the bands for source localization.

Van den Bogaert *et al*. [14] investigated horizontal source localization with bilateral hearing aids. The stimuli used were of 200 ms duration and included 1/3 octave low frequency noise with centre frequency of 500 Hz, and high frequency noise with centre frequency of 5.0 kHz for normal hearing subjects and 3.15 kHz for hearing impaired subjects. They also used 1 s broadband stimulus consisting of telephone ring. An array of 13 speakers was placed at a distance of 1 m from the listener, at angles varying from -90^0 to $+90^0$ relative to the listener, with a spacing of 15^0. Subjects were asked to keep their head still and pointed to 0^0. Sounds were presented by randomly selecting one of the speakers. The subject's task was to identify the speaker from which the target sound was perceived. Hearing impaired subjects were tested with and without their hearing aids. The study showed that although bilateral hearing aids distort the inter-aural cues degrading the localization ability, hearing impaired listeners could use inter-aural difference cues for source localization. In localization experiments involving presentation from an array of speakers, the head position has a large influence on the response and hence subject needs to maintain steady head position throughout the test. This problem can be overcome by using head related transfer functions (HRTFs), to simulate directionality. HRTFs describe the transfer of acoustical signal from a source in the free field to the eardrum of both the ears and can be used to generate spatial sounds for the investigation [15].

Murase *et al*. [16] studied the effect of dichotic listening on source localization, by dividing the speech spectrum into two bands (based on the formant frequencies of Japanese vowels i.e. with frequency boundary of 0.8 kHz and 1.6 kHz), and presenting them dichotically, for hearing impaired and normal hearing subjects. Stimuli included three speech sentences used in daily conversation. To simulate directionality, speech signals were convolved with HRTFs of a dummy head and torso obtained at five directions (-90^0, -45^0, 0^0, 45^0, and 90^0). The convolved speech signal was processed using one of the two pairs of low pass and high pass filters (with crossover frequency of 0.8 kHz and 1.6 kHz), for binaural dichotic presentation. Listening tests were conducted on six normal hearing and three hearing impaired subjects using a pair of headphones. It was reported that normal hearing listeners generally perceived two sound images corresponding to both low and high frequency bands. For the hearing impaired listeners, the sound images localized towards the side of the low frequency band.

In the above work [16], speech was divided into only two bands, and lateralization was more significant towards low frequency band, thus affecting the localization. Earlier investigations in our laboratory on binaural dichotic presentation have used comb filters based on auditory critical bands and the objective of present investigation is to study the effect of this scheme on source localization. As a first step, we have conducted the study with normal hearing subjects. Processing using HRTFs has been used for generating spatial sounds with presentation using headphones. CIPIC HRTF database [17] [18], a public domain database, provides HRTFs for different combination of azimuth and elevation. The detailed method of HRTF measurement along with anthropometric parameters is available in [17]. We have used the HRTFs, for one of the standard subjects from this database, for 0^0 elevation (horizontal plane) and frontal azimuth angle varying from -90^0 (extreme left) to $+90^0$ (extreme right) in steps of 10^0, to generate spatial sounds in this study.

METHODOLOGY

The signal processing scheme for studying source localization is shown in Fig. 1. Input signal is digitized at 10 k samples/s and 16-bit quantization. The HRTFs in CIPIC database are available as 200-point impulse responses at 44.1 k samples/s. Filtering with each HRTF involved up-sampling the input signal at this rate, convolving with the corresponding impulse response, and down-sampling the signal again. For diotic presentation, the two filtered outputs are presented to the two ears through the left and right channels of sound card using headphones. For dichotic presentation, the HRTF outputs are processed through a pair of critical band based linear phase comb filters, with complimentary magnitude response [7] - [9]. FIR filters were designed using frequency sampling technique (512 point response, f_s= 10 k samples/s) with a maximum gain ripple of 1 dB in the pass band, minimum stop band attenuation of 29 dB, transition band of 45-55Hz, and inter-band crossover gain of 4-6 dB below the pass band gain.

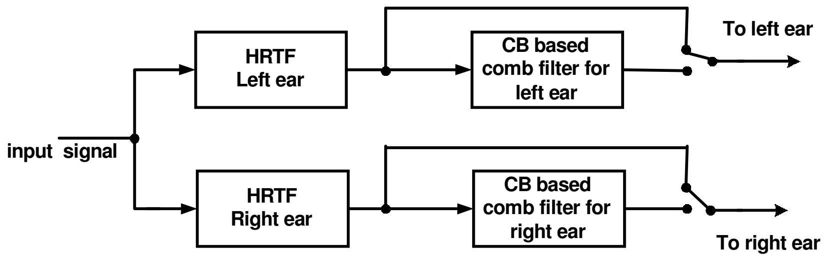

Figure 1.- Signal processing for studying source localization

Listening tests for studying the source localization were conducted on six subjects with normal hearing. For studying localization with broadband signals, we have used (i) sound of breaking glass, (ii) syllable /aba/, and (iii) white noise. For studying the effect of signal bandwidth on source localization, white noise was band-pass filtered with center frequency of 1500 Hz and bandwidth of 1 octave, 1/3 octave, and 1/6 octave. For a detailed study of source localization, three experiments were conducted: (i) source direction identification, (ii) left/center/right identification, and (iii) left/center/right discrimination threshold. The same subjects participated in all the three experiments. Sound was presented binaurally at the most comfortable level as selected by the individual subjects. In each of the three experiments, the stimuli used were processed (a) for diotic presentation, i.e. with only HRTFs and (b) for dichotic presentation, i.e. HRTFs followed by comb filters. In further description in this paper, the diotic and dichotic presentations are referred to as conditions A and B respectively.

Experiment I: Direction identification
The sounds were processed with HRTFs corresponding to 0^0 elevation and azimuth angles of 0^0, $\pm30^0$, $\pm60^0$, and $\pm90^0$. These directions were displayed in a chart kept in front of the subject (Fig.2). Stimuli processed for different angles were presented in a random order, with each angle repeated five times. The subject identified the perceived direction of the source as one of these seven angles. At the end of the test, the responses were tabulated as a presentation angle vs response angle (stimulus-response) matrix. Tests were conducted with the three broadband stimuli (breaking glass, syllable /aba/, and broadband noise). The basic objective of this experiment was to compare identification scores under the two processing conditions.

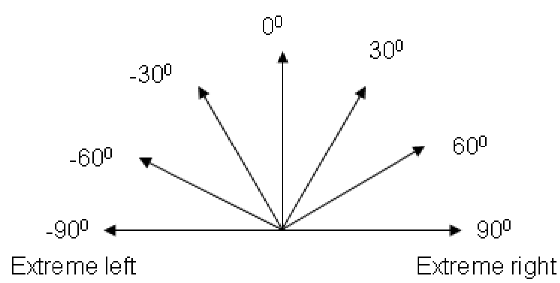

Figure 2.- Seven choice angles used in Experiment I

Experiment II: Left/center/right identification
This experiment is similar to Experiment I, except that the presentation angles were in the range of -90^0 to $+90^0$ at intervals of 10^0, and the subject identified the perceived direction as (i) left, (ii) center, or (iii) right. The three percentage responses were plotted against the azimuth angle of the source, for comparing the sharpness in the left-center and right-center transitions under the two conditions.

Experiment III: Left/center/right discrimination threshold
The objective was to determine the smallest value of azimuth angle on the left and right side for source to be localized on the corresponding side of the center, using threshold detection technique. The azimuth angles used for presentation were in $\pm90^0$ range, with a step of 10^0 as in Experiment II. Presentation was started from one of the extreme positions. The subject identified

the presented sound as left, center, or right. In case of correct response, the angle was decreased by 10^0. For incorrect response, it was increased by 20^0. Presentation of a specific angle was limited to five times. At the end of the test, the responses were used to determine thresholds for left-center and right-center discrimination. These tests were conducted for broadband noise and noise band-pass filtered with center frequency of 1500 Hz and bandwidth corresponding to 1 octave, 1/3 octave, and 1/6 octave, in order to study the effect of signal bandwidth on direction discrimination.

RESULTS AND DISCUSSION

For Experiment I, the direction identification scores were tabulated for each of the three broadband sounds. The results are summarized in Table 1 for all the subjects. Under diotic listening (condition A), the scores for the three broadband stimuli were found to be similar. With dichotic listening (condition B), there is a moderate decrease in scores for all the three stimuli. For broadband noise, stimulus-response matrix, with responses by all the subjects merged together is shown in Table 2. We see that with diotic presentation, the responses are mainly along diagonal, showing that HRTFs were able to successfully generate spatial perception. It is seen that with dichotic presentation, there is a larger spread in error, but responses follow the diagonal. Results for the other two broadband sounds showed the same pattern.

Table 1. - Experiment I: Direction identification scores (%)

Sub.	Breaking glass		Syllable /aba/		Broad-band noise	
	A	B	A	B	A	B
PNK	86	63	77	49	86	77
JAY	57	46	74	46	54	37
DSA	54	40	46	34	51	51
RSA	49	49	66	37	57	49
RSH	49	38	54	40	63	49
SGK	54	29	43	40	69	69
Mean	58	44	60	41	63	55
Std.	14	12	14	6	13	15
Med.	54	43	60	40	60	50

Table 2.- Experiment I: Identification scores for perceived angle vs target angle. Test material: broadband noise. No. of total presentation for each angle = 30 (5 presentations×6 subjects)

		A (Diotic)							B (Dichotic)						
		Perceived angle (^{o})							Perceived angle (^{o})						
		-90	-60	-30	0	30	60	90	-90	-60	-30	0	30	60	90
Target angle (^{o})	-90	20	9	1	-	-	-	-	17	9	4	-	-	-	-
	-60	6	17	7	-	-	-	-	3	15	12	-	-	-	-
	-30	-	10	19	1	-	-	-	2	13	14	1	-	-	-
	0	-	-	8	22	-	-	-	1	1	9	19	-	-	-
	30	-	-	-	-	23	5	2	-	-	-	1	21	4	4
	60	-	-	-	-	8	18	4	-	-	-	1	9	16	4
	90	-	-	-	-	3	14	13	-	-	-	3	2	13	12

For Experiment II, a plot of averaged (across the subjects) percentage score for left, right, and center identification, as a function of azimuth angle is shown in Figure 3. For diotic presentation (condition A), we see a sharp transition in response with azimuth angle, indicating that HRTFs successfully provide spatial perception. For dichotic presentation (condition B) there is a decrease in the recognition scores and slight smearing of responses. The separation angle between crossover points for left/center and right/center identification were obtained from these plots for individual subjects, under the two processing conditions, and these values are tabulated in Table 3. Average increase in the separation angles between the thresholds after splitting was 7.8^0, 16.8^0, and 33^0 for broadband noise, sound of breaking glass, and syllable /aba/ respectively.

Results of Exp.III, showing the mean of the left/center and right/center discrimination thresholds, are tabulated in Table 4. It is seen that the difference in the threshold values under the two conditions is larger for narrowband noise.

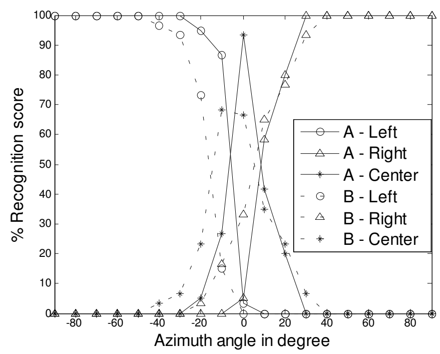

Figure 3.- Experiment II: percentage scores (avg. across subjects) for left, center, and right identification for the diotic (A) and dichotic (B) presentations. Test material: broadband noise

Table 3.- Experiment II: Seperation between left/center and right/center identification cross-over under two processing conditions, for different test material

Sub.	Breaking glass		Syllable /aba/		Broadband noise	
	A	B	A	B	A	B
PNK	15	39	15	44	28	36
JAY	10	17	10	36	21	19
DSA	20	33	18	50	10	21
RSA	13	56	10	30	10	21
RSH	10	22	10	41	10	31
SGK	28	30	36	96	41	39
mean	16.0	32.8	16.5	49.5	20.0	27.8
Std.	7.0	13.8	10.1	23.8	12.7	8.6
med.	14.0	31.5	12.5	42.5	15.5	26.0

Table 4.- Experiment III: Mean of left/center and right/center discrimination thresholds (deg.) for various test material. A: diotic (without spectral splitting), B: dichotic (with spectral splitting)

Sub.	Breaking glass		Syllable /aba/		Broadband noise		1 oct. noise		1/3 oct. noise		1/6 oct. noise	
	A	B	A	B	A	B	A	B	A	B	A	B
PNK	10	15	10	25	10	20	10	20	10	15	25	30
JAY	10	10	25	30	10	10	10	10	10	20	20	40
DSA	10	20	15	25	10	15	10	15	10	25	10	30
RSA	10	20	10	15	10	10	10	15	10	15	10	25
RSH	10	10	10	30	10	15	10	10	10	10	15	15
SGK	10	10	15	20	15	15	15	15	15	25	20	15
Mean	10.0	14.2	14.2	24.2	10.8	14.2	10.8	14.1	10.8	18.3	16.7	25.8
Std.	0	4.9	5.8	5.8	2.0	3.8	2.0	3.7	2.0	6.1	6.1	9.7
Med.	10.0	12.5	12.5	25.0	10.0	15.0	10.0	15.0	10.0	17.5	17.5	27.5

CONCLUSIONS

Earlier investigations using critical bandwidth based spectral splitting have shown that this scheme is beneficial in speech perception for persons with moderate bilateral sensorineural hearing loss. The objective of the present study was to investigate the effect of this scheme on source localization in horizontal plane. Three separate experiments were conducted. In Experiment I, we studied azimuth angle identification scores for diotic and dichotic presentations. It was observed that dichotic presentation slightly degraded the scores. In Experiment II, left/center/right identification was compared under the two conditions and it was observed that, dichotic condition reduced the sharpness in response transition with azimuth angle. In Experiment III, left/center/right discrimination thresholds were determined for noise stimulus with different bandwidths. It was concluded that difference in the mean thresholds for the two processing conditions were marginal for sounds with bandwidths in excess of 1/3 octave. Thus the experiments showed that even though source localization is moderately affected by dichotic presentation, subjects were able to perceive direction for broadband sounds, by using ITD and ILD cues across the bands in the comb filters used for spectral splitting.

In the present study, listening tests were conducted on normal hearing subjects only. For a thorough investigation, a similar study should be carried out with hearing impaired subjects.

References

[1] B. C. J. Moore: An Introduction to the Psychology of Hearing. 4th ed. London: Academic, 1997

[2] B. R. Glasberg, B. C. J. Moore: Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. Journal of the Acoustical Society of America **79** (1986) 1020-1033

[3] A. E. Carney, D. A. Nelson: An analysis of psychophysical tuning curves in normal hearing and pathological ears. Journal of the Acoustical Society of America **73** (1983) 268-278

[4] P. E. Lyregaard: Frequency selectivity and speech intelligibility in noise. Scand. Audiol. Suppl., **15** (1982) 113-122

[5] T. Lunner, S. Arlinger, J. Helgren: 8-channel digital filter bank for hearing aid use: preliminary results in monaural, diotic and dichotic modes. Scand. Audiol. Suppl. **15** (1993) 75-81

[6] D. S. Chaudhari, P. C. Pandey: Dichotic presentation of speech signal with critical band filtering for improving speech perception. Proceedings of IEEE/ICASSP 1998 (Seattle, Washington) AE 3.1

[7] A. N. Cheeran, P. C. Pandey: Dichtic presentation for binaural hearing aids using perceptually balanced critical bandwidth based comb filters. Proceedings of SCI 2003 (Orlando, Florida) V-360-365

[8] A. N. Cheeran: Speech processing with dichotic presentation for binaural hearing aids for moderate bilateral sensorineural loss. Ph.D. dissertation, School of Biosciences and Bioengineering, IIT Bombay. 2005

[9] P. N. Kulkarni, P. C. Pandey, D. S. Jangamashetti: Perceptually balanced filter response for binaural dichotic presentation to reduce the effect of spectral masking. Journal of the Acoustical Society of America **120** (2006) 3253

[10] W. Chung, S. Carlile, P. Leong: A performance adequate computational model for auditory localization. Journal of the Acoustical Society of America **107** (2000) 432-445

[11] V. Best, S. Carlile, C. Jin, A. van Schaik: The role of high frequencies in source localization. Journal of the Acoustical Society of America **118** (2005) 353-363

[12] H. J. Simon: Bilateral amplification and sound localization: then and now. Journal of Rehabilitation Research and Development **42** (2005) 117-132

[13] J. C. Middlebrooks: Narrow-band sound localization related to external ear acoustics. Journal of the Acoustical Society of America **92** (1992) 2607-2624

[14] T. Van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, J. Wouters: Horizontal localization with bilateral hearing aids: without is better than with. Journal of the Acoustical Society of America **119** (2006) 515-526.

[15] E. H. A. Langedijk, A. W. Bronkhorst: Contribution of spectral cues for human sound localization. Journal of the Acoustical Society of America **112** (2002) 1583-1596

[16] A. Murase, F. Nakajima, S. Sakamoto, Y. Suzuki, T. Kavase: Effect of sound localization with dichotic-listening digital hearing aids. Proceedings of 18th International Congress on Acoustics (ICA 2004, Kyoto, Japan) 1519 - 1522

[17] V. R. Algazi, R. O. Duda, D. M. Thompson: The CIPIC HRTF database. Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (Mohonk, New York, 2001) 99 – 102

[18] CIPIC HRTF Database, University of California, Davis, Cal. (2001). Website http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm