A DISCRIMINATIVE POST-FILTER FOR SPEECH ENHANCEMENT IN HEARING AIDS

Ying-Hui Lai[1], Syu-Siang Wang[1], Pei-Chun Li[2], and Yu Tsao[1]

Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan
Department of Audiology and Speech-Language Pathology, Mackay Medical College, New Taipei City, Taiwan

{jackylai, sypdbhee}@citi.sinica.edu.tw; ankh_li@mmc.edu.tw; yu.tsao@citi.sinica.edu.tw

ABSTRACT

For hearing aid (HA) devices, speech enhancement (SE) is an essential unit aiming to improve signal-to-noise ratio (SNR) and quality of speech signals. Previous studies, however, indicated that user experience with current HAs was not fully satisfactory in noisy environments, suggesting that there is still room for improvement of SE in HA devices. This study proposes a novel discriminative post-filter (DPF) approach to further enhance the SNR and quality of SE processed speech signals. The DPF uses a filter to increase the energy contrast (discrimination) of speech and noise segments in a noisy utterance. In this way, SNR and sound quality of speech signals can be improved, and annoying musical noises can be suppressed. To verify the effectiveness of DPF, the present study integrates DPF with a previously proposed generalized maximum *a posteriori* spectral amplitude estimation (GMAPA) SE method. Experimental results demonstrated that when comparing to GMAPA alone, this integration can further improve output SNR and perceptual evaluation of speech quality (PESQ) scores and effectively suppress musical noises across various noisy conditions. Due to its low-complexity, low-latency, and high-performance, DPF can be suitably integrated in HA devices, where computational efficiency, power consumption, and effectiveness are major considerations.

Index Terms— hearing aid, speech enhancement, discriminative post-filter (DFP), GMAPA algorithm, spectral restoration

1. INTRODUCTION

Sensorineural hearing loss (SNHL) is a common type of hearing loss in clinical, and hearing aid (HA) is the most popular method to help SNHL individuals improve their communication ability [1, 2]. In an HA device, speech enhancement (SE) is an essential unit, which aims to improve signal-to-noise ratio (SNR) and sound quality of speech signals. However, previous studies indicated that degraded speech quality caused by noise remains a key issue to HA, even when the SE approach is used [3, 4]. This issue directly affects the satisfaction of HA users and suggests that there is still room for improvement of SE in HA devices.

SE approaches can be roughly divided into two classes: supervised and unsupervised approaches. Supervised SE approaches prepare an acoustic structure using the prior information of noise type, signal-to-noise ratio (SNR), and/or speaker identity, to facilitate the online enhancement process. Nonnegative matrix factorization (NMF) [5], sparse coding [6], deep autoencoder [7] [8], and deep neural network [9] are successful supervised approaches. Although the supervised algorithms can provide satisfactory noise reduction performance, suitable prior information is required to achieve the optimal performance. Such information is usually inaccessible beforehand.

On the other hand, unsupervised SE approaches are more favorable for the tasks where the prior information about the acoustic condition is not available. Among the unsupervised SE approaches, spectral restoration is a notable class. Spectral restoration estimates a gain function for performing noise reductions in the spectral domain in order to obtain clean speech spectra from noisy speech spectra. Well-known spectral restoration approaches include minimum mean-square error (MMSE) spectral estimator [8], maximum *a posteriori* spectral amplitude (MAPA) estimator [9, 10], and maximum likelihood spectral amplitude (MLSA) estimator [11]. More recently, a generalized maximum *a posteriori* spectral amplitude estimation (GMAPA) algorithm was proposed for spectral restoration [12, 13]. The GMAPA algorithm optimally specifies the prior information scale according to the SNR of degraded speech to calculate the gain function and has been confirmed to perform well in both high and low SNR conditions. Previous studies also verified that the GMAPA algorithm provides efficacious noise reduction performance while causing small distortions when compared to other related SE methods (i.e., MMSE, MAPA, and MLSA) [12]. Meanwhile, it has been verified that the GMAPA algorithm provides significant benefits when integrated in an wide-dynamic range compression (WDRC) or adaptive WDRC [14] amplification scheme of HA device [13, 15].

Although SE approaches can effectively improve the SNR of input speech signals, another issue remains unsolved: musical tones appear during silence segments [16]. The annoying musical noises seriously degrade the satisfaction of HA users. This study proposes a discriminative post-filter (DPF), which is placed after an SE process to further improve SNR and sound quality of enhanced speech signals. The proposed DPF approach devices a non-linear function, which specifies a smaller gain for soft sounds (i.e., noise segments) and a larger gain for loud sounds (i.e., speech segments) [17]. Because an SE process can already provide improved SNR, the DPF approach can further enhance the energy contrast (discrimination) of speech and noise segments in a noisy utterance. A larger energy contrast (discrimination) of speech and noise segments correspond to the processed speech signals with higher output SNR and better sound quality. Moreover, by suppressing low SNR segments, the annoying musical noises can be suppressed effectively. We evaluated the proposed DPF approach with objective evaluations using PESQ [18] and long-term SNR [19]. Evaluation results confirm that the DPF approach can be suitably integrated with the GMAPA algorithm to further enhance output SNR and sound quality.

2. RELATED WORKS

This section reviews the spectral restoration framework and the GMAPA algorithm for speech enhancement.

2.1. Spectral restoration framework

In the time domain, a noisy speech signal, $y[n]$, is the sum of a clean speech, $s[n]$, and a noise signal, $v[n]$:

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

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

$$\begin{gathered} Y[m,l] = S[m,l] + V[m,l], \\ \text{where } 1 \leq m \leq M;\ 0 \leq l \leq L-1, \end{gathered} \tag{2}$$

where l is the frequency bin for ω_l, where $\omega_l = 2\pi l/L$; $S[m,l]$ and $V[m,l]$ are the speech and noise spectra, respectively; *L* and *M* are the number of frequency bins and frames, respectively.

Figure 1 shows the spectral restoration process, which can be divided into noise tracking and gain estimation stages. The noise tracking stage computes noise power from the noisy speech, $Y[m,l]$, to obtain *a priori* SNR ξ_k and *a posteriori* SNR γ_k, which are defined as $\xi_k = \sigma_s^2/\sigma_v^2$ and $\gamma_k = Y_k^2/\sigma_v^2$, where $\sigma_s^2 = E[|S|^2]$ and $\sigma_v^2 = E[|V|^2]$; the subscript *k* denotes the amplitude part of a signal. In the following, ξ_k and γ_k are denoted as ξ and γ, respectively, for simplicity. Then, the gain estimation stage calculates a gain function, $G[m,l]$, based on *a priori* and *a posteriori* SNR statistics, to obtain enhanced speech, $\hat{S}[m,l]$, by filtering $Y[m,l]$ through $G[m,l]$. In the following, we denote $Y[m,l]$, $S[m,l]$, $V[m,l]$, and $G[m,l]$, respectively, as Y, S, V, and G, for simplicity. Finally, after the IFFT process, we then obtain the enhanced speech, $\hat{s}[n]$.

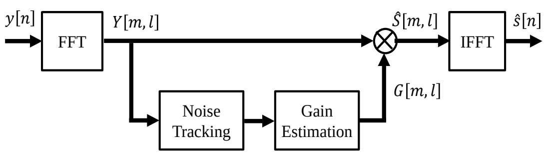

Figure 1: *Block diagram of a spectral restoration system.*

By decomposing noisy and clean speech specta, Y and S in (2), into amplitude and phase parts, we have

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

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

where $Y_k = |Y|$, $S_k = |S|$, $\theta_Y = \angle Y$, and $\theta_S = \angle S$. To restore S from Y, we first estimate the phase of clean speech spectrum by [10]:

$$exp\big(j\hat{\theta}_S\big) = exp(j\theta_Y). \tag{5}$$

Full details about the phase estimation can be found in [8]. Accordingly, the clean speech spectrum is estimated as

$$\hat{S} = \hat{S}_k exp(j\theta_Y), \tag{6}$$

where $\hat{S}_k$ is the enhanced speech amplitude.

2.2. The GMAPA algorithm

Recently, the GMAPA spectral restoration algorithm [12] has been proposed to estimate the spectral amplitude, $\hat{S}_k$ by:

$$\hat{S}_k = \arg\max_{S_k} J_{GMAPA}(S_k), \tag{7}$$

where $J_{GMAPA}(S_k)$ is the GMAPA cost function:

$$J_{GMAPA}(S_k) = \ln\{p[Y|S_k]\ (p[S_k])^{\alpha}\}, \tag{8}$$

where α is the prior information scalar, which is optimally determined according to the SNR of the testing condition. A sigmoid function was used to optimally determine the scale of α for G_{GMAPA} in Eq. (8) for each utterance according to

$$\alpha = \frac{\alpha_{max}}{1 + exp[-b(\bar{\gamma} - c)]}\,, \tag{9}$$

where α_{max} is the maximum value (upper bound) for α, *b* and *c* are coefficients of the sigmoid function, and $\bar{\gamma}$ is the mean of the *a posteriori* SNR for a given utterance; $\bar{\gamma} = 1/T\sum_{t=1}^{T}\gamma^{(t)}$, where *T* is the total number of frames for this utterance, and $\gamma^{(t)}$is the mean *a posteriori* SNR for the *t*-th frame. In our previous study, the parameter set $\{\alpha_{max}, b, c\}$ is determined by a set of training data from Aurora-4 [20], and the method to determine $\{\alpha_{max}, b, c\}$ is based on a quality objective metric: speech distortion index (SDI) [12].

After some derivations, the GMAPA-based gain function, G_{GMAPA}, can be expressed as

$$G_{GMAPA} = \frac{\xi + \sqrt{\xi^2 + (2\alpha - 1)(\alpha + \xi)\xi/\gamma}}{2(\alpha + \xi)}. \tag{10}$$

Finally, the estimated clean speech spectrum for the GMAPA estimator can be written as

$$\begin{aligned} \hat{S} &= \hat{S}_k \cdot exp\big(j\theta_{Y_k}\big) \\ &= G_{GMAPA} \cdot Y. \end{aligned} \tag{11}$$

3. DISCRIMINATIVE POST-FILTER (DPF)

This section presents the proposed DPF algorithm and introduces the integration of DPF and GMAPA, termed GMAPA+DPF in this study.

3.1. The proposed DPF algorithm

Figure 2 shows an input/output (I/O) function of DPF in the dB sound pressure level (SPL) domain.. The area below the knee point (T) corresponds to the expansion area; an expansion ratio (ER) is applied to the input signals. Meanwhile, the area above the knee point (T) corresponds to the linear area; a fixed gain (or zero gain) is applied to the input signals. More specifically, in the linear area, when the input signal is changed by 5 dB, the output signal is also changed by 5 dB. In the expansion area, when the input signal is changed by 5 dB, the output signal is changed by 10 dB. If the knee point is set properly (i.e., exactly separating the levels of speech and noise segments), DPF will effectively increase the energy contrasts. Furthermore, because the linear amplification is used when the level of input signals is higher than the knee point, less speech distortions will be incurred by DPF. The DPF algorithm designs the filter, $F(\cdot)$, by:

$$\hat{S}'[m,l] = F(\hat{S}[m,l]) = \frac{\sum_{m=1}^{M}\sum_{l=1}^{L}(\hat{S}[m,l])}{\sum_{m=1}^{M}\sum_{l=1}^{L}(\rho[m,l]\hat{S}[m,l])}\rho[m,l]\hat{S}[m,l]\,, \quad (12)$$

where $\hat{S}[m,l]$ is the SE enhanced speech signal from Eq. (6); $\hat{S}'[m,l]$ is the output speech signals processed by SE and DPF; the denominator term in Eq. (12) normalizes the output speech signals to the same power level as that of $\hat{S}[m,l]$; $\rho[m,l]^{(c)}$ is the poster filter coefficient, which is estimated by

$$\rho[m,l]^{(c)} = \begin{cases} 1 & , \quad if\ p_m^{(c)} > T^{(c)} \\ 10^{\{[(1-ER)\times(T^{(c)}-p_m^{(c)})]/20\}} & , \quad if\ 0 < p_m^{(c)} < T^{(c)} \end{cases}$$
$$\text{for } l \in \{f_L^{(c)}, f_H^{(c)}\} \quad (13)$$

where $f_L^{(c)}$ and $f_H^{(c)}$ denote the lower- and higher- bounds, respectively, of the *c*-th frequency channel (each frequency channel covers a band of frequency bins); $p_m^{(c)}$ denotes the level of input signals (in dB SPL) assigned to the *m*-th frame; $T^{(c)}$ is the knee point of expansion of *c*-th frequency channel; the parameter ER represents the expansion ratio. Based on the poster filter coefficient in Eq. (13), the I/O function of DPF in dB SPL domain becomes Figure 2.

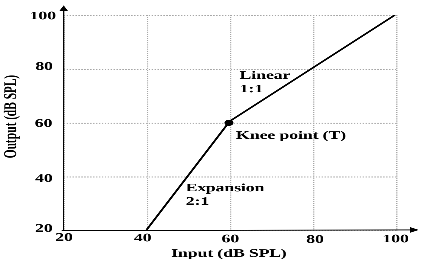

Figure 2: *The I/O function of DPF; the area below the knee point (T) stands for the expansion area (ER=2), and the area above the knee point (T) stands for the linear gain area.*

3.2. The integration of GMAPA and DPF

This study integrates the DPF approach with the GMAPA algorithm by following the procedure demonstrated in Figure 3. Comparing to the original spectral restoration system presented in Figure 1, the DPF approach is placed after the GMAPA spectral restoration process. From Eqs. (11) and (12), the final output spectra of GMAPA+DPF becomes:

$$\hat{S}'[m,l] = F(G[m,l]\cdot Y[m,l]), \quad (14)$$

where $F(\cdot)$ is the filter function defined in Eqs. (12) and (13). Finally, by performing IFFT on $\hat{S}'[m,l]$, we can obtain the enhanced speech, $\hat{s}'[n]$.

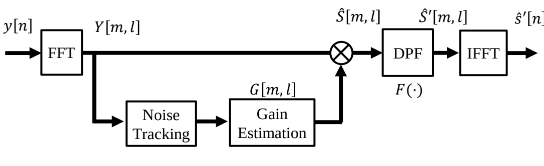

Figure 3: *The proposed GMAPA+DPF framework.*

4. EXPERIMENT

This section presents the experimental setup and results of the proposed GMAPA+DPF approach.

4.1. Test signals

The Mandarin sentences chosen from Mandarin hearing in noise test (MHINT) database [21] were used to prepare the testing speech sentences, denoted as "S". The pink noise were used and denoted as "N" in this study. To prepare noisy test set, two steps were implemented. First, the root-mean-square sound intensity of each signal was normalized to 65 dB SPL, corresponding to a moderate speech level. Next, "S" and "N" were combined at five SNR values (-10, -5, 0, 5, and 10 dB). The "S" and "N" signals were adjusted simultaneously by the same absolute amounts to produce different SNRs. For example, when the input SNR was increased by 10 dB, "S" was increased by 5 dB, and "N" was decreased by 5 dB. Finally, the combined sounds were normalized to 65 dB SPL.

4.2. Measurement procedure

We performed objective evaluations to measure the performance of SNR and sound quality of GMAPA+DPF. In addition, we intend to observe the correlations of ER values of DPF algorithm with output SNR and sound quality performances. The results of the original noisy signals, and GMAPA were tested to compare with GMAPA+DPF. In this study, a single channel DPF (*c*=1; $f_L^{(c)}$=0 Hz, $f_H^{(c)}$=8k Hz in Eq. (13)) was used. The knee point (T) of DPF was set according to the mean level of each input noisy signals. For example, if the mean level of input signal is 60 dB SPL, then T will be set to 60 in DPF. In addition, the parameters of α_{max}, b, and *c* in GMAPA algorithm were set to 2.0, -1.0 and 11, respectively.

In this study, the ER values in GMAPA+DPF were set to 2, 3, 4, and 5, respectively, which were denoted as G+DPF(2), G+DPF(3), G+DPF(4), and G+DPF(5) in the following discussion. In addition, the results of GMAPA algorithm was also presented for comparison. For each approach, 150 testing utterances (30 Mandarin sentences × 5 SNR levels) were used to compare the output long-term SNR and sound quality performance.

4.3. Methods of objective evaluation

In this study we used the separation technique of the long-term SNR [19] and perceptual evaluation of speech quality (PESQ) [18, 22] measurement to compare performances. The separation technique of long-term SNR was developed by Hagerman and Olofsson [19] that was used to evaluate the output SNR performance by non-linear processing (e.g., compression amplification scheme of HAs [14], and speech enhancement [13]). It was used to extract the speech and noise components from processed sounds. A higher output SNR value corresponds to a better SNR performance. Next, the PESQ algorithm was used to evaluate the sound quality performance. The PESQ algorithm is the current industry standard for the objective prediction of one-way speech quality, and was accepted in February 2001 by the Telecommunication Standardization Sector of the International Telecommunication Union (ITU-T) as objective speech quality measurement standard P.862 [18]. The PESQ algorithm uses two input signals to compute the speech quality (i.e., the unprocessed speech sample and processed speech), and score ranges from –0.5 (worst) to 4.5 (best). More details of the long-term SNR and PESQ measurements can be found in [19, 22, 23].

4.4. Experiment results

4.4.1. Objective evaluation

Table I and II show the results of long-term SNR and PESQ, respectively, of original noisy, GMAPA, G+DPF(2), G+DPF(3), G+DPF(4), and G+DPF(5). For the long-term SNR results, a higher value indicates better SNR performance. From Table I, the proposed G+DPF (i.e., ER from 2 to 5) algorithms achieved higher SNR values than original noisy and GMAPA for all of the test conditions, suggesting that GMAPA+DPF can provides better speech intelligibility in different noisy conditions [14, 24]. Furthermore, it is noted that using a larger ER can obtain a higher output SNR value, especially in low SNR conditions (-10 dB to 0 dB).

For the PESQ results, a higher PESQ score represents better sound quality [18] in Table II. The results in Table II show that G+DPF (i.e., ER from 2 to 5) outperformed original noisy and GMAPA alone in various test conditions. Moreover, the results indicated that the sound quality is decreased when increasing the ER values. The reason might result from that a higher ER value can cause distortions easily for speech segments than a lower ER. Accordingly, a higher ER value may deteriorate the performance of sound quality. In addition, from the above results, we observe a trade-off between output SNR and sound quality performance when deciding a suitable ER value in the DPF algorithm.

Table I. *Output SNR values for original noisy and five approaches in different SNRs.*

	-10 dB	-5 dB	0 dB	5 dB	10 dB
Original	-10	-5	0	4.99	9.98
GMAPA	0.40	8.83	13.60	17.03	20.29
G+DPF(2)	1.82	9.80	14.09	17.30	20.49
G+DPF(3)	2.07	9.96	14.15	17.33	20.50
G+DPF(4)	2.22	10.03	14.16	**17.34**	**20.51**
G+DPF(5)	**2.31**	**10.07**	**14.17**	**17.34**	**20.51**

Table II. *PESQ values for original noisy and five algorithms in different SNRs.*

	-10 dB	-5 dB	0 dB	5 dB	10 dB
Original	0.99	1.07	1.20	1.40	1.69
GMAPA	1.19	1.16	1.62	1.91	2.09
G+DPF(2)	**1.90**	**1.96**	**2.24**	**2.48**	**2.79**
G+DPF(3)	1.80	1.93	2.13	2.36	2.63
G+DPF(4)	1.77	1.87	2.05	2.26	2.54
G+DPF(5)	1.74	1.81	1.98	2.19	2.50

4.4.2. Spectrogram analysis

A spectrogram shows the spectral representations of a time-varying signal and is often used to analyze frequency and level properties of speech signals [25]. Figure 4 illustrates six sub-figures, showing the spectrograms of: (a) original noisy speech at 0 dB SNR; (b) GMAPA, (c) G+DPF(2), (d) G+DPF(3), (e) G+DPF(4), and (f) G+DPF(5). The original noisy spectrogram was collected from a male voice in Mandarin, saying "*He can play table tennis very well*". It can be noted that sub-figures (c) to (f), namely G+DPF(2) to G+DPF(5), can provide higher SNR performances than sub-figures (a) and (b), namely the original speech and the GMAPA algorithms alone. The sub-figures (c) to (f) also show that the proposed DPF remove the residual noise in the silence segments, and thus the annoying musical tones are effectively suppressed. Real subject tests will be conducted in our future study to quantify this advantage of DPF.

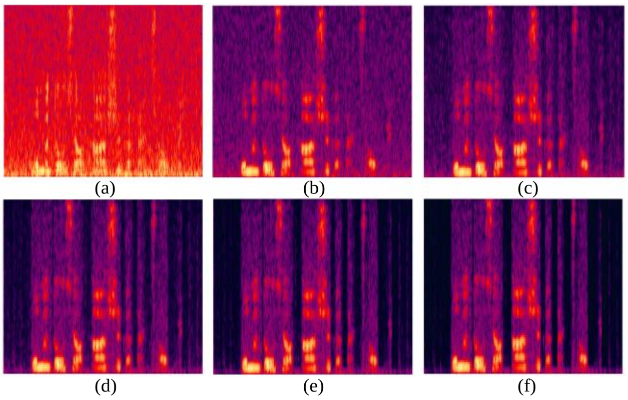

Figure 4: *Spectra of (a) original 0 dB SNR noisy speech; (b) GMAPA; (c) G+DPF(2); (d) G+DPF(3); (e) G+DPF(4); (f) G+DPF(5).(For the demo files please refer to [1]).*

5. DISCUSSION AND CONCLUSION

This paper proposes a novel DPF approach to further improve the SNR and sound quality of enhanced speech signals. The proposed DPF approach uses a non-linear function to enhance the energy contrast (discrimination) between speech and noise segments in a noisy utterance. To verify the proposed DPF approach, this study integrates the proposed DPF approach with the GMAPA speech enhancement algorithm. The integration approach is termed GMAPA+DPF. The objective evaluation results from output SNR and PESQ confirmed that GMAPA+DPF outperformed GMAPA alone, verifying that the DPF approach can further improve enhanced signals to achieve higher SNR and better sound quality. Moreover, musical noises are effectively suppressed. Because of its efficient computation, low latency, and satisfactory sound quality, the DPF approach can be suitably integrated in HA devices.

The experimental results indicate that a higher expansion rate (ER) enables DPF to achieve higher output SNR values while decreased PESQ scores. The results suggested that there is a trade-off between SNR and sound quality. In the future, we will further incorporate the adaptive adjustment scheme to optimize the parameters of knee point (T) and ER to achieve maximal benefits for specific tasks. Meanwhile, we will explore the correlation of T and ER with intelligibility of processed speech signals.

6. ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Technology of Taiwan under Project MOST103- 2221-E-001-003.

[1] http://goo.gl/4PDGk6

7. REFERENCES

[1] H. Dillon, *Hearing aids (second edition),* Thieme, New York, 2012.

[2] S. Gatehouse, G. Naylor, and C. Elberling, “Benefits from hearing aids in relation to the interaction between the user and the environment,” *International Journal of Audiology,* vol. 42, pp. 77-85, 2003.

[3] S. Kochkin, “MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing,” *The Hearing Journal,* vol. 63, pp. 19-20, 2010.

[4] S. Kochkin, “Consumers rate improvements sought in hearing instruments,” *Hearing Review,* vol. 9, pp. 18-22, 2002.

[5] K. W. Wilson, B. Raj, P. Smaragdis, and A. Divakaran, “Speech denoising using nonnegative matrix factorization with priors. ” in *Proc. ICASSP 2008*, pp. 4029-4032.

[6] C. D. Sigg, T. Dikk, and J. M. Buhmann, “Speech enhancement using generative dictionary learning,” *IEEE Transactions on Audio, Speech, and Language Processing,* , vol. 20, pp. 1698-1712, 2012.

[7] X. Lu, Y. Tsao, S. Matsuda, and C. Hori, “Speech enhancement based on deep denoising autoencoder,” in *Proc. Interspeech 2013, pp.* 2013.

[8] X. Lu, Y. Tsao, S. Matsuda, and C. Hori, “Ensemble modeling of denoising autoencoder for speech spectrum restoration, ” in *Proc. Interspeech 2014*, pp. 885-889,

[9] Y. Xu, J. Du, L. Dai, C.-H. Lee, “An experimental study on speech enhancement based on deep neural networks,” *IEEE Signal Processing Letters*, vol. 21, pp. 65-68, 2014.

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

[11] T. Lotter, and P. Vary, “Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model,” *EURASIP Journal on Applied Signal Processing,* vol. 2005, pp. 1110-1126, 2005.

[12] B. Fodor, and T. Fingscheidt, ”Speech enhancement using a joint map estimator with Gaussian mixture model for (non-) stationary noise. ” in *Proc. ICASSP 2012*, pp. 4768-4771.

[13] P. Scalart, "Speech enhancement based on a priori signal to noise estimation." in *Proc. ICASSP 1996*, pp. 629-632.

[14] Y. C. Su, Y. Tsao, J. E. Wu, and F. R. Jean, “Speech enhancement using generalized maximum a posteriori spectral amplitude estimator,” in *Proc. ICASSP* 2013, pp. 7467 - 7471, 2013.

[15] Y.-H. Lai, Y.-C. Su, Y. Tsao, and S.-T. Young, "Evaluation of generalized maximum a posteriori spectral amplitude (GMAPA) speech enhancement algorithm in hearing aids." in Proc. ISCE, pp. 245-246, 2013.

[16] M. Li, H. G. McAllister, N. D. Black, and T. A. De Perez, “Perceptual time-frequency subtraction algorithm for noise reduction in hearing aids,” *IEEE Transactions on Biomedical Engineering, ,* vol. 48, pp. 979-988, 2001.

[17] Y.-H. Lai, P.-C. Li, K.-S. Tsai, W.-C. Chu, and S.-T. Young, “Measuring the Long-Term SNRs of Static and Adaptive Compression Amplification Techniques for Speech in Noise,” *Journal of the American Academy of Audiology,* vol. 24, pp. 671-683, 2013.

[18] Y.-H. Lai, Y. Tsao, and F. Chen, “A Study of Adaptive WDRC in Hearing Aids under Noisy Conditions,” *International Journal of Speech & Language Pathology and Audiology,* vol. 1, pp. 43-51, 2013.

[19] H. J. Steeneken, and T. Houtgast, “A physical method for measuring speech-transmission quality,” *The Journal of the Acoustical Society of America,* vol. 67, pp. 318-326, 1980.

[20] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, “Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs.” in *Proc. ICASSP* 2001, pp. 749-752.

[21] 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.

[22] D. Pearce, “Aurora Working Group: DSR Front End LVCSR Evaluation AU/384/02,” Mississippi State University, 2002.

[23] L. L. Wong, S. D. Soli, S. Liu, N. Han, and M.-W. Huang, “Development of the Mandarin hearing in noise test (MHINT),” *Ear and hearing,* vol. 28, pp. 70-74, 2007.

[24] J. G. Beerends, A. P. Hekstra, A. W. Rix, and M. P. Hollier, “Perceptual evaluation of speech quality (pesq) the new itu standard for end-to-end speech quality assessment part ii: psychoacoustic model,” *Journal of the Audio Engineering Society*, vol. 50, pp. 765-778, 2002.

[25] H. Sun, L. Shue, and J. Chen, “Investigations into the relationship between measurable speech quality and speech recognition rate for telephony speech. ” in Proc. ICASSP 2004, pp. 865-868.

[26] G. Naylor, and R. B. Johannesson, “Long-term signal-to-noise ratio at the input and output of amplitude-compression systems,” *Journal of the American Academy of Audiology,* vol. 20, pp. 161-171, 2009.

[27] S. Haykin, *Advances in spectrum analysis and array processing (vol. III),* Prentice-Hall, Inc., Canada, 1995.