WAVELET PACKETS-BASED SPEECH ENHANCEMENT FOR HEARING AIDS APPLICATION

Jiming Yang and Sridhar Krishnan
Dept. of Electrical and Computer Engineering, Ryerson University, Toronto, ON, M5B 2K3
[jyang, krishnan]@ee.ryerson.ca

1. INTRODUCTION

Hearing loss is the fastest growing chronic disability is Canada. Three million Canadian have certain degree of hearing impairment. Sensorineural hearing loss is the most common type of hearing impairments, which typically results in increased hearing thresholds and increased perceived loudness growth. As a result, impaired listeners frequently experience difficulty understanding speech, especially under noisy conditions [1].

Fourier domain methods such as spectral subtraction (S.S.) and Ephraim-Malah filter are effective reducing noise in speech processing. Wavelet-based methods are explored to reduce computational complexity and to achieve better noise reduction performance. Wavelet packet transform (WPT) has good time-frequency localization; it may not require overlapped signal windows thus process less data. Flexible multi-resolution analysis can be easily achieved to decompose signal according to critical bands. Some of the WPTs also have very reasonable computation complexity.

Loudness compression compensation uses the idea of hearing model. The purpose of a model based hearing aid is to prewarp the signal so that a hearing impaired listeners will hear it just as a normal hearing individual would hear it. It is accomplished by combining the normal auditory model with the inverse of the impaired model. The sound is processed with the model for a normal hearing subject followed by the inverse model for hearing impaired subjects. This enables close normal hearing perception for a impaired hearing.

In this paper, we propose a robust hearing aids system which will simultaneously conduct hearing compensation and noise removal by using critical band WPT. With a reasonable complexity, the new speech enhancement scheme can improve perception in noise for hearing-impaired listeners. The proposed speech enhancement scheme can be used to improve next generation hearing aids device performance

2. METHOD

Figure 1 shows the block diagram of the proposed speech enhancement scheme. First the noisy speech time series is decomposed according to psychoacoustic critical bands by using WPT. Wavelet coefficients (WCs) of noisy speech are perceptually weighted through a weighting function incorporating masking properties [2]. The denoised WCs are then compressed to compensate recruitment of loudness problem of hearing impaired on the Compression stage. IWPT transform the processed signal back to time domain. To evaluate the processed speech on normal hearing subjects, Hearing loss simulation introduces effect of hearing loss to both original noisy signal and processed signal to simulate certain types of hearing impairment.
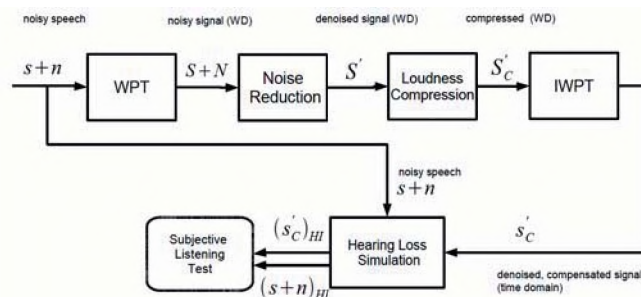

Figure 1. Block diagram of proposed method

2.1 Modified PTFS noise reduction

In wavelet domain, noisy input-to-noise ratio(INR) is estimated in each critical band. Based on the ratio, a weighting function $H(m,n)=f(INR)$ is developed. The function incorporates properties of masking phenomenon. In a noise dominant band, minimum weight is given to eliminate noise. In a signal dominant band, WCs are not modified under the assumption that noise is masked by signal. In a signal noise co-dominant band, WCs are weighted by the multiplying the weighting function $H(m,n)$.

2.2 Loudness compression

Compression projects the WCs of input signal in normal listener's dynamic range to the impaired listener's. So that the loudness perceived by the impaired listener is equal to that of normal listener. $C^{*} = T_{im}+(C- T_{nor})\Delta^{*}/\Delta$ (C and C^{*} are WCs before and after compression, T_{im} and T_{nor} are hearing thresholds, Δ and Δ^{*} are dynamic ranges of normal and hearing impaired) [3]. Wavelet-based compression can be easily modified to the subject specific model of loudness recruitment.

2.3 Hearing loss simulation

The objective of this stage is to emulate hearing loss effect on normal hearing subjects. Only normal hearing

subjects are needed to d evaluate the enhancing results. Spectral smearing [4] can be used to simulate loss of frequency selectivity. Adding shaped noise raises the hearing threshold of certain bands. In this experiment, shaped noise is added to simulate the effect of high frequency hearing loss.

3. RESULTS

In the experiment, the noisy speech is obtained by corrupting the clean speech with white Gaussian noise of different SNR setting. Average segment SNR gain and subjective listening test are used to evaluate the processing results.
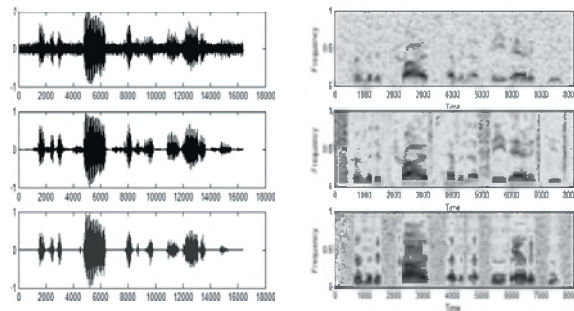

Figure 2. Waveform and spectrogram of noisy (5dB), clean and denoised signals (top to bottom)

Table 1. SegSNR improvement for enhanced speech by modified PTFS, conventional wavelet thresholding (soft), and spectral subtraction(SS)

Input SNR(dB)	SegSNR Gain(dB)		
	Modified PTFS	DWT	S.S.
-5	7.3	6.4	5.5
0	6.4	3.8	4.2
5	4.6	2.4	2.6

The waveforms (left) and spectrogram (right) of the enhanced speech are shown in Figure 2. It shows the algorithm efficiently removes the background noise. But the technique seems to attenuate the signal more in low SNR region.

The proposed noise reduction algorithm is compared with wavelet soft thresholding and spectral subtraction methods at various SNR level. Average Segment SNR gain in Table 1 demonstrates the modified PTFS method outperforms other two methods in all different SNR level. DWT and SS performances are relatively close.
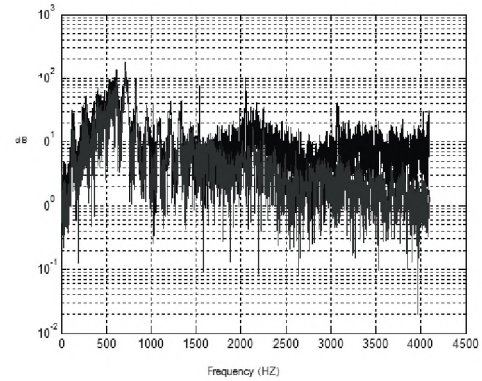

Figure 3 Spectral plot of signal before (light) and after (dark) loudness compression

Informal subjective listening test results with 11 subjects show the denoised speech from modified PTFS sounds natural, when removing the noise dramatically. S.S. is able to remove noise more than DWT but tends to introduce significant amount of musical noise.

Figure 3 illustrates that the compression stage introduces gain in high band thus compensating high frequency loss while keeps lower band signal relatively unchanged.

Listening test shows enhanced signal after hearing loss simulation produce much better speech quality than the unprocessed one. However, at low SNR level (-5dB), improvement is less significant.

4. CONCLUSION

Our proposed method removes background noise without introducing noticeable perceptual distortion. It also compensates frequency dependent loudness recruitment. This work advances WPT as an adequate and flexible choice for time-frequency speech processing. The complete system can serve as an efficient and effective way to develop and evaluate algorithm for hearing aids application.

REFERENCES

[1] Levitt, H(2001). Noise reduction in hearing aids: a review. Journal of rehabilitation research and development. Vol.38, No.1

[2] Li, M, MacAllister, H.G., Black, N.D, De Perez T.A.(2001) Perceptual time frequency subtraction algorithm for noise reduction in hearing aids. IEEE transactions on biomedical engineering, Vol.48, No.9

[3] Drake, L.A., Rutledge J.C. (1993). Wavelet analysis in recruitment of loudness compensation. IEEE Transactions on Signal Processing, Vol.41, No.12

[4] Baer, T, and Moore, B (1993). Effects of spectral smearing on the intelligibility of sentences in noise. J. Acoust. Soc. Am, 94, 1229:1241