LETTER

Sub-Band Noise Reduction in Multi-Channel Digital Hearing Aid

Qingyun WANG[†a)], ***Nonmember***, **Ruiyu LIANG**[†], ***Member***, **Li JING**[††], **Cairong ZOU**[††], ***and*** **Li ZHAO**[††], ***Nonmembers***

SUMMARY Since digital hearing aids are sensitive to time delay and power consumption, the computational complexity of noise reduction must be reduced as much as possible. Therefore, some complicated algorithms based on the analysis of the time-frequency domain are very difficult to implement in digital hearing aids. This paper presents a new approach that yields an improved noise reduction algorithm with greatly reduce computational complexity for multi-channel digital hearing aids. First, the sub-band sound pressure level (SPL) is calculated in real time. Then, based on the calculated sub-band SPL, the noise in the sub-band is estimated and the possibility of speech is computed. Finally, a posteriori and a priori signal-to-noise ratios are estimated and the gain function is acquired to reduce the noise adaptively. By replacing the FFT and IFFT transforms by the known SPL, the proposed algorithm greatly reduces the computation loads. Experiments on a prototype digital hearing aid show that the time delay is decreased to nearly half that of the traditional adaptive Wiener filtering and spectral subtraction algorithms, but the SNR improvement and PESQ score are rather satisfied. Compared with modulation frequency-based noise reduction algorithm, which is used in many commercial digital hearing aids, the proposed algorithm achieves not only more than 5dB SNR improvement but also less time delay and power consumption.

key words: *sub-band noise reduction, multi-channel digital hearing aid, adaptive Wiener filtering, modulation frequency-based noise reduction*

1. Introduction

In digital hearing aids, raising the noise reduction efficiency is vital to enhance the experience of the hearing-impaired person. Auditory cognition is influenced by the signal-to-noise ratio (SNR) of the speech, especially for hearing-impaired person [1]. In noisy situations, even loud speech may not be understood by the listener [2]. The noise overwhelms the auditory cognition of the hearing-impaired person and disturbs the perception processing of the speech in the neural system [3]. Some excellent noise reduction algorithms have been proposed in recent years. Reference [4] describes some DFT-domain based single-microphone noise reduction methods for speech enhancement. The adaptive Wiener filtering algorithm has been developed to reduce noise by tracing the noise spectral adaptively [5]. Andrew L. Maas from Stanford University proposed a new noise reduction algorithm based on deep neural network and achieved excellent denoising performance [6]. Besides, orthogonal decomposition [7] and wavelet transform [8] methods are exploited to improve the performance of the noise reduction. Lamentably, most of the algorithms are not fit for digital hearing aid because they fail to satisfy the real-time requirement and power dissipation limitation.

In order to improve noise reduction efficiency, a new approach is proposed for multi-channel digital hearing aid. Through the sub-band noise estimation, the proposed algorithm adjusts the gain function of each channel to suppress the noise. By replacing the FFT and IFFT transforms by the known SPL, the proposed algorithm greatly reduces the computation loads. In order to acquire the real time performance in the real acoustic environment, the proposed algorithm has been programmed and transplanted into a prototype hearing aid. Experiments of the proposed, spectral subtraction, adaptive Wiener filtering and modulation frequency-based algorithms show that SNR and PESQ (the perceptual evaluation of speech quality, PESQ) of the proposed algorithm are rather satisfied, and just slightly less than those of the adaptive Wiener filtering algorithm. Moreover, the proposed algorithm has the least time delay.

2. Methods

2.1 Diagram of Multi-Channel Digital Hearing Aid

Diagram of multi-channel digital hearing aid processing with the proposed sub-band noise reduction algorithm is shown in Fig. 1. Digital speech signal, which is received by microphone and converted by ADC, is preemphasized firstly to make the spectrum flat and avoid excessive loss of high frequency component in the following noise reduction. Correspondingly, the processed signal is deemphasized before output to speaker. As shown in Fig. 1, the preemphasized signal is analyzed by 6-order IIR filter bank to 16 channel sub-band signals.

2.2 Sub-Band Noise Reduction Algorithm

In sub-band k, the noisy speech signal $\mathbf{y}_k$ is segmented to frames. For the m-th frame, the noisy speech signal $\mathbf{y}(m,k)$ can be expressed as

$$\mathbf{y}(m,k) = \mathbf{s}(m,k) + \mathbf{n}(m,k) \quad (1)$$

where $\mathbf{s}(m,k)$ denotes speech signal and $\mathbf{n}(m,k)$ denotes

Manuscript received July 22, 2015.
Manuscript revised September 26, 2015.
Manuscript publicized October 14, 2015.

[†]The authors are with School of Communication Engineering, Nanjing Institute of Technology, Jiangsu Nanjing, 211167, China.
[††]The authors are with School of Information Science and Engineering, Southeast University, Jiangsu Nanjing, 210096, China.
a) E-mail: wangqingyun@vip.163.com

DOI: 10.1587/transinf.2015EDL8162

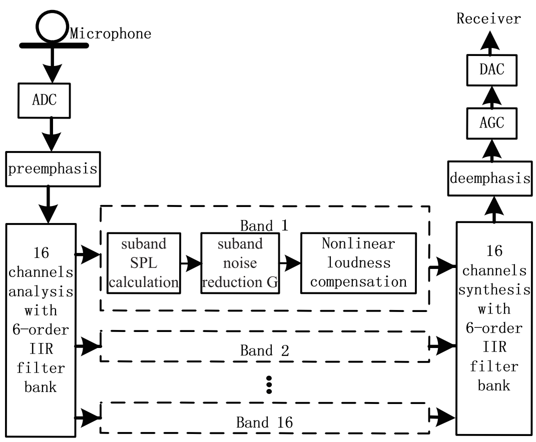

Fig. 1 Diagram of multi-channel digital hearing aid with sub-band noise reduction.

noise. The sub-band noise reduction unit improves the quality of the speech by adjusting the sub-band gain $G(m,k)$. The adjustable gain $G(m,k)$ is defined as

$$G(m,k) = \frac{SNR_{prio}(m,k)}{1 + SNR_{prio}(m,k)} \tag{2}$$

where $SNR_{prio}(m,k)$ is the a priori SNR of frame m and sub-band k. When the quality of noisy speech deteriorates, $G(m,k)$ decreases to inhibit noise. $SNR_{prio}(m,k)$ is estimated by the decision-directed method [8], that is

$$\begin{aligned} SNR_{prio}(m,k) = \eta \cdot & \frac{\|\hat{\mathbf{s}}(m-1,k)\|^2}{\|\hat{\mathbf{n}}(m,k)\|^2} \\ & + (1-\eta) \cdot max\left(SNR_{post}(m,k) - 1, 0\right) \end{aligned} \tag{3}$$

where $\|\hat{\mathbf{s}}(m-1,k)\|^2$ is the power of the estimated previous pure speech signal. $\|\hat{\mathbf{n}}(m,k)\|^2$ is the power of the estimated noise. The coefficient $\eta \in [0,1]$ decides the performance of noise reduction. The closer to 1 the value is, the more successfully the musical noise is inhibited, but the more severely the distortion of the speech is. $SNR_{post}(m,k)$ is the a posteriori SNR. In traditional noise reduction method, $SNR_{post}(m,k)$ is calculated by the spectrums of input signal and the estimated noise through FFT. Here, to decrease the computational complexity, the a posteriori SNR of frame m and sub-band k is modified as

$$SNR_{post}(m,k) = \frac{\|\mathbf{y}(m,k)\|^2}{\|\hat{\mathbf{n}}(m,k)\|^2} \tag{4}$$

$\|\mathbf{y}(m,k)\|^2$ and $\|\hat{\mathbf{n}}(m,k)\|^2$ are the powers of input signal and the estimated noise, respectively. As shown in Fig. 1, sub-band SPL is calculated by Eq. (5).

$$SPL(m,k) = 20\lg \frac{\|\mathbf{y}(m,k)\|^2}{R \times p_{ref}} \tag{5}$$

Here, $p_{ref} = 20 \times 10^{-5} Pa$ is the reference sound pressure. R is the constant decided by the system feature. In multi-channel digital hearing aid, $\|\mathbf{y}(m,k)\|^2$ and $SPL(m,k)$ must be calculated before nonlinear loudness compensation. Therefore, the computational complexity of Eq. (4) is decided only by the calculation of $\|\hat{\mathbf{n}}(m,k)\|^2$.

Four steps are exploited to estimate the power of the noise $\|\hat{\mathbf{n}}(m,k)\|^2$. For step 1, the power of the sub-band signal $\|\mathbf{y}(m,k)\|^2$ is calculated by

$$P(m,k) = \alpha \cdot P(m-1,k) + (1-\alpha) \cdot \|\mathbf{y}(m,k)\|^2 \tag{6}$$

where α is the coefficient of smoothness and $\alpha \in [0,1]$.

For step 2, the minimum power of the sub-band signal is traced by

$$\begin{cases} P_{\min}(m,k) = \gamma \cdot P_{\min}(m-1,k) \\ \quad + \frac{1-\gamma}{1-\beta}(P(m,k) - \beta \cdot P(m-1,k)) \\ \qquad\qquad P_{\min}(m-1,k) < P(m,k) \\ P_{\min}(m,k) = P(m,k) \qquad\qquad \text{else} \end{cases} \tag{7}$$

Here, β and γ are empirical coefficients.

For step 3, $I(m,k)$ is calculated to decide if speech exists in sub-band signal.

$$I(m,k) = \begin{cases} 1, & P(m,k)/P_{\min}(m,k) > \delta \\ 0, & P(m,k)/P_{\min}(m,k) \le \delta \end{cases} \tag{8}$$

where δ is the threshold. Then the possibility of speech in the sub-band is calculated by

$$p(m,k) = \xi \cdot p(m-1,k) + (1-\xi) \cdot I(m,k) \tag{9}$$

where ξ is the coefficient of possibility updating.

For step 4, the power of the estimated sub-band noise is updated by

$$\begin{cases} \|\hat{\mathbf{n}}(m,k)\|^2 = \psi(m,k)\|\hat{\mathbf{n}}(m-1,k)\|^2 \\ \qquad\qquad + (1-\psi(m,k))\|\mathbf{y}(m,k)\|^2 \\ \psi(m,k) = \alpha + (1-\alpha) \cdot p(m,k) \end{cases} \tag{10}$$

where α is the coefficient of smoothness defined by Eq. (6). Having the estimated $\|\hat{\mathbf{n}}(m,k)\|^2$, the power of the estimated pure speech is derived as

$$\|\hat{\mathbf{s}}(m,k)\|^2 = \|\mathbf{y}(m,k)\|^2 - \|\hat{\mathbf{n}}(m,k)\|^2 \tag{11}$$

By substitution of Eqs. (5)~(11) into Eqs. (2)~(3), the gain function $G(m,k)$ is acquired and the noise is reduced. The coefficients η, α, β, γ and ξ should be adjusted carefully. Good tradeoff, not only between the speech quality and the distortion, but also between the response speed and the estimation accuracy, is made by the adjustment of the coefficients. In the experiments in Sects. 3.1, 3.2 and 3.3, we set the coefficient values as $\eta = 0.9$, $\alpha = 0.7$, $\beta = 0.96$, $\gamma = 0.998$, $\delta = 2$ and $\xi = 0.2$.

3. Simulations and Experiments

3.1 Experimental Arrangement

In multi-channel digital hearing aid, there are two main real-time methods to suppress the background noise in sub-bands. The first one is named the modulation frequency-based method, such as the noise reduction in Octicon Syncro digital hearing aid. Because the computation of the modulation depth is mild, it costs low time and power consumption, but the improvement of the speech quality is not satisfied. Another method is based on the adaptive Wiener filtering [5]. Utilizing the spectrum estimation of noise, it has achieved satisfied performance of noise reduction, but costs high computational complexity and power consumption.

Performances of noise reduction between the proposed method, the above two schemes and other traditional algorithms were compared. The experiments were conducted in the silent room. The audio scene simulation software SurroundRouter was used to generate testing acoustic scenes. Four kinds of noises (white, speech babble, tank, destroyer engine) were superimposed to the clean speech signals, which was sampled by 16kHz sampling rate. The prototype digital hearing aid was located in the center of the speaker array. In the digital hearing aid, the speech signal was analyzed to 16 channels. The frequency bands of 16 channels were divided according to the human auditory with reference to the gammatone filter bank. The noise reduction algorithms were programmed and transplanted to the auditory DSP SOC in the prototype digital hearing aid. The length of the frame in sub-bands was 32, 64 or 128 samples and the frame shift was set to 25% of the frame length, so the time delay caused by frame segmentation was lower than several milliseconds. In order to acquire the performance of the algorithms, we designed the analysis software running in the computer combined with the prototype digital hearing aid by the serial data bus.

3.2 Waveform and Spectrogram Improvement

We compared the waveform and spectrogram improvements for spectral subtraction, adaptive Wiener filtering, modulation frequency-based, the proposed algorithms. Fifty speech segments were tested with four kinds of noises. One group of the spectrograms of pure speech, noisy speech (SNR = 5dB), the outputs of the above four algorithms are shown in Fig. 2.

In the spectrograms, the color of the image indicates the amplitude of the particular frequency at the particular time. The colors from the smallest amplitude to the largest amplitude are deep blue, green, orange, yellow and red. From Fig. 2, the adaptive Wiener filtering method acquires the deepest blue background and the clearest tracks of the pitches and formants, which means the smallest background noise and the cleanest output. The proposed method achieves the second best performance but the computational complexity is much smaller than that of the adaptive Wiener filtering method. The residual noise of the modulation frequency-based method is the largest, while that of the spectral subtraction is moderate.

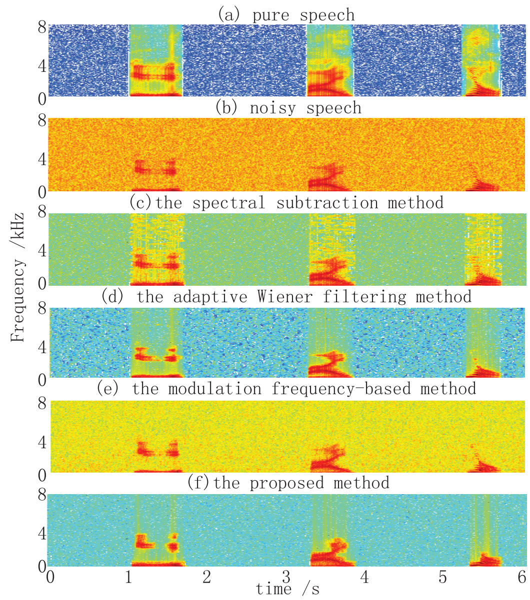

Fig. 2 Spectrograms of four noise reduction algorithms.

3.3 SNR and PESQ Improvement

SNR and PESQ are important for speech quality evaluation. PESQ assesses speech quality by estimating the overall loudness difference between the noise-free and processed signals [10]. PESQ score is computed as a linear combination of the averaged symmetrical disturbance value d_{SYM} and the averaged asymmetrical disturbance value d_{ASYM} as Eq. (12).

$$PESQ = 4.5 - 0.1d_{SYM} - 0.0309d_{ASYM} \qquad (12)$$

The parameters d_{SYM} and d_{ASYM} are calculated by the reference [11]. SNR and PESQ improvements of the above methods with different noises and different input SNRs are compared when frame length is 32, 64 and 128 samples, and the result of 128 samples is shown in Table 1. Here, fifty speech segments with SNR = 5dB, 10dB are tested for each kind of noise and the values are the average. Result is similar when frame length is 32 or 64 samples.

From Table 1, among the above four noise reduction algorithms, the adaptive Wiener filtering algorithm achieves the best SNR and PESQ improvements. When the SNR of the input speech is 10dB, the best 10dB increase is obtained with white noise. The improvement of the proposed algorithm is the second best. The spectral subtraction algorithm follows, and the modulation frequency-based algorithm achieves the worst. The difference of PESQ of the above four algorithms is milder than that of the SNR. The best PESQ improvement (nearly 0.98) is accessed by the adaptive Wiener filtering algorithm when input SNR = 10dB. The second best PESQ improvement (about 0.95) is accessed by the proposed algorithm.

Table 1 SNR and PESQ improvements when frame length is 128 samples.

Methods	Input SNR	white		tank		babble		engine	
		ΔSNR	ΔPESQ	ΔSNR	ΔPESQ	ΔSNR	ΔPESQ	ΔSNR	ΔPESQ
spectral subtraction	5	2.86	0.46	3.36	0.26	1.71	0.67	2.25	0.50
	10	3.56	0.22	3.06	0.09	2.39	0.24	2.98	0.32
adaptive Wiener filtering	5	9.12	0.89	7.02	0.81	5.59	0.67	6.99	0.93
	10	10.33	0.88	9.34	0.77	8.78	0.39	9.57	0.97
modulation frequency-based	5	2.51	0.37	2.09	0.38	1.34	0.36	1.59	0.47
	10	3.23	0.18	2.70	0.21	2.34	0.27	3.12	0.34
the proposed	5	8.89	0.94	6.59	0.63	5.49	0.67	6.74	0.81
	10	9.88	0.59	9.12	0.62	8.37	0.32	9.21	0.94

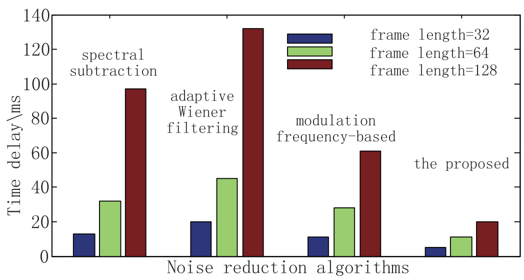

Fig. 3 Time delay in the prototype digital hearing aid.

3.4 Computational Complexity and Real-Time Performance

In digital hearing aid, the time delay from the input to the output must be limited within 40 milliseconds. Low computational complexity means not only high real time, but also low power consumption of the battery.

For spectral subtraction and the adaptive Wiener filtering algorithms, FFT and IFFT transforms are required and the cost of them in each channel is $O(N \log_2 N)$ multiplications. In the modulation frequency-based algorithm, the estimated noise is acquired by the tracking of the amplitude modulation depth. The computation complexity without FFT and IFFT is smaller than that of spectral subtraction and adaptive Wiener filtering methods, but the quality of the output signal decreases apparently. In the proposed algorithm, by replacing spectrum estimation by the power of the sub-band signals, the computational complexity is the lowest. The time delay from the analysis unit to the synthesis unit of the above four algorithms in the prototype digital hearing aid is shown in Fig. 3. Here the FFT length is set to the same as frame length.

4. Conclusions

In the present study, a novel sub-band noise reduction algorithm in the multi-channel digital hearing aid is introduced. The main focus of the work is to reduce the complexity and the time delay as well as increase the quality of the output speech. In the proposed algorithm, the noise is estimated in each sub-band by replacing the noise spectrum with the power of the sub-band signal iteratively. Compared with other noise reduction algorithms in experiments, the proposed algorithm yields excellent speech quality and real time response.

Acknowledgments

This research project was founded in part by National Natural Science Foundation (No. 61375028, 61301219), Natural Science Foundation of Jiangsu Province (No. BK20130241) and the Scientific Research Funds of Nanjing Institute of Technology (No. ZKJ201202).

References

[1] S. Kalluri and L.E. Humes, "Hearing technology and cognition," American Journal of Audiology, vol.21, no.2, pp.338–343, 2012.

[2] T. Lunner and E. Sundewall-Thorén, "Interactions between cognition, compression, and listening conditions: Effects on speech-in-noise performance in a two-channel hearing aid," J. American Academy of Audiology, vol.18, no.7, pp.604–617, 2007.

[3] T. Lunner, M. Rudner, and J. Rönnberg, "Cognition and hearing aids," Scandinavian Journal of Psychology, vol.50, no.5, pp.395–403, 2009.

[4] R.C. Hendriks, T. Gerkmann, and J. Jensen, "DFT-domain based single-microphone noise reduction for speech enhancement: A survey of the state of the art," Synthesis Lectures on Speech and Audio Processing, vol.9, no.1, pp.1–80, 2013.

[5] M.A.A. El-Fattah, M.I. Dessouky, A.M. Abbas, S.M. Diab, E.-S.M. El-Rabaie, W. Al-Nuaimy, S.A. Alshebeili, and F.E.A. El-Samie, "Speech enhancement with an adaptive Wiener filter," International Journal of Speech Technology, vol.17, no.1, pp.53–64, 2014.

[6] A.L. Maas, Q.V. Le, T.M. O'Neil, et al., "Recurrent neural networks for noise reduction in robust ASR," Proc. 13th International Conference on INTERSPEECH. pp.22–25, 2012.

[7] J. Benesty, J. Chen, Y.A. Huang, and T. Gaensler, "Time-domain noise reduction based on an orthogonal decomposition for desired signal extraction," J. Acoust. Soc. Am., vol.132, no.1, pp.452–464, 2012.

[8] R. Aggarwal, J.K. Singh, V.K. Gupta, S. Rathore, M. Tiwari, and A. Khare, "Noise reduction of speech signal using wavelet transform with modified universal threshold," International Journal of Computer Applications, vol.20, no.5, pp.14–19, 2012.

[9] C. Breithaupt and R. Martin, "Analysis of the decision-directed SNR estimator for speech enhancement with respect to low-SNR and transient conditions," IEEE Trans. Audio, Speech, Language Process., vol.19, no.2, pp.277–289, 2011.

[10] J.Y. Ma, Y. Hu, and P.C. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions," J. Acoust. Soc. Am., vol.125, no.5, pp.3387–3405, 2009.

[11] Y. Hu and P.C. Loizou, "Evaluation of objective quality measures," IEEE Trans. Audio, Speech, Language Process., vol.16, no.1, pp.229–238, 2008.