EVALUATION OF GENERALIZED MAXIMUM A POSTERIORI SPECTRAL AMPLITUDE (GMAPA) SPEECH ENHANCEMENT ALGORITHM IN HEARING AIDS

Ying-Hui Lai[1], Yu-Cheng Su[2], Yu Tsao[2], Shuenn-Tsong Young[1]

[1] Institute of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan
[2] Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan

ABSTRACT

Speech enhancement is an important segment in digital hearing aids, which aims to improve signal-to-noise (SNR) level of received speech signals and thus enhance speech intelligibility for hearing-loss individuals. Recently, we proposed a generalized maximum a posteriori spectral amplitude (GMAPA) speech enhancement algorithm. The proposed GMAPA algorithm has been confirmed effective in a series of objective evaluations and speech recognition tests. In this study, we conduct experiments and observe that GMAPA also provides clear long-term SNR increases in a simulated hearing-aids testing condition. The result demonstrates that GMAPA can be suitably applied in digital hearing aids.

Index Terms— Speech Enhancement, Noise Reduction, GMAPA, Digital Hearing Aids, Long-term SNR.

INTRODUCTION

Background noise can seriously damage speech intelligibility for people with hearing loss. Therefore, speech enhancement is usually deployed and plays an important role in digital hearing aids to improve the signal-to-noise (SNR) level of received speech signals [1]. Recently, we proposed a generalized maximum a posteriori spectral amplitude (GMAPA) [2] speech enhancement algorithm. GMAPA adopts an adjustable scale of prior information to compute the gain function for speech enhancement: in high SNR conditions, GMAPA uses a smaller scale to maintain the speech quality; on the other hand, in low SNR conditions, GMAPA can remove noise components effectively by using a larger scale. A mapping function is designed to determine the optimal scale of the prior information according to the SNR of the received speech data. Our previous study has shown that GMAPA provides superior performance to several popular speech enhancement methods in both objective evaluations and speech recognition tests [2].

In this study, we intend to investigate the applicability of GMAPA on digital hearing aids. For the experiments, we combine GMAPA with the wide-dynamic-range compression (WDRC) [3] amplification to simulate the processing of digital hearing aids. Then, we test the mean difference of the long-term SNR between using and not using GMAPA on speech signals, which are contaminated with different noises for the high-frequency hearing loss cases. Experimental results demonstrate that GMAPA significantly improve the long-term SNR score, suggesting that GMAPA is suitable to be integrated into digital hearing aids.

THE GMAPA SPEECH ENHANCMENT ALGORITHM

This section introduces the proposed GMAPA algorithm and the mapping function to determine the scale of prior information.

Signal Analysis

In the time domain, a noisy speech signal, $y[n]$, is composed of a clean speech, $s[n]$, corrupted by additive a noise signal, $v[n]$, as

$$y[n] = s[n] + v[n], \tag{1}$$

where n denotes the time index. In the frequency domain, the noisy speech spectrum, $Y[m,l]$, can be expressed as

$$Y[m,l] = S[m,l] + V[m,l], 0 \le l \le L-1\,, \tag{2}$$

where l is the frequency bin correspond to the frequency ω_l, where $\omega_l = \frac{2\pi l}{L}$, $l = 0,1,\dots,L-1$; m is the frame index; $S[m,l]$ and $V[m,l]$ are speech and noise spectrums, respectively.

Fig. 1 shows the overall speech enhancement system, which can be decomposed to two parts: noise tracking and gain estimation. The noise tracking determines noise information from the noisy speech. Then, the gain estimation calculates a gain function, $G[m,l]$, based on the estimated noise information, to reconstruct the speech, $\hat{S}[m,l]$, by filtering $Y[m,l]$ through $G[m,l]$. In the following discussion, we denote $Y[m,l]$, $S[m,l]$, $V[m,l]$, and $G[m,l]$, respectively, as Y, S, V, and G, for simplicity.

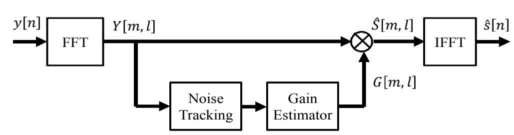

Fig. 1. Block diagram of a speech enhancement system.

GMAPA Algorithm

To perform speech enhancement, we first decompose Y and S into spectral amplitude and phase parts:

$$Y = Y_k exp(j\theta_{Y_k})\,, \tag{3}$$

$$S = S_k exp(j\theta_{S_k})\,, \tag{4}$$

where $Y_k = |Y|$, $S_k = |S|$, $\theta_{Y_k} = \angle Y$, and $\theta_{S_k} = \angle S$.

For GMAPA [2], the spectral amplitude, $\hat{S}_k$, is calculated by

$$\hat{S}_k = \arg\max_{S_k} J_{GMAPA}(S_k), \tag{5}$$

where $J_{GMAPA}(S_k)$ is the cost function and can be expressed as

$$J_{GMAPA}(S_k) = \ln\{p[Y|S_k]\;(p[S_k])^{\alpha}\}. \tag{6}$$

By differentiating $J_{GMAPA}(S_k)$ in Eq. (6) with respect to S_k and equating the result to zero, we can obtain

$$\hat{S}_k = \frac{\xi + \sqrt{\xi^2 + (2\alpha-1)(\alpha+\xi)\xi/\gamma}}{2(\alpha+\xi)} Y_k, \tag{7}$$

where $\hat{S}_k$ is the enhanced speech, $\xi = \sigma_s^2/\sigma_v^2$ and $\gamma = Y_k^2/\sigma_v^2$ are the *a priori* and *a posteriori* SNRs. Thus, GMAPA gain function is

$$G_{GMAPA} = \frac{\xi + \sqrt{\xi^2 + (2\alpha - 1)(\alpha + \xi)\xi/\gamma}}{2(\alpha + \xi)}. \quad (8)$$

With the estimated spectral amplitude, S_k, we follow the suggestion from [4, 5] to estimate the phase, θ_{S_k}, of S:

$$exp(j\hat{\theta}_{S_k}) = exp(j\theta_{Y_k}). \quad (9)$$

Finally, the clean speech spectrum can be obtained as

$$\hat{S} = \hat{S}_k exp(j\theta_{Y_k}). \quad (10)$$

Determining the Scale of Prior Information

GMAPA adopts a sigmoid function [6] to determine optimal α:

$$\alpha = \frac{\alpha_{max}}{1 + exp[-b(\bar{\gamma} - c)]}, \quad (11)$$

where α_{max} is the maximum value for α; b and c are coefficients of the sigmoid function; $\bar{\gamma}$ is the mean of *a posteriori* SNR. Fig. 2 shows the relationship between the scale α and the $\bar{\gamma}$. We can optimally determine α_{max}, b and c based on a set of development data.

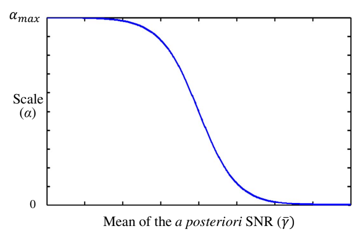

Fig. 2. The relationship between α and $\bar{\gamma}$ of the mapping function.

EXPERIMENT

Experimental Setup

Figure 3 shows the block diagram of the experiment framework. The evaluation measurement is the long-term SNR scores between using and not using GMAPA for the same hearing-loss audiogram simulated by a platform, which was developed using the LabVIEW software. In this study, the same set of speech data was used throughout all measurements; it was concatenated Mandarin utterances lasting for 10 seconds, along with three noise signals: range-hood, babble, and traffic noises, which were also concatenated for 10 seconds. The sound intensity of each signal was normalized to 65 dB SPL. Additionally, speech and noise were combined at four SNR levels (-2, +2, +6, or +10 dB) to simulate typical daily situations. The combined signals were processed by the GMAPA algorithm and the WDRC amplification. Here the National Acoustic Laboratories-Nonlinear Version 1 prescription [7] was used for the WDRC amplification process. The separation technique of the long-term SNR developed by Hagerman and Olofsson [8] was used to extract the speech and noise components.

Results and Discussion

Table 1 shows the obtained long-term SNR scores, where Y^a and N^b, respectively, represent using and not using GMAPA. From Table 1, the overall mean±SD values of the output long-term SNR were 11.0±5.5 dB and 3.7±3.6 dB for using and not using GMAPA, respectively, across the four SNR levels. A paired-sampled *t*-test was performed and showed that the difference between using and not using GMAPA was significant (t=7.2, p<0.001). The results confirmed that GMAPA can bring huge gain on speech intelligibility in the simulated hearing-aids noisy conditions.

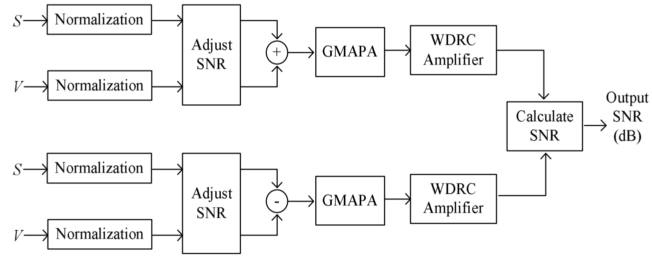

Fig. 3. Block diagram of the overall processing to obtain the output SNR, where S and V denote the speech and noise, respectively.

Table 1. Mean difference of the long-term SNR (in dB).

Input SNR (dB)	Range-hood noise		Babble noise		Traffic noise	
	Y^a	N^b	Y^a	N^b	Y^a	N^b
-2	8.9	-0.4	0.4	-1.6	9.6	-0.3
+2	11.9	2.7	3.7	1.3	12.9	2.7
+6	14.9	5.7	7.0	4.2	15.9	5.6
+10	17.7	9.1	10.2	7.2	18.9	8.7
Mean	**13.4**	**4.3**	**5.3**	**2.8**	**14.3**	**4.2**

Y^a: using GAMAP; N^b: not using GMAPA.

CONCLUSION

This study investigates the applicability of the GMAPA speech enhancement algorithm for hearing aids. Experimental results show that GMAPA can achieve clear improvements on long-term SNR evaluations when tested with noisy speech contaminated by three types of noise, at four SNR levels (-2, 2, 6, and 10 dB), suggesting that GMAPA can benefit speech intelligibility for hearing-aid users in noisy environments. The clinical trial of real listening tests will be conducted in the near future.

REFERENCE

[1] J. M. Kates, *Digital hearing aids, Chapter 8.*: Cambridge Univ Press, 2008.

[2] Y. Tsao, Y. C. Su, J. E. Wu, and F. R. Jean, "Speech enhancement using generalized maximum a posteriori spectral amplitude estimator," in Proc. *ICASSP,* 2013.

[3] T. Venema., *Compression for clinicians, 2nd Edition*: Singular Publishing Group, 2006.

[4] J. Chen, *Fundamentals of Noise Reduction in Spring Handbook of Speech Processing*: Springer, 2008.

[5] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 32, pp. 1109-1121, 1984.

[6] M. Zhang, S. Vassiliadis and J. G. Delgado-Frias, "Sigmoid generators for neural computing using piecewise approximations," *IEEE Transactions on Computers*, vol. 45, pp. 1045-1049, 1996.

[7] H. Dillon, *Hearing aids, Chapter 6*: Thieme, 2001.

[8] B. Hagerman and A. Olofsson, "A method to measure the effect of noise reduction algorithms using simultaneous speech and noise," *Acta Acustica united with Acustica,* vol. 90, pp. 356-361, 2004.