

# An adaptive multichannel loudness compensation method

Wang Xia<sup>1</sup> Liang Ruiyu<sup>2</sup> Wang Qingyun<sup>2</sup> Shen Hongming<sup>3</sup> Zhao Li<sup>1</sup> Zou Cairong<sup>1</sup>

(<sup>1</sup>Key Laboratory of Underwater Acoustic Signal Processing of Ministry of Education, Southeast University, Nanjing 210096, China)

(<sup>2</sup>School of Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China)

(<sup>3</sup>School of Electronic and Information, Nantong University, Nantong 226019, China)

**Abstract:** To alleviate the conflict between audibility and distortion in the conventional loudness compensation method, an adaptive multichannel loudness compensation method is proposed for hearing aids. The linear and wide dynamic range compression (WDRC) methods are alternately employed according to the dynamic range of the band-passed signal and the hearing range (HR) of the patient. To further reduce the distortion caused by the WDRC and improve the output signal to noise ratio (SNR) under noise conditions, an adaptive adjustment of the compression ratio is presented. Experimental results demonstrate that the output SNR of the proposed method in babble noise is improved by at least 1.73 dB compared to the WDRC compensation method, and the average speech intelligibility is improved by 6.0% and 5.7%, respectively, compared to the linear and WDRC compensation methods.

**Key words:** loudness compensation; wide dynamic range compression; gammatone filter banks; hearing aids

**doi:** 10.3969/j.issn.1003-7985.2016.02.002

Sensorineural hearing losses are characterized by an elevated hearing threshold, reduced dynamic range of hearing and decreased spectral resolution. To address the deficit of the elevated hearing threshold, the linear amplification method was usually used in early hearing aids to compensate for the loss of loudness<sup>[1]</sup>. However, unappreciated perception was produced due to insufficient gain related to the soft sounds and excessive gain for loud sounds. Nowadays, the nonlinear amplification method, i. e., the wide dynamic range compression (WDRC) method, has been widely used in digital hearing aids<sup>[2]</sup>. The WDRC method ensures audibility through the high gain for soft sounds and low gain for loud sounds. However, the benefits of improved audibility are offset by the

distortion caused by compression<sup>[3]</sup>. Compared to the WDRC method, the linear amplification method can provide higher intelligibility as long as the systems provide sufficient gain to ensure that the speech sounds fall above the impaired threshold<sup>[4]</sup>, and it can achieve superior performance than WDRC when the input speech level equals 65 dB<sup>[5]</sup>. Souza et al.<sup>[6]</sup> found that the WDRC method can decrease the input SNRs while the linear amplification method cannot, which provides a potential explanation regarding the issue why the WDRC method sometimes failed to show superiority to linear amplification for speech in background noise. As a result, the linear method combined with the WDRC method can improve the intelligibility of speech.

In this paper, an adaptive multi-channel loudness compensation method is proposed to reduce the speech distortion at the premise of maintaining audibility, and to further improve speech intelligibility. The loudness compensation method is determined by the hearing range (HR) of the patient and the dynamic range of the signal from the 5th to the 10th channels. Moreover, an adaptive WDRC method is presented to decrease the distortion introduced by the conventional WDRC method and increase the output SNR under noisy conditions.

## 1 Proposed Strategy

The block diagram of the one-channel loudness compensation method is illustrated in Fig. 1. It contains three parts: 1) The preprocessing module, in which the input signal is decomposed into band signals by gammatone filter banks, and the envelope of the signal is extracted; 2) The decision module, in which the compensation method is determined by the dynamic range of the channel signal and the HR of the patient; 3) The compensation module, in which the linear amplification method or the WDRC method is adopted for loudness compensation.

### 1.1 Preprocessing module

Recently, gammatone filter banks have been widely used in hearing device design to decompose the input signal into channel signals<sup>[7-8]</sup> due to the acceptable delay and computational efficiency. The impulse response of the gammatone filter is expressed as

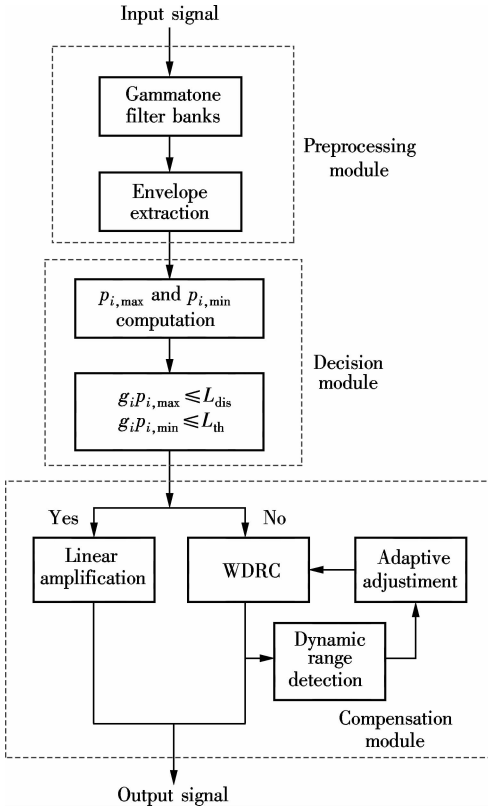
$$g_i(t) = A^n t^{n-1} e^{-2\pi B_f t} \cos(2\pi f_i t + \varphi_i) U(t) \quad 1 \leq i \leq N \quad (1)$$

**Received** 2015-10-13.

**Biographies:** Wang Xia (1980—), female, graduate; Zhao Li (corresponding author), male, doctor, professor, zhaoli@seu.edu.cn.

**Foundation items:** The National Natural Science Foundation of China (No. 61301219, 61375028, 61273266), the Natural Science Foundation of Jiangsu Province (No. BK20130241), the Fundamental Research Funds for the Central Universities (No. 2242013K30010), the Science and Technology Program of Nantong (No. BK2014002).

**Citation:** Wang Xia, Liang Ruiyu, Wang Qingyun, et al. An adaptive multichannel loudness compensation method [J]. Journal of Southeast University (English Edition), 2016, 32(2): 141 – 145. doi: 10.3969/j.issn.1003-7985.2016.02.002.



**Fig. 1** Block diagram of proposed loudness compensation method in one channel

where  $N$  is the number of the filter banks;  $A$  is the gain coefficient of the filter, which usually equals 1;  $f_i$  is the centre frequency of the filter;  $U(t)$  is the unit step function;  $\varphi_i$  is the initial phase, which is usually set to be 1;  $n$  is the order of the filter, which is usually set to be 4;  $B_i$  is the gammatone attenuation factor and its value is related to the bandwidth of the filter. The relationship between the attenuation factor and filter bandwidth can be expressed as

$$B_i = b \text{ERB}(f_i) \quad (2)$$

where  $b$  is a constant and it usually equals 1.019;  $\text{ERB}(f_i)$  is the equivalent rectangular bandwidth (ERB) and it can be expressed as<sup>[9]</sup>

$$\text{ERB}(f_i) = 24.7 \times \left( 4.37 \times \frac{f_i}{1000} + 1 \right) \quad (3)$$

The center frequency is uniformly divided in the ERB domain and it can be mapped to a linear frequency. The mapping function between the linear frequency  $f_{\text{lin}}$  and ERB frequency  $f_{\text{ERB}}$  is expressed as<sup>[10]</sup>

$$f_{\text{ERB}} = 21.4 \lg(4.37 \times 10^{-3} f_{\text{lin}} + 1) \quad (4)$$

16-channel gammatone filter banks are applied to decompose the signal in the frequency between 0 and 8 kHz.

The envelope of the channel signal is obtained by the

Hilbert transformation which is a common method for envelope extraction<sup>[11-12]</sup>. The Hilbert transformation of a real signal  $s(t)$  is expressed as

$$\hat{s}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau \quad (5)$$

$z(t)$  is the analytic signal related to  $s(t)$  and it is defined as

$$z(t) = s(t) + j\hat{s}(t) \quad (6)$$

The envelope of the real signal  $s(t)$  is the module of  $z(t)$  and it is denoted as

$$x(t) = \sqrt{s^2(t) + \hat{s}^2(t)} \quad (7)$$

## 1.2 Decision module

In the decision module,  $p_{i,max}$  and  $p_{i,min}$  of the input signal from the 5th to the 10th channels are computed. The formulas for computing SPL are shown as

$$p_{i,max} = 20 \lg \left( \frac{(\max_{1 \leq n \leq N} |x_i(n)|)^2}{p_0} \right) \quad 5 \leq i \leq 10 \quad (8)$$

$$p_{i,min} = 20 \lg \left( \frac{(\min_{1 \leq n \leq N} |x_i(n)|)^2}{p_0} \right) \quad 5 \leq i \leq 10 \quad (9)$$

where  $x_i(n)$  is the envelope of the  $i$ -th channel signal;  $N$  is the length of the channel signal;  $p_0$  is the basic sound pressure, which is 20  $\mu\text{Pa}$ . Then,  $p_{i,max}$  and  $p_{i,min}$  are compensated by the linear gain. As long as each compensated SPL from the 5th to the 10th channels lies within the hearing range, the linear amplification method is adopted, i. e.,

$$g_i p_{i,max} \leq L_{dis}, \quad g_i p_{i,min} \geq L_{th} \quad 5 \leq i \leq 10 \quad (10)$$

where  $L_{dis}$  and  $L_{th}$  denote the discomfort level and the hearing threshold of the patient, respectively. Otherwise, the WDRC method is adopted for loudness compensation. Usually, the discomfort level is difficult to measure, so it is estimated using a simple method<sup>[5]</sup>. If  $L_{th}$  is less than 60 dB,  $L_{dis}$  is set to be 105 dB. Otherwise,  $L_{dis}$  is set to be 105 dB plus half of the loss that exceeds of 60 dB. Since the frequency band between 500 and 2 000 Hz contains 70% speech information<sup>[13]</sup>, the signals of the fifth to the tenth channels are selected in the decision.

## 1.3 Compensation module

The decision module described in Section 1.2 determines the amplification method adopted for loudness compensation. The linear and WDRC methods are adaptively adopted. The disadvantage of the conventional WDRC is that the compression curve is fixed once the fitting is finished. Levitt<sup>[14]</sup> stated that personal difference cannot be omitted during compression. In order to effectively utilize the HR, reduce the signal distortion and improve the out-

put SNR under noisy conditions, an adaptive WDRC method is presented for loudness compensation. The input-output curve of the adaptive WDRC for one channel is shown in Fig. 2. Symbols  $L_k$  and  $H_k$  denote the low knee-point and high knee-point, respectively.  $L_g$  and  $H_g$  are the corresponding gains of  $L_k$  and  $H_k$ . The compression curve can be adjusted in accordance with the maximum and minimum SPL of the output signal. The adjustment rules are described as follows: 1) Decreasing compression-ratio (CR) rule. When the maximum output SPL is less than  $L_{dis}$  and the minimum output SPL is more than  $L_{th}$ , the gain

of high knee-point  $H_g$  is increased and the low knee-point gain  $L_g$  is simultaneously decreased unless CR equals 1. 2) Comfort rule. As long as the maximum output SPL exceeds  $L_{dis}$ ,  $H_g$  is decreased unless  $H_g$  equals the initially prescribed gain. 3) Audibility rule. When the minimum output SPL drops below  $L_{th}$ ,  $L_g$  is increased unless  $L_g$  equals the initial prescribed gain. The gain is increased at a rate of  $\alpha$  and reduced at a rate of  $\beta$ . According to the outlined rules, the adaptive compression curve can be adjusted in the shadow area.

2 Experimental Results

2.1 Listeners

Five sensorineural hearing impaired (HI) patients with hearing loss from mild to moderately severe participate in the experiment (2 female, 3 male), whose ages are between 53 and 74 years and all are native Chinese speakers. Audiometric values of different frequencies for the test ears are given in Tab. 1. The column labeled HA indicates hearing-aid usage in the test ear, and the column labeled PTA describes the pure-tone average thresholds at 0.5, 1, 2, and 4 kHz. Each value is measured by the Aurical hearing device made in Denmark.

Tab. 1 Audiometric thresholds for the test ears of 5 HI listeners

ID	Age	Gender	HA	PTA	Frequency/kHz											
					0.125	0.25	0.5	0.75	1	1.5	2	3	4	6	8	
1	53	Female	No	25	10	15	20	15	15	20	25	30	40	55	60	
2	65	Male	No	39	25	20	25	30	30	35	45	50	55	65	70	
3	68	Male	Yes	48	20	25	30	35	40	50	55	60	65	80	90	
4	70	Male	No	50	30	30	35	35	40	50	50	65	70	75	85	
5	74	Female	Yes	58	35	45	50	50	55	60	60	60	65	80	90	

2.2 Simulation results

Test signals are from the speech corpus for the listening test made by Jiangsu Province Hospital. Each signal is sampled at 44.1 kHz. The corpus includes 5 sentence lists and each list consists of 25 sentences spoken by a female talker and a male talker. The test signal in this experiment is from List 1 lasting for 5 s. It is resampled to 16 kHz. The output of each compensation method for the hearing loss of subject 1 is plotted in Fig. 3. In this simulation, parameters  $\alpha$  and  $\beta$  are set to be 2.5 and 5 dB/s, respectively. The NL-R prescription formula is employed in the linear amplification method and the NAL-NL1 prescription formula is employed in the WDRC and the proposed methods. It is noted that the linearly amplified signal shown in Fig. 3(b) is similar to the original signal shown in Fig. 3(a); while the WDRC compensated signal shown in Fig. 3(c) is different from the original signal due to the essential character of linear compensation which amplifies the signals with different levels by the same gain. As can be seen from Fig. 3(d), the output SNR of the proposed method is improved compared to that of the WDRC method due to adaptively adjusting the CR.

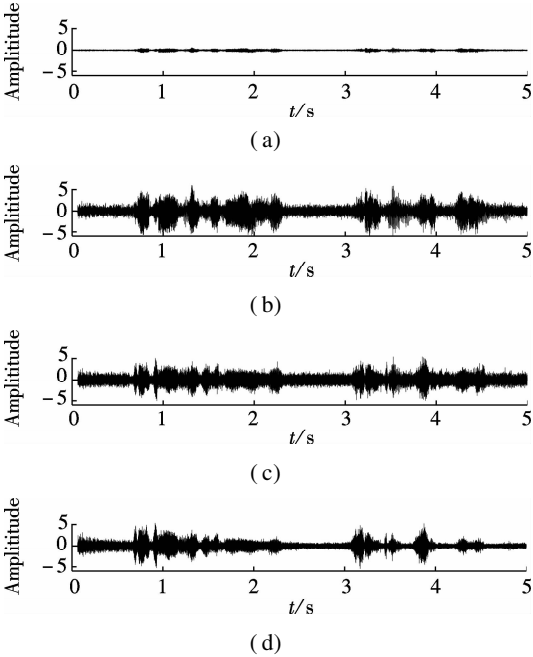


Fig. 3 Comparisons of speech waveforms with different compensation methods. (a) Unprocessed signal; (b) Signal processed by the linear amplification method; (c) Signal processed by the WDRC method; (d) Signal processed by the proposed method

The improved output SNR of the proposed method is given in Tab.2 with different input SPLs and different input SNRs, compared to the WDRC method. The speech and babble noise are combined to form two SNRs with 15 and 5 dB, respectively. In addition to the two SNRs, three different speech intensities are used. The input speech levels are adjusted to three output levels: 55, 65 and 75 dB which denote the soft signal, moderate signal and loud signal in real environments. It can be seen from Tab.2 that the output SNR of our proposed method is improved by at least 1.73 dB compared to the traditional WDRC method. The improvement increases with the increase in input SPL because the method decreases LLG more for high level input speech than that for low level input speech. The improvement is reduced with the decrease in input SNR. The reason is that when the input SNR decreases, the noise SPL increases and the difference between the noise and the noisy speech decreases. Therefore, the improvement is reduced.

Tab.2 Improved output SNR dB			
Input SNR/dB	Input SPL/dB		
	55	65	75
15	3.35	3.18	3.56
5	1.73	2.66	2.97

2.3 Subjective listening test

Test signals come from the same corpus mentioned in Section 2.2. The speech is inputted to the processing system using three different amounts of babble noise: no noise, a SNR of 5 dB and a SNR of 15 dB. Furthermore, three different intensities are used, which are 55, 65 and 75 dB. Prescription formulas and parameters of  $\alpha$  and  $\beta$  are set to be the same as those employed in the above simulation. Hearing-aid simulations are implemented off-line using the Matlab program, which are adjusted according to the individual’s hearing loss.

The speech intelligibility of linear compensation, WDRC compensation and our proposed method are compared in this experiment, which are expressed by the recognition rate of the word. On each intelligibility trial, listeners hear a sentence randomly drawn from the test conditions. The stimuli are presented monaurally via Sennheiser HD650 headphones in an anechonic chamber.

The intelligibility scores averaged over talkers and listeners with different SPLs, SNRs and processing types are shown in Tab.3. The symbol  $\infty$  in Tab.3 means that the input signal is pure speech. Compared with the three amplification schemes, our proposed method produces the highest intelligibility due to the combination of the linear and adaptive WDRC methods. The intelligibility is improved by 6.0% and 5.7%, respectively, on average compared to the linear amplification and WDRC processing methods. In addition, the average intelligibility de-

crease by 2.93% after babble noise is added to the input speech (see Tab.3). After adding more noise, the SNR decreases from 15 to 5 dB, the average intelligibility further decreases by 6.25%. This indicates that the noise affects the system performance significantly.

Tab.3 Speech intelligibility scores averaged over talkers and listeners

Input SNR/ dB	Input SPL/ dB	Linear method	WDRC method	Proposed method
$\infty$	55	80.4	84.7	88.6
	65	90.3	87.5	94.2
	75	87.6	88.9	94.8
15	55	79.7	82.5	87.8
	65	85.6	82.7	90.3
	75	86.5	85.7	89.9
5	55	73.6	76.3	81.8
	65	80.7	77.6	84.4
	75	77.3	78.5	84.2

3 Conclusion

An adaptive multichannel loudness compensation method is proposed for hearing aids to alleviate the conflict between audibility and distortion in the conventional methods. First, the input signal is decomposed into channel signals by gammatone filter banks. Secondly, the linear and WDRC methods are selected to compensate for the loss of loudness according to the dynamic range of the band-passed signal and the HR of the patient. The distortion is further reduced when the compression CR is adaptively adjusted. Experimental evaluation indicates that the output SNR in babble noise is improved by at least 1.73 dB compared to the WDRC compensation method, and the speech intelligibility is improved by 6.0% and 5.7%, respectively, on average compared to the linear and WDRC compensation methods.

References

[1] Walden B E, Surr R K, Cord M T, et al. Comparison of benefits provided by different hearing aid technologies[J]. *Journal of the American Academy of Audiology*, 2000, **11** (10): 540 – 560.

[2] Strelcyk O, Li N, Rodriguez J, et al. Multichannel compression hearing aids: Effect of channel bandwidth on consonant and vowel identification by hearing-impaired listeners[J]. *The Journal of the Acoustical Society of America*, 2013, **133** (3): 1598 – 1606. DOI: 10.1121/1.4789894.

[3] Marriage J E, Moore B C J, Stone M A, et al. Effects of three amplification strategies on speech perception by children with severe and profound hearing loss[J]. *Ear and Hearing*, 2005, **26**(1): 35 – 47.

[4] Kates J M. Understanding compression: modeling the effects of dynamic-range compression in hearing aids[J]. *International Journal of Audiology*, 2010, **49**(6): 395 – 409. DOI: 10.3109/14992020903426256.

[5] Kates J M. A hearing aid with improved compression:

Europe, 12165500. 5[P]. 2013-10-30.

[6] Souza P E, Jenstad L M, Boike K T. Measuring the acoustic effects of compression amplification on speech in noise[J]. *The Journal of the Acoustical Society of America*, 2006, **119**(1): 41 – 44.

[7] Cao L, Li R, Shi Y, et al. Loudness compensation method based on human auditory for digital hearing aids[C]// *2014 7th International Conference on Biomedical Engineering and Informatics (BMEI)*. Dalian, China, 2014: 335 – 340.

[8] Rasetshwane D M, Gorga M P, Neely S T. Signal-processing strategy for restoration of cross-channel suppression in hearing-impaired listeners[J]. *IEEE Transactions on Biomedical Engineering*, 2014, **61**(1): 64 – 75. DOI: 10.1109/TBME.2013.2276351.

[9] Smith J O III, Abel J S. Bark and ERB bilinear transforms[J]. *IEEE Transactions on Speech and Audio Processing*, 1999, **7** ( 6 ): 697 – 708. DOI: 10.1109/89.799695.

[10] Glasberg B R, Moore B C J. Derivation of auditory filter shapes from notched-noise data[J]. *Hearing Research*, 1990, **47**(1): 103 – 138.

[11] Santos J F, Cosentino S, Hazrati O, et al. Objective speech intelligibility measurement for cochlear implant users in complex listening environments[J]. *Speech Communication*, 2013, **55**(7): 815 – 824. DOI: 10.1016/j.specom.2013.04.001.

[12] Biesmans W, Vanthornhout J, Wouters J, et al. Comparison of speech envelope extraction methods for EEG-based auditory attention detection in a cocktail party scenario [C]// *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan, Italy, 2015: 5155 – 5158.

[13] Zhang H. *Hearing aids*[M]. Beijing: People's Medical Publishing House, 2004: 8 – 12. (in Chinese)

[14] Levitt H. Compression amplification[M]// *Compression: From Cochlea to Cochlear Implants*. New York: Springer-Verlag, 2004: 153 – 183.

# 一种自适应多通道响度补偿方法

王 侠<sup>1</sup> 梁瑞宇<sup>2</sup> 王青云<sup>2</sup> 申红明<sup>3</sup> 赵 力<sup>1</sup> 邹采荣<sup>1</sup>

(<sup>1</sup> 东南大学水声信号处理教育部重点实验室, 南京 210096)

(<sup>2</sup> 南京工程学院通信工程学院, 南京 211167)

(<sup>3</sup> 南通大学电子信息学院, 南通 226019)

**摘要:**为减小助听器系统中传统响度补偿方法中语音可听和畸变之间的矛盾,提出了一种自适应的多通道响度补偿方法.该方法根据带通信号的动态范围和患者的听阈交替使用线性方法和动态范围压缩方法.为进一步减小动态范围压缩引起的畸变,并提高噪声环境下的输出信噪比,给出了一种自适应调整压缩比的方法.实验结果表明,与 WDRC 动态范围压缩补偿方法相比,所提方法在语噪声环境下的输出信噪比至少提高了 1.73 dB;与线性和动态范围压缩补偿方法相比,所提方法的平均语音可懂度分别提升了 6.0% 和 5.7%.

**关键词:**响度补偿;动态范围压缩;gammatone 滤波器组;助听器

**中图分类号:**TN912.3